

UNIVERSIDADE FEDERAL FLUMINENSE

Flávio Miranda de Farias

**Autoria Imersiva em Realidade Virtual para
Aplicações Multimeídia Interativas**

NITERÓI

2023

UNIVERSIDADE FEDERAL FLUMINENSE

Flávio Miranda de Farias

Autoria Imersiva em Realidade Virtual para Aplicações Multemídia Interativas

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Computação da Universidade Federal Fluminense como requisito para a obtenção do Grau de Doutor em Computação. Área de concentração: Ciência da Computação.

Orientadora:
Débora Christina Muchaluat Saade

NITERÓI

2023

Ficha catalográfica automática - SDC/BEE
Gerada com informações fornecidas pelo autor

F224a Farias, Flávio Miranda de
Autoria Imersiva em Realidade Virtual para Aplicações
Mulsemídia Interativas / Flávio Miranda de Farias. - 2023.
182 f.: il.

Orientador: Débora Christina Muchaluat Saade.
Tese (doutorado)-Universidade Federal Fluminense, Instituto
de Computação, Niterói, 2023.

1. Autoria 360 graus. 2. Cenários interativos. 3. Realidade
Virtual. 4. Mulsemídia,. 5. Produção intelectual. I. Saade,
Débora Christina Muchaluat, orientadora. II. Universidade
Federal Fluminense. Instituto de Computação.III. Título.

CDD - XXX

FLÁVIO MIRANDA DE FARIAS

Autoria Imersiva em Realidade Virtual para Aplicações Mulsemídia Interativas

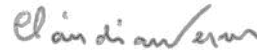
Tese de Doutorado apresentada ao Programa de Pós-Graduação em Computação da Universidade Federal Fluminense como requisito parcial para a obtenção do Grau de Doutor em Computação. Área de concentração: Ciência da Computação.

Aprovada em abril de 2023.

BANCA EXAMINADORA



Prof^ª. Débora Christina Muchaluat Saade, Doutora. –
Orientadora, UFF



Prof^ª. Claudia Maria Lima Werner, Doutora. – UFRJ



Prof. Esteban Walter Gonzalez Clua, Doutor. – UFF



Prof. Joel André Ferreira dos Santos, Doutor. – CEFET/RJ



Prof^ª. Luciana Cardoso de Castro Salgado, Doutora. – UFF

Niterói

2023

Dedico este trabalho a minha família, o meu maior incentivo de estudo, além de meus parentes e minha Orientadora Débora, que com sua competência, carisma e conhecimento veio me conduzindo nesta construção de conhecimento.

Resumo

Existem diversas propostas na literatura para a criação de aplicativos multimídia com conteúdo 2D. Entretanto, essas propostas não fornecem soluções adequadas para conteúdo imersivo tridimensional. Além disso, há pouco consenso sobre a metodologia de autoria de mídia 360°, que inclui efeitos sensoriais sincronizados com os objetos de mídia. Quando se necessita produzir conteúdo imersivo e interativo com suporte a efeitos sensoriais, geralmente são necessárias soluções individuais criadas para cada caso, utilizando várias ferramentas diferentes. Para preencher essa lacuna, esta tese propõe realizar a autoria de forma imersiva em um ambiente de Realidade Virtual (RV) para a construção de aplicações multimídia interativas. A proposta é implementada no sistema AMUSEVR, um ambiente de realidade virtual para autoria de aplicativos multimídia 360°. O AMUSEVR pode ser usado como uma alternativa exclusiva para autoria ou como um possível complemento a outras ferramentas de autoria multimídia disponíveis. O sistema proposto permite compor várias cenas 360° interativas com mídias sensoriais, organizando objetos diretamente em um espaço 3D. Além disso, o AMUSEVR permite que os usuários verifiquem o comportamento espaço-temporal da aplicação por meio de seu modo de pré-visualização, usando a tecnologia VR em todas as fases do projeto. Para avaliar a proposta, foram realizados três experimentos de autoria com a ferramenta proposta utilizando a metodologia *Goal Question Metric* (GQM) por meio da aplicação dos questionários SUS (*System Usability Scale*), UEQ (*User Experience Questionnaire*), *Cognitive Dimensions of Notations* (CDN) e com resultados analisados pelo método estatístico Mann-Whitney. Os experimentos foram realizados com 10, 20 e 42 participantes, obtendo resultados positivos.

Palavras-chave: Autoria 360°, Cenários interativos, Realidade Virtual, Multemídia, AMUSEVR.

Abstract

Numerous tools can be found in the literature for creating multimedia applications with 2D content. However, these tools do not provide adequate solutions for creating three-dimensional immersive content. Additionally, there is little consensus regarding methodology for 360° multimedia development, including synchronized sensory effects with media objects. When producing interactive immersive content with sensory effects, one typically has to use specific tools for each individual case or manage multiple tools. To address this gap, this thesis proposes immersive authoring in Virtual Reality (VR) environments for building interactive multimedia applications. Our proposal is implemented within AMUSEVR, a VR environment for the development of 360° multimedia applications. It can be used as a standalone alternative for authoring or as a complement to currently available multimedia tools. The proposed system allows for the creation of multiple scenes containing sensory media, with direct manipulation of objects in 3D space. Furthermore, the VR preview mode in AMUSEVR allows users to monitor spatio-temporal behavior throughout all stages of project development. To evaluate our proposal, three authoring experiments were carried out with the proposed tool using the Goal Question Metric (GQM) methodology through the application of the SUS (System Usability Scale), UEQ (User Experience Questionnaire), Cognitive Dimensions of Notations (CDN) and with results analyzed by the Mann-Whitney statistical method. The experiments were carried out with 10, 20 and 42 participants, obtaining positive results.

Keywords: 360° Authoring, Interactive Scenarios, VR environment, Multimedia, AMUSEVR.

Lista de Figuras

2.1	Representação simplificada do eixo de continuidade de Realidade Virtual. [45]	8
2.2	Layouts para vídeo 360 monoscópico e estereoscópico	9
2.3	Vídeo 2D Equirretangular.[62]	10
2.4	Vídeo mapa de cubo 2D.[62]	10
2.5	Projeção em esfera.	11
2.6	Projeção em plano de fundo.	12
2.7	Exemplo de aplicação com efeitos sensoriais.[14, 19]	16
2.8	Esquema centralizado de controle de atuadores multissensoriais.[51]	17
3.1	STEVE 2.0 de Mattos et al. [26].	27
3.2	Ambiente de execução da aplicação mulsemídia utilizando NLC 4.0 [38]. . .	28
3.3	Exemplo de interação com os olhos apresentado em [4].	29
3.4	Ferramenta de autoria bidimensional de Coelho et al. [11].	31
3.5	Execução da ferramenta autoria declarativa de Horst et al. [36].	34
3.6	Resultado da extração de objetos em mídias de 360 graus de Takas et al. [61].	35
3.7	Sistema FlowMatic de Zhang et al. [65].	37
3.8	Execução da ferramenta autoria em modo visualização de Mendes et al. [44].	38
3.9	Ferramenta de autoria de Park et al. [47].	39
3.10	Ferramenta de visualização para gestão de Cordeil et al. [17].	40
3.11	Proposta de Souza et al. [58] para segunda tela da TV digital.	41
4.1	Cenário ilustrativo.	53

4.2	Caso de Uso.	54
4.3	Cena 360 graus com diferentes objetos de mídia.	55
4.4	Diagrama de classes.	56
4.5	Diagrama de componentes.	58
5.1	Menu e controles.	62
5.2	Hieraquia no projeto.	64
5.3	Fluxograma para edição de cenas.	65
5.4	Fluxograma de criação, edição ou remoção de mídias.	66
5.5	Tela de edição.	68
5.6	Janela de edição de um objeto de vídeo em detalhes.	69
5.7	Edição de Áudio 3D.	70
5.8	Edição de Imagem em 2D.	71
5.9	Edição de Imagem em 360 graus.	71
5.10	Edição de Objetos interativos.	72
5.11	Edição de <i>Piture-in-Picure</i> - PIP.	72
5.12	Edição de Efeito Sensorial - ES.	73
5.13	Edição de Mensagem Textual.	74
5.14	Edição de Vídeo em 2D.	75
5.15	Edição de Vídeo em 360 graus.	75
5.16	Modo de edição.	76
5.17	Objetos de mídia 2D em uma cena de 360 graus.	77
5.18	Dependência posicional dos objetos em cena.	78
5.19	Exemplo de 6DoF.	80
5.20	Esquema do Arduino.	83
5.21	Fluxo de interatividade.	85
5.22	Cena com vídeo interativo.	86

5.23	Exemplo de portabilidade de um documento MultiSEL.	88
6.1	Diagrama direcionado com Objetivos, Questões e Métricas (GQM) do experimento com AMUSEVR Vive.	97
6.2	Participante no primeiro experimento.	99
6.3	Exibição das tarefas no menu virtual.	100
6.4	Número de respostas às afirmações do questionário do SUS (1 - discordo totalmente; 5 - concordo totalmente).	102
6.5	Escalas de Aceitabilidade e Adjetivos (pontuação SUS).	102
6.6	Avaliação de desempenho no UEQ (<i>benchmark</i>).	104
6.7	Grafo Direcionado para o modelo GQM definido.	112
6.8	Tempo médio para Q2.	115
6.9	Média de sucesso em acertos para Q4.	116
6.10	Média de sucesso em acertos para Q4 por elemento XML (experientes). . .	116
6.11	Média de sucesso em acertos para Q4 por elemento XML (Novatos). . . .	117
6.12	CDN - Grupo de experientes em XML, fatores negativos.	120
6.13	CDN - Grupo de novatos em XML, fatores negativos.	120
6.14	CDN - Grupo de experientes em XML, fatores positivos.	121
6.15	CDN - Grupo de novatos em XML, fatores positivos.	122
6.16	Diagrama direcionado com Objetivos, Questões e Métricas (GQM) do experimento com AMUSEVR Oculus.	128
6.17	Terceiro experimento com efeitos sensoriais.	129
6.18	Número de respostas às afirmações do questionário do SUS (1 - discordo totalmente; 5 - concordo totalmente).	132
6.19	Escalas de Aceitabilidade e Adjetivos associadas à pontuação SUS.	133
6.20	Comparativo de pontuação SUS dos experimentos.	133
6.21	<i>Benchmark</i> UEQ para AMUSEVR Oculus.	134
6.22	<i>Benchmark</i> UEQ comparativo entre AMUSEVR Oculus e Vive.	135

6.23	Tempo médio para os grupos em Q4.	136
6.24	Cliques do AMUSEVR versus linhas do MultiSEL.	139
6.25	Gráfico comparativo quantitativo de Q9-Q16 para o grupo de experientes. .	140
6.26	Gráfico comparativo quantitativo de Q9-Q16 para o grupo de novatos. . . .	141

Lista de Tabelas

2.1	Tipos de relações temporais suportadas por MultiSel. [25]	19
3.1	Protocolos de pesquisa na base do IEEE.	23
3.2	Resumo e destaque nos trabalhos relacionados	42
5.1	Características da mídia segundo seu tipo.	67
6.1	Resumo comparativo dos experimentos.	93
6.2	Os objetivos definidos para os experimentos AMUSEVR Vive	94
6.3	Questionário do SUS.	95
6.4	Questionário UEQ-S.	96
6.5	Perguntas personalizadas para G3.	96
6.6	Resultado UEQ-S obtido.	103
6.7	Dados UEQ-S obtidos.	103
6.8	Respostas dos participantes.	105
6.9	Tempo dos usuários experientes e novatos em VR para concluir o experimento.	105
6.10	Tempo para usuários experientes e novatos em ferramentas de autoria multimídia para concluir o experimento.	105
6.11	Objetivos do experimento MultiSEL	108
6.12	Perguntas do G1	109
6.13	Perguntas G2 - Questionário CDN	111
6.14	Respostas das perguntas Q1, Q2, Q3 e Q4.	114
6.15	Respostas das questões do objetivo G2.	118
6.16	Objetivos para os experimentos com o AMUSEVR.	124
6.17	Perguntas personalizadas para G3.	125

6.18 Perguntas personalizadas para G4.	126
6.19 Perguntas personalizadas para G5.	127
6.20 Tarefas comparativas - AMUSEVR x MultiSEL.	131
6.21 Resultado UEQ-S obtido.	134
6.22 Respostas sobre experiência dos usuários.	136
6.23 Tempo de conclusão do experimento por usuários experientes e novatos. . .	136
6.24 Comparativo da preferência dos usuários sobre métodos de autoria.	137
6.25 Comparação do número de cliques no AMUSEVR com o número de linhas no arquivo MultiSEL.	138
6.26 Resumo Q9-Q16.	140
7.1 Tabela com trabalhos já publicados.	147

Sumário

1	Introdução	1
1.1	Questão de Pesquisa	2
1.2	Objetivos	3
1.3	Contribuições da Tese	4
1.4	Estrutura do Texto	5
2	Conceitos Básicos	6
2.1	Tecnologias Imersivas	6
2.2	Mídias em 360 graus	9
2.3	Tecnologias Interativas	12
2.4	Efeitos Sensoriais	14
2.5	Linguagem MultiSEL	18
3	Trabalhos Relacionados	22
3.1	Metodologia da Revisão da Literatura	22
3.2	Autoria Mulsemídia em 2D	23
3.2.1	Saleme et al.	23
3.2.2	Guedes et al.	25
3.2.3	Mattos et al.	26
3.2.4	Linguagem NCL 4.0	28
3.3	Outros Trabalhos de Autoria Imersiva em 2D	29
3.4	Autoria Imersiva em 3D	30

3.4.1	Coelho et al.	30
3.4.2	Saleme, Covaci, Comsa, Ghinea et al.	32
3.4.3	Horst et al.	33
3.4.4	Takacs et al.	34
3.5	Outros Trabalhos de Autoria Imersiva em 3D	36
3.6	Comparação dos Trabalhos Relacionados	41
4	Ambiente de Autoria Totalmente Imersiva	50
4.1	Requisitos do Ambiente de Autoria Totalmente Imersiva	50
4.2	Especificação do Ambiente de Autoria Totalmente Imersiva	53
4.3	Avaliação do Ambiente de Autoria	59
5	AMUSEVR	60
5.1	Desenvolvimento com Unity	60
5.2	Características da ferramenta	65
5.2.1	Modos de uso da ferramenta	65
5.2.2	Objetos de Mídias	66
5.2.3	Posicionamento Espacial dos Objetos	76
5.2.4	Seis graus de liberdade	79
5.2.5	Efeitos sensoriais	81
5.2.6	Gerenciamento de mídias e sincronização	83
5.2.7	Interatividade	85
5.2.8	Gerenciamento de Arquivos	86
5.2.9	Projetos utilizando MultiSEL	87
5.3	Porte para versão do Oculus Quest 2	90
6	Avaliação	92
6.1	Experimento com AMUSEVR Vive	93

6.1.1	Metodologia	94
6.1.1.1	G1 - Questões e Métricas	95
6.1.1.2	G2 - Questões e Métricas	95
6.1.1.3	G3 - Questões e Métricas	96
6.1.1.4	Grafo GQM	97
6.1.2	Procedimentos com Participantes	98
6.1.3	Resultados	101
6.1.3.1	G1 - Análise dos Resultados	101
6.1.3.2	G2 - Análise dos Resultados	102
6.1.3.3	G3 - Análise dos resultados	104
6.1.4	Conclusões	105
6.1.5	Análise Subjetiva	106
6.1.6	Discussão	106
6.2	Experimento com a Linguagem MultiSEL para AMUSEVR Oculus	107
6.2.1	Metodologia	108
6.2.1.1	G1 - Questões e Métricas	108
6.2.1.2	G2 - Questões e Métricas	109
6.2.1.3	Grafo GQM	112
6.2.2	Procedimentos com Participantes	112
6.2.3	Resultados	113
6.2.3.1	G1 - Análise dos resultados	113
6.2.3.2	G2 - Análise dos Resultados	117
6.2.4	Análise Subjetiva	122
6.2.5	Discussão	123
6.3	Experimento com AMUSEVR Oculus com Renderização de Efeitos Sensoriais	123
6.3.1	Metodologia	124

6.3.1.1	G1 - Questões e Métricas	124
6.3.1.2	G2 - Questões e Métricas	125
6.3.1.3	G3 - Questões e Métricas	125
6.3.1.4	G4 - Questões e Métricas	126
6.3.1.5	G5 - Questões e Métricas	127
6.3.1.6	Grafo GQM	128
6.3.2	Procedimentos com Participantes	128
6.3.3	Experimento Comparativo para Esforço de Autoria - AMUSEVR x MultiSEL	130
6.3.4	Resultados	132
6.3.4.1	G1 - Análise dos Resultados	132
6.3.4.2	G2 - Análise dos Resultados	134
6.3.4.3	G3 - Análise dos resultados	135
6.3.4.4	G4 - Análise dos Resultados	137
6.3.4.5	G5 - Análise dos resultados	139
6.3.5	Análise Subjetiva	142
6.3.6	Discussão	142
6.4	Considerações Finais	144
7	Conclusão	146
7.1	Publicações	147
7.2	Trabalhos futuros	148
	Referências	150
	Apêndice A – Códigos utilizados no Arduino	156
A.1	Código Ardino para controle de Efeitos Sensoriais	156
	Apêndice B – Códigos utilizados no experimento MultiSEL	157

B.1	Código Inicial entregue aos participantes	157
B.2	XML esquema MultiSEL para Projetos AMUSEVR	157
B.3	Gabarito para Correção da atividade prática do experimento	162

Capítulo 1

Introdução

A evolução de aplicativos multimídia trouxe novos tipos de conteúdo, como imagens e vídeos 360 graus, que proporcionam ao consumidor um senso aumentado de imersão no conteúdo midiático. Portanto, eles trazem a promessa de novas narrativas nas comunicações, educação e marketing, entre outras áreas [57]. Deve-se notar que há algumas limitações nesta tecnologia, pois a liberdade de orientação encontrada no ambiente de RV (Realidade Virtual) introduz imprevisibilidade na experiência do usuário. O público corre o risco de perder elementos essenciais ou detalhes, devido aos métodos de interação com o conteúdo, bem como à falta de experiência dos produtores de conteúdo e do público eventual [57].

Considerando mídia 360, uma limitação é o fato de que o usuário geralmente está fixado no centro do aplicativo. Isso pode funcionar para conteúdo de cinema (filmes e similares), onde se senta em uma localização fixa, mas nem sempre funciona para museus interativos, por exemplo. Para atender à demanda por liberdade de movimento, a RV tem sido usada, pois oferece uma mistura de diferentes tipos de conteúdo ao longo de 6 eixos de movimento. Além disso, múltiplos objetos de mídia podem ser localizados espacialmente na apresentação e, usando dispositivos controladores, os usuários podem interagir com o conteúdo, permitindo que evoluam de espectadores passivos para participantes ativos [14, 31].

Para a RV se proliferar como uma tecnologia que oferece uma experiência de usuário aprimorada [57], precisamos fornecer ferramentas para produção desse tipo de conteúdo para autores que não são especialistas em programação. No entanto, as ferramentas disponíveis ainda usam técnicas de autoria herdadas de tecnologias tradicionais voltadas para mídia 2D. Alguns usam o enfoque em construir apresentações baseadas em linhas do tempo [11, 12], outros em relações entre objetos de mídia [26]. Não é surpreendente

que fornecer um ambiente de autoria que permita manipular o conteúdo de mídia 360 em espaço 3D ainda seja um desafio.

Com base na indústria de desenvolvimento de jogos eletrônicos, que é um ambiente atual para a criação de conteúdo RV [57], o trabalho de [44] recomenda o uso da RV para produzir e posicionar objetos dentro do conteúdo, a fim de melhorar a produção e fornecer maior fidelidade no produto final. Para este tipo de tecnologia, motores de jogos, como o Unity, oferecem bom suporte de produção para conteúdo RV, além de suporte para mídias padrão, como imagem, vídeo e áudio, interação do usuário e até integração com efeitos sensoriais [47, 53].

De fato, a integração com efeitos sensoriais se torna desejável, já que a maior parte da comunicação humana é não verbal e frequentemente usamos outros sentidos, além da visão e audição, para compreender o mundo ao nosso redor, socializar e nos entreter [18]. Além disso, podemos interpretar melhor o conteúdo oferecido por uma aplicação combinando diferentes sentidos [57]. Esse tipo de conteúdo que envolve mídias tradicionais como áudio e vídeo, e ainda efeitos sensoriais, é chamado de conteúdo mulsemídia (*mulsemedia*) como descrito em [33]. Neste trabalho, propomos a autoria de efeitos sensoriais e eventos interativos sincronizados com conteúdo multimídia tradicional, como o modelo conceitual MultiSEM [26] e o editor gráfico STEVE [1]. No entanto, há menos aplicações e metodologias para este propósito direcionadas a RV e a conteúdo 360. Portanto, oferecer um ambiente de autoria que atenda à variedade de conteúdos e tecnologias [25] para aplicações imersivas, interativas e multissensoriais, permitindo ao autor ter a sensação de imersão no conteúdo que será proporcionado ao usuário consumidor, ainda é um desafio.

1.1 Questão de Pesquisa

Visando preencher lacunas encontradas na literatura, esta tese propõe e investiga o conceito de autoria imersiva de aplicações multissensoriais interativas com conteúdo 360°. Essas aplicações são chamadas de mulsemídia (*MulSeMedia - Multiple Sensorial Media*) [33]. Através desta pesquisa, busca-se responder a seguinte questão:

"Qual é a experiência do autor na criação de aplicações multissensoriais interativas em 360° usando um ambiente de autoria imersivo?"

Para investigar e responder a essa questão de pesquisa, foi proposto um ambiente de autoria imersivo para aproximar o autor do ambiente de apresentação final do usuário. A proposta foi implementada em um ambiente de RV chamado AMUSEVR (*Authoring 360*

MUltimedia and Sensory Effects in VR).

O ambiente de autoria foi avaliado com usuários através de experimentos para criação e execução de aplicações multissensoriais interativas buscando responder à nossa questão de pesquisa. A principal contribuição deste trabalho é propor e avaliar se fornecer imersão para o autor da aplicação e não apenas para os usuários finais enriquece e facilita o processo de autoria.

1.2 Objetivos

Com o objetivo de facilitar o desenvolvimento de projetos mulsemídia para ambientes imersivos, esta tese propõe uma nova abordagem para a criação de conteúdo mulsemídia com base em um ambiente de autoria imersiva. A proposta foi implementada em um ambiente integrado de autoria e execução chamado AMUSEVR (*Authoring 360° MUltimedia and Sensory Effects in VR*), que oferece imersão em realidade virtual para a criação do conteúdo. Essa abordagem baseia-se no desenvolvimento e visualização em tempo de execução em ambiente de realidade virtual, com objetivo de facilitar o trabalho do autor para criação de conteúdo mulsemídia.

AMUSEVR foi desenvolvido utilizando o motor gráfico Unity, como uma plataforma para desenvolvimento de aplicações em realidade virtual. Para permitir o intercâmbio de projetos desenvolvidos com AMUSEVR e outros sistemas de autoria mulsemídia, foi utilizado um formato textual baseado em XML (*eXtensible Markup Language*) para representação das aplicações criadas com AMUSEVR, permitindo importação e exportação de projetos. Para isso, foi utilizada a linguagem declarativa MultiSEL (*Multimedia Sensory Effect Language*) [24], que servirá como formato de intercâmbio entre os modos de autoria e execução do sistema proposto.

Para responder à questão de pesquisa indicada na Seção 1.1, outro objetivo da tese é a avaliação do Ambiente de Autoria Totalmente Imersiva (AATI), através de testes de usabilidade e experiência do usuário, utilizando questionários bem conhecidos na literatura como SUS (*System Usability Scale*) [3, 10] e UEQ (*User Experience Questionnaire*) [40, 55] adaptados para o sistema AMUSEVR.

Um outro objetivo do trabalho é realizar uma avaliação da linguagem MultiSEL, através do *framework* de Dimensões Cognitivas de Notações (*Cognitive Dimensions of Notations*) [9] e outras questões personalizadas com objetivo de comparar a autoria declarativa com a autoria imersiva em realidade virtual.

Para o planejamento das avaliações realizadas, foi utilizado o modelo GQM (*Goal Question Metric*) [6], que permite organizar e estruturar os experimentos, facilitando sua compreensão e reprodutibilidade.

1.3 Contribuições da Tese

As contribuições desta tese incluem:

- A proposta de uma nova abordagem de autoria imersiva em realidade virtual para a produção de aplicações mulsemídia interativas, que suporta sincronização espaço-temporal para múltiplos objetos de mídia e efeitos sensoriais, incluindo mídias 360°;
- A implementação desta abordagem em um ambiente chamado AMUSEVR desenvolvido com a plataforma Unity. Este ambiente é capaz não somente de facilitar a criação de aplicações mulsemídia com conteúdo 360° aos autores, mas também de executar tais aplicações, sendo utilizado pelos usuários finais;
- A avaliação da qualidade de experiência do autor ao utilizar a abordagem de autoria imersiva proposta;
- O refinamento e avaliação da linguagem declarativa MultiSEL idealizada para especificação de conteúdo mulsemídia, como formato de interoperabilidade entre ambientes de autoria e execução;
- A demonstração prática de uma nova abordagem de autoria e formatação espaço-temporal para múltiplos objetos de mídia em dispositivos de realidade virtual.

Por outro lado, não está no escopo deste trabalho:

- Incorporação de objetos ou cenários tridimensionais (3D);
- Detalhamento na implementação de equipamentos atuadores para renderização de efeitos sensoriais físicos;
- Captura ou edição de qualquer tipo de mídia tradicional.

1.4 Estrutura do Texto

O restante da tese está organizado da seguinte maneira. O Capítulo 2 apresenta uma revisão da literatura existente sobre imersão e mulsemídia, destacando conceitos fundamentais, bem como sua importância e aplicação em ambientes imersivos.

O Capítulo 3 fornece uma visão geral dos trabalhos relacionados e como eles foram utilizados para inspirar e fundamentar a proposta de ambiente de autoria imersivo.

O Capítulo 4 descreve detalhadamente o ambiente de autoria totalmente imersiva, incluindo sua arquitetura e funcionalidades.

O Capítulo 5 apresenta a implementação do ambiente de autoria AMUSEVR, incluindo a descrição de suas funcionalidades e uso da Unity como plataforma de desenvolvimento.

O Capítulo 6 apresenta os resultados das avaliações realizadas, incluindo testes de usabilidade e experiência do usuário, bem como os resultados dos testes comparativos com outras ferramentas.

Por fim, o Capítulo 7 apresenta as conclusões finais, bem como sugestões para trabalhos futuros.

Capítulo 2

Conceitos Básicos

Este capítulo apresenta conceitos básicos necessários para o entendimento das tecnologias e temas abordados na pesquisa.

2.1 Tecnologias Imersivas

As tecnologias imersivas são aquelas que sugerem ao usuário uma sensação de estar presente ou inserido no conteúdo apresentado. O nível de imersão pode ser completo ou parcial, gerando vários tipos de sensações como: emoções como prazer, satisfação e até incômodo. Essas sensações normalmente são induzidas pelo autor do projeto como abordado em [59, 8]. A capacidade de imersão pode ser dividida segundo sua origem de duas formas como em [23, 21]:

- Imersividade natural – induzida normalmente por fármacos e drogas alucinógenas que levem ao usuário a ter sensação de estar vivendo em cenários paralelos; e
- Imersividade artificial – induzida através de equipamentos e tecnologias diversas de forma sintética.

As tecnologias imersivas artificiais oferecem diversas formas de imersão ao usuário, sendo algumas delas descritas em trabalhos como [23, 21, 22]. Essas técnicas incluem:

- Tecnologias de estimulação eletrônica cerebral, que visam modificar o estado cognitivo do usuário para proporcionar uma sensação de imersão;
- Tecnologias de imersão através de cavernas virtuais ou ambientes preparados, que

utilizam projeção de imagens e som para criar uma sensação de estar dentro de um ambiente diferente da realidade;

- Tecnologias de realidade aumentada ou estendida, que adicionam elementos virtuais ao mundo real, permitindo uma sensação de interação com esses elementos;
- Tecnologias de imersão através de equipamentos vestíveis, como óculos de realidade virtual e outros dispositivos que permitem ao usuário estar dentro do ambiente virtual.

Cada uma dessas opções é indicada para públicos, cenários ou aplicabilidades específicas, sendo que o mesmo conteúdo pode ser ofertado utilizando mais de uma delas. Atualmente a tecnologia mais usada para conteúdos imersivos é a Realidade Virtual (RV ou VR - *Virtual Reality*) através de capacetes vestíveis chamados de HMD (*Head-Mounted Display*), devido à redução de custo e popularização dos mesmos. O principal mercado dos produtores de conteúdo imersivo atualmente é a indústria de jogos eletrônicos, mas essas tecnologias também estão sendo aplicadas em outras áreas como treinamento, educação e entretenimento. Além da citada RV, há outras tecnologias imersivas que podemos citar como [22]:

- Telepresença: é uma tecnologia que utiliza câmeras de vídeo e microfones remotos para envolver e projetar o usuário profundamente no mundo virtual. Como exemplos, podemos citar robôs interativos de atendimento remoto e robôs de exploração planetária.
- Realidade Aumentada (RA): é uma tecnologia que utiliza elementos reais ou partes do mundo real e adiciona elementos virtuais para misturar com a realidade. Existem dois tipos de RA: o *see-through* (ver através) e o *non see-through* (não ver através). Ambos são uma combinação da RV de simulação com a telepresença, juntando o ambiente real com o ambiente virtual.
- RV de Projeção: normalmente implementada em cavernas de imersão, na RV de Projeção o usuário está fora do mundo virtual, mas pode se comunicar com personagens ou objetos dentro dele. A imagem do usuário é capturada e projetada em uma tela que representa um mundo virtual, geralmente utilizada em museus virtuais.
- Realidade Estendida (XR - *eXtended Reality*): Quando se combinam as tecnologias de RV, RA e outras tecnologias para aprofundar a imersão, elas formam um novo

conceito chamado Realidade Estendida (XR). Essa nova terminologia engloba o que não se encaixa exclusivamente em uma das tecnologias citadas anteriormente. Além disso, a XR também oferece outros tipos de interação além da audiovisual e háptica, permitindo maiores sensações através de simuladores.

Dentro da literatura acadêmica, um uma vertente de pensamento que utiliza do conceito apresentado no artigo [45], o autor propõe uma classificação para exibições que usam realidade aumentada, com base em um continuum que varia de displays completamente reais a displays completamente virtuais, com a realidade aumentada em algum ponto intermediário. O autor descreve as características de cada tipo de exibição, como a realidade aumentada, a realidade virtual e a realidade mista, e discute suas aplicações potenciais em diversas áreas, incluindo educação, entretenimento, design e treinamento. Ele também aborda algumas das tecnologias necessárias para a realização dessas exibições, como rastreamento de posição e orientação, reconhecimento de objetos, visão computacional e processamento de imagens.

Neste trabalho de pesquisa, foi utilizado óculos de realidade virtual para criar um ambiente imersivo, o que o colocaria no extremo direito do eixo de classificação do autor, que corresponde a exibições completamente virtuais, conforme ilustrado na Figura 2.1. Nesse tipo de exibição, o usuário é completamente imerso em um ambiente virtual, com pouca ou nenhuma conexão com o mundo real. Isso pode ser particularmente útil em cenários de treinamento ou em jogos, onde o usuário precisa se concentrar exclusivamente em uma tarefa ou objetivo específico. No entanto, também apresenta desafios em termos de ergonomia e conforto, bem como limitações técnicas em relação à qualidade gráfica e de som.

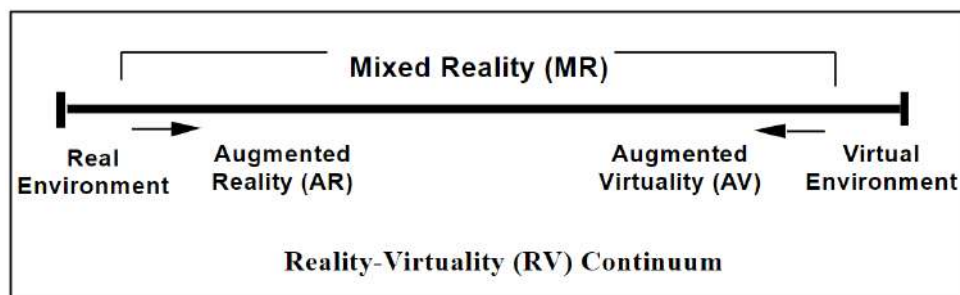


Figura 2.1: Representação simplificada do eixo de continuidade de Realidade Virtual. [45]

2.2 Mídias em 360 graus

Uma das técnicas para facilitar a imersão do usuário é prover preenchimento do campo de visão em todas as direções, para tal, entre outras, são utilizadas imagens e vídeos em 360 graus. Essas mídias normalmente são capturadas com câmeras panorâmicas de propósito especial. Existem vários tipos de câmeras de vídeo omnidirecionais (referente à capacidade de gravar em todas as direções). Independentemente da técnica de captura, posteriormente, esses vídeos são agrupados em uma etapa de pós-processamento para criar um vídeo panorâmico de 360 graus, geralmente em uma parametrização latitude-longitude, também conhecido como formato equirretangular [63].

Em relação ao tipo de vídeo, eles normalmente são gravados, assim como áudios, nos formatos monoscópico e estereoscópico, como ilustrado na Figura 2.2. O formato monoscópico é mais comumente utilizado em aplicações exibidas em 2D, como navegadores e aplicativos de navegação, enquanto os formatos estereoscópicos são mais comumente utilizados em equipamentos de Realidade Virtual, onde uma imagem é enviada para cada olho do usuário. No entanto, atualmente, os equipamentos de Realidade Virtual são projetados para suportar ambos os formatos de forma nativa.

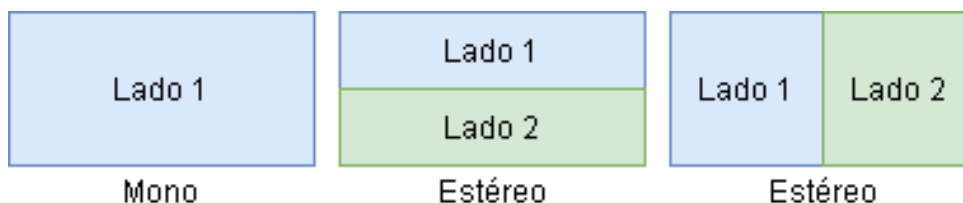


Figura 2.2: Layouts para vídeo 360 monoscópico e estereoscópico

Para cada lado do vídeo apresentado na Figura 2.2, existem vários layouts diferentes disponíveis. Os layouts mais comuns são o equirretangular (longitude e latitude), ilustrado na Figura 2.3, e o mapa de cubo (6 quadrados), ilustrado na Figura 2.4 [62].

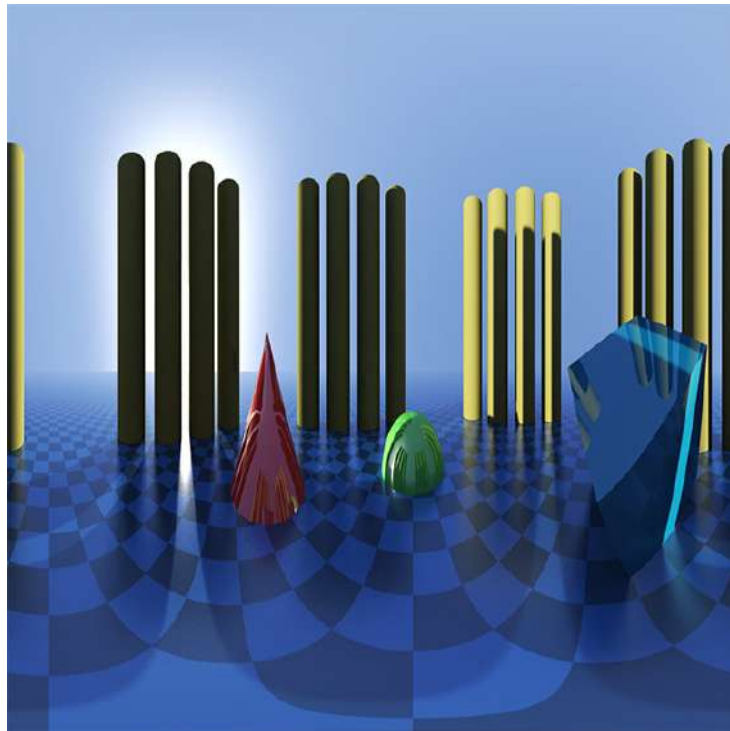


Figura 2.3: Vídeo 2D Equirretangular.[62]

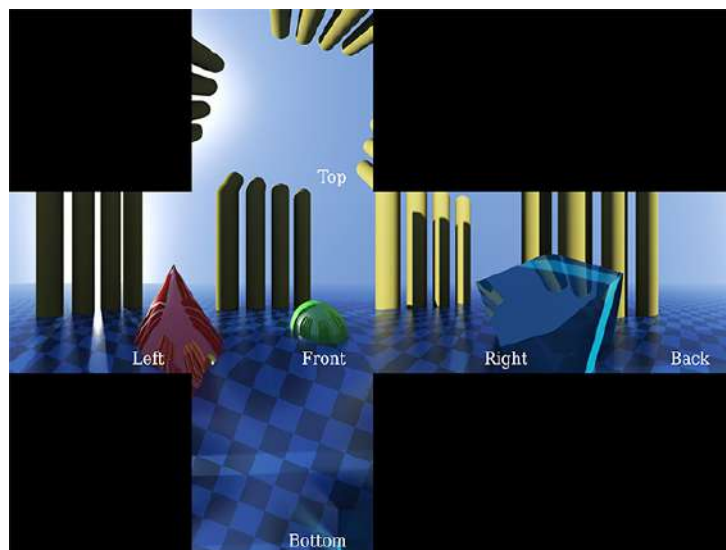


Figura 2.4: Vídeo mapa de cubo 2D.[62]

Por fim, cada vídeo ou imagem para ser reproduzido deve ser projetado em alguma moldura conhecida como projeção. Assim como na cartografia, existem vários tipos de projeções. As mais comuns são a projeção cúbica, esférica, bidimensional e de plano de fundo [64]. A Figura 2.5 apresenta um exemplo de projeção interna em esfera, onde a câmera que representa o campo de visão do usuário fica inserida no centro da esfera. Já a Figura 2.6, trata-se de um exemplo de projeção em plano de fundo, na qual o player

de vídeo se encarrega de preencher sempre o fundo do campo de visão do usuário com uma parte do conteúdo correspondente. A técnica de projeção em cubo é semelhante à da esfera e a bidimensional é usada em aplicações como o Google Maps¹.



Figura 2.5: Projeção em esfera.

¹ "<https://www.google.com.br/maps>"



Figura 2.6: Projeção em plano de fundo.

2.3 Tecnologias Interativas

As tecnologias imersivas e interativas são abordagens que visam criar experiências mais envolventes e memoráveis para os usuários. As tecnologias imersivas, como Realidade Virtual, Realidade Aumentada e Realidade Mixta, buscam criar uma sensação de imersão ou mergulho para o usuário, através de dispositivos e técnicas específicas. Já as tecnologias interativas permitem ao usuário interagir com os sistemas ou dispositivos de alguma forma, seja através de toque, gestos, voz, movimento ou pensamento, proporcionando uma experiência mais natural e intuitiva. Ambos os tipos de tecnologias são importantes para criar uma interação envolvente e imersiva para o usuário [50].

Para entender melhor as tecnologias interativas, é importante saber que existem conteúdos midiáticos no qual o usuário participa de forma passiva, não podendo interferir no desenvolvimento do conteúdo, portanto, não são interativos. Por outro lado, quando o usuário tem a capacidade de atuar no conteúdo, seja fazendo escolhas, controlando um personagem ou simplesmente mudando a direção de visualização do conteúdo, chamamos

de mídia interativa. Considerando que para cada tecnologia midiática oferecida existe uma ou mais formas de interação, e deixando um pouco de lado a interface de usuário (UI – *User Interface*), existem alguns modos mais comuns de controlar e interagir com o conteúdo midiático [50], por exemplo:

- Interações tradicionais - Controles remotos, teclado, mouse, controles de videogames;
- Interações verbais - Comandos de voz;
- Interações baseadas em captura e análise de imagens e vídeos - Comandos gestuais, expressões faciais, corporais, sensores; e
- Interações baseadas em dispositivos vestíveis - Pulseiras, capacetes de RV/RA/XR com ou sem equipamentos complementares.

O uso do tipo de interação deve ser planejado e compatível com o tipo de mídia e cenário. Por exemplo, não é recomendável oferecer controles verbais para mídias que sejam exibidas em locais de muito público, pois pode haver interferência sonora. Outro exemplo de uso inapropriado é oferecer comandos de reconhecimento de gestos em locais com grande incidência de raios solares, pois a luz solar pode causar interferência em alguns equipamentos que usam tecnologia de rastreamento por luz infravermelha.

O modelo de controles interativos para mídias como vídeos, áudios, jogos e texto está bem estabelecido e difundido. No entanto, quando se trata de conteúdos inseridos em ambientes de realidade virtual, não é possível determinar uma única ferramenta que abranja todas as possibilidades. Entre as tecnologias interativas utilizadas em algumas aplicações de realidade virtual, podemos citar:

- Simuladores de armas – amplamente utilizados em jogos de tiro;
- Gestos e comando de voz – além de comando de voz, também inclui rastreamento do corpo ou de partes dele; e
- Controladores – dispositivos de controle que geralmente acompanham os kits de capacetes de RV/RA/XR.

Dentro de um ambiente virtual, existem diversas técnicas para interagir com o conteúdo, tais como:

- Interação física: utilizando controles físicos ou representações virtuais das mãos para pegar e manipular objetos;
- Seleção e clique: similares aos comandos de um mouse, permitindo selecionar e clicar em elementos da interface;
- Rastreamento de olhar: utilizando tecnologias de rastreamento de olhos para detectar e responder aos movimentos do olhar do usuário;
- Movimento: permitir ao usuário se deslocar fisicamente no ambiente virtual para ativar eventos e interagir com elementos do cenário; e
- Comando de voz: utilizando comandos de voz para invocar ações e interagir com elementos do conteúdo.

Além dessas técnicas tradicionais, há também novas abordagens sendo desenvolvidas, como o controle mental [41], o rastreamento de mãos [29] e o rastreamento de olhos [46]. Isso indica que a área de interação em realidade virtual está em constante evolução e é um campo de pesquisa em constante expansão que demanda investigações específicas para cada tipo de interação.

2.4 Efeitos Sensoriais

Efeitos sensoriais são estímulos aos sentidos do usuário, como visão, audição, tato, olfato e paladar, a fim de criar uma experiência mais envolvente e imersiva. Estes efeitos podem ser utilizados em diversas áreas, como entretenimento, publicidade, educação e treinamento, e podem ser combinados com outras tecnologias como realidade virtual, realidade aumentada e interatividade para criar uma experiência completa. Efeitos sensoriais podem incluir desde iluminação e som, até aromas e degustação de alimentos, o objetivo é criar uma experiência imersiva que envolva diferentes sentidos. Os efeitos sensoriais vêm sendo abordados por diversos autores, em muitos casos, de forma complementar a realidade virtual, como descrito por [47, 53, 11].

O estudo de mídias com enfoque em estimular os diferentes sentidos humanos através de efeitos sensoriais é uma área ampla e em constante evolução, como evidenciado por diversos trabalhos científicos, como os citados por Covaci et al. [20], Comcsa et al. [14], Saleme et al. [52, 51, 53] e Farias et al. [27]. Dentre as tecnologias que têm sido estudadas para essa finalidade, destacam-se os equipamentos atuadores que se concentram em

emitir sensações relacionadas a umidade, vento, cheiro e calor. Embora existam vários equipamentos disponíveis no mercado para essas finalidades, ainda há desafios relacionados à falta de padronização na comunicação com esses equipamentos, nos formatos de armazenamento e até mesmo nos conceitos, como discutido por Saleme et al. em suas pesquisas [52, 51, 53].

Dentre os trabalhos citados, é possível destacar questões de problemas comuns enfrentadas durante a implementação de sistemas mulsemídias, tais como:

- Padronização da comunicação entre os equipamentos;
- Formatos de armazenamento de dados;
- Falta de consenso nos conceitos utilizados;
- Integração com outros sistemas e dispositivos;
- Considerações éticas e de privacidade.

Além disso é importante também considerar em uma implementação:

- Posicionamento dos atuadores - o usuário/participante pode se deslocar pelo cenário. Com isso, o efeito emitido pelo atuador deve ser levado junto ao usuário como parte da vestimenta ou simplesmente ser fixado em um ponto do mundo físico? Se somente a cabeça do usuário virar, o efeito deve ser ativado/desligado? A Figura 2.7, retirada de [14, 19], mostra um atuador de emissão de odores fixado à cabeça do usuário que pode ou não ser ativado, de acordo com a direção para onde o usuário estiver olhando, caso a apresentação seja configurada dessa forma. Na mesma figura, pode-se observar um atuador de vento, que é fixo no ambiente real e não pode ter sua posição alterada no ambiente virtual.
- Intensidade - É um atributo comum à maioria dos atuadores. Dependendo do atuador, podem ser implementadas variações de velocidade do vento, nível de vibração, intensidade de odor, entre outros. No entanto, quando considera-se a liberdade de locomoção do usuário, a emissão de um vento, cheiro ou similares deve ser representada ou recriada espacialmente no mundo imersivo? Por exemplo, quando o usuário se afasta de um ponto virtual que emite vento, a intensidade do vento deve ser alterada? Sabe-se que algumas manifestações, como vento e cheiro, no mundo real, se espalham volumetricamente pelo espaço. O mesmo é muito difícil de simular em um

cenário imersivo, e caso a apresentação não seja bem planejada, pode atrapalhar a experiência imersiva.

- Exposição – durante uma apresentação multissensorial, o acionamento dos equipamentos de efeitos sensoriais possui peculiaridades que requerem um planejamento distinto para cada situação. Por exemplo, considerando o efeito de cheiro, os odores podem ficar impregnados no epitélio olfatório presente nas narinas do usuário por mais tempo do que o desejado ou programado no contexto da apresentação. A intensidade de um efeito frio, por exemplo, pode afetar a sensação posterior de um efeito de calor. Isso tudo desconsiderando o tempo de entrega que o equipamento atuador leva para fornecer o conteúdo ao usuário. Se esse conteúdo demorar a ser entregue ou percebido, pode perder sincronismo com o conteúdo audiovisual, por exemplo.



Figura 2.7: Exemplo de aplicação com efeitos sensoriais.[14, 19]

Os efeitos sensoriais podem ser criados através de capturas por sensores ou gerados de forma sintética através de ferramentas de autoria. Para ambos os casos, é necessário sua especificação para posterior execução em um player compatível. Os formatos atuais que dão suporte a essa tecnologia são principalmente os seguintes:

- O MPEG-V², que é um formato padrão bem difundido e robusto, porém possui limitações pois não consegue se antever a produção de novos equipamentos e efeitos sensoriais [51]; e

²<https://mpeg.chiariglione.org/standards/mpeg-v>

- Os formatos não-padronizados textuais possuem a característica principal de serem baseados em script ou linguagem de marcação textuais. São abertos e de fácil manipulação, porém não tão bem difundidos para essa finalidade. Diversos autores têm implementações usando linguagens específicas como [44, 53, 26].

Como há muitos tipos de equipamentos de diversos fabricantes que emitem até mesmo os mesmos efeitos sensoriais de forma diferente. Aliado a isso, existem diversos tipos de conexão, tais como cabos, *Bluetooth*³ e até *Wi-Fi*⁴. E, para complicar ainda mais, existem diversos tipos de protocolos de comunicação, tanto idealizados pelos desenvolvedores dos equipamentos, quanto pelos desenvolvedores das ferramentas.

Como não há um consenso, alguns autores se debruçaram em algumas soluções, como Saleme et al. [51] que baseiam uma de suas propostas em centralizar o controle em um servidor para gerenciamento individual de cada atuador, como ilustrado na Figura 2.8. No entanto, há outros exemplos baseados em implementações de servidores *Message Queuing Telemetry Transport* (MQTT)⁵, como os trabalhos [27, 4], ou em controle direto via API dos equipamentos atuadores, como descrito em [44].

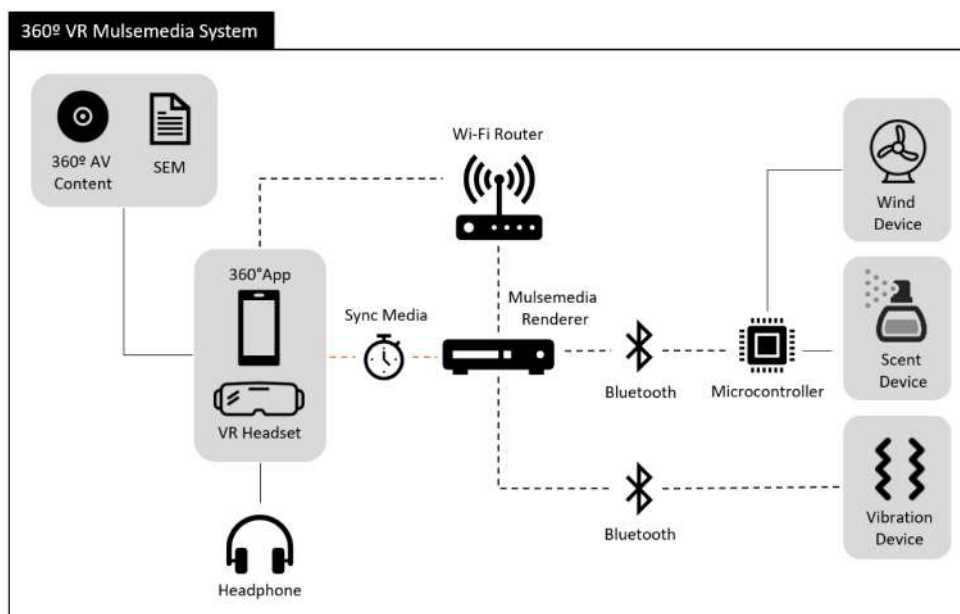


Figura 2.8: Esquema centralizado de controle de atuadores multissensoriais.[51]

É importante observar que as questões apresentadas requerem considerações específicas para cada aplicação de mídia interativa. Além disso, é importante lembrar que a

³<https://www.bluetooth.com/>

⁴www.wi-fi.org










⁵<https://mqtt.org/>

pesquisa sobre atuação em diferentes sentidos humanos através de efeitos sensoriais ainda está em desenvolvimento, portanto, é importante estar atento às novas descobertas e tendências no campo. Além disso, é importante mencionar a importância de uma boa documentação e padronização dos equipamentos utilizados para garantir a compatibilidade e facilidade de uso na implementação de projetos mulsemídia.

2.5 Linguagem MultiSEL

A linguagem MultiSEL (*Multimedia Sensory Effect Language*) [24] é baseada no Modelo MultiSEM (*Multimedia Sensory Effect Model*) [26]. Ela permite a especificação de documentos multimídia com mídias 360° utilizando uma abordagem declarativa, seguindo o paradigma de outras linguagens multimídia, como HTML [16], NCL (*Nested Context Language*) [48] e SMIL (*Synchronized Multimedia Integration Language*) [15]. MultiSEL também oferece suporte a diversos tipos de relacionamentos temporais, os quais podem ser acessados através dos conectores vistos na Tabela 2.1

Tabela 2.1: Tipos de relações temporais suportadas por MultiSel.[25]

Relações	Simbolos	Conector	Hipermídia	Descrição
<i>Starts</i>		<i>onBeginStart</i>		Os nós começam quando o nó principal começa
<i>Starts With Delay</i>		<i>onBeginStartDelay</i>		Os nós começam com atraso quando o nó primário começa
<i>Finishes</i>		<i>onEndStop</i>		Os nós terminam quando o nó primário termina
<i>Finishes With Delay</i>		<i>onEndStopDelay</i>		Os nós terminam com atraso quando o nó primário termina
<i>Meet</i>		<i>onEndStart</i>		Os nós começam quando o nó principal termina
<i>Meets With Delay</i>		<i>onEndStartDelay</i>		Os nós começam com atraso quando o nó primário termina
<i>Met By</i>		<i>onBeginStop</i>		Os nós terminam quando o nó principal começa
<i>Met By With Delay</i>		<i>onBeginStopDelay</i>		Os nós terminam com atraso quando o nó primário começa
<i>Before</i>		<i>onEndStartDelay</i>		Apresentar nós sequencialmente com um atraso entre eles quando o nó primário terminar

No documento MultiSEL exemplificado na Listagem 2.1, uma cena 360° é representada pelo elemento *scene*. Essa cena pode incluir elementos como objetos de mídia (*media*), elementos de efeitos sensoriais (*effect*), elementos de relacionamento de sincronização e interatividade (*relation*), entre outros. O atributo *primaryComponent* da cena (*scene*), identifica seu componente interno a ser iniciado quando a cena começa a ser renderizada

no ambiente VR. Os autores também podem criar relacionamentos entre as cenas para especificar um projeto completo. MultiSEL fornece tipos de relação predefinidos usando conectores MultiSEM [26].

```

1 <multisel id="360MulsemmediaPresentation">
2   <scene id="scene1" primaryComponent="back_video">
3     <port id="port_image" component="2D_image"/>
4     <media id="back_video" src="rio.mp4" type="360video">
5       <property name="background" value="true"/>
6       <property name="loop" value="false"/>
7     </media>
8     <media id="2D_image" src="portal.jpeg" type="image">
9       <property name="position" value="0,2,6"/>
10    </media>
11    <relation type="starts">
12      <primary component="back_video"/>
13      <secondary component="2D_image"/>
14    </relation>
15  </scene>
16  <relation type="onSelectionStart" keyCode="TRIGGER">
17    <primary component="scene1" interface="port_image"/>
18    <secondary component="scene2"/>
19  </relation>
20  <scene id="scene2" primaryComponent="image_background">
21    ...
22    <effect id="wind_effect" type="windEffect">
23      <property name="intensityValue" value="50%"/>
24      <property name="position" value="0,6,12"/>
25    </effect>
26    <relation type="starts" delay="10s">
27      <primary component="image_background"/>
28      <secondary component="wind_effect"/>
29    </relation>
30  </scene>
31 </multisel>

```

Listing 2.1: Exemplo de documento MultiSEL.

Ainda usando o exemplo de código MultiSEL dado na Listagem 2.1, este documento define duas cenas: *scene1* e *scene2*. A cena 1 começa com a mídia *back_video* como vídeo em 360° como pano de fundo. Ela também contém uma relação (*relation*) cujo tipo é *starts*. Esta relação define que o componente secundário (*2D_image*) começa a ser exibido quando o primário (*back_video*) começa. Dentro do elemento raiz deste documento (*multisel*), é especificada uma relação entre as duas cenas. Essa relação define que a *scene2* começa quando a mídia *2D_image* da *scene1*, identificada através de seu elemento *port* com valor *port_image*, é selecionada usando a tecla de gatilho (*TRIGGER*) presente nos controles do equipamento de RV, como o HTC Vive. Em relação à *scene2*, ela contém um efeito de vento que começa com um atraso de 10 segundos, a partir do

momento em que a imagem de pano de fundo *image_background* da *scene2* começa.

A linguagem MultiSEL oferece versatilidade e adaptabilidade para novos tipos de mídia 2D ou 3D, bem como suporte a efeitos sensoriais, diferentes cenas mulsemídia, hierarquização, interações e relacionamentos de sincronização entre mídias. Devido a essas características, esta tese optou por adotar a linguagem MultiSEL como formato textual. MultiSEL foi escolhida para ser usada como formato de intercâmbio entre o Modo Autor e o Modo Visualizador do sistema AMUSEVR, como será detalhado no Capítulo 5. Um autor pode criar ou editar um aplicativo imersivo em 360° usando o Modo Autor, exportar a especificação para o formato MultiSEL e distribuí-lo posteriormente aos usuários finais para experimentá-lo usando o Modo Visualizador. O Modo Visualizador permite importar o documento MultiSEL e executá-lo no ambiente VR.

MultiSEL também pode ser usada por outra ferramenta de autoria mulsemídia para criar uma aplicação com conteúdo 360°, que pode ser posteriormente executada no Modo Visualizador do AMUSEVR.

Capítulo 3

Trabalhos Relacionados

Este capítulo aborda estudos relacionados às propostas de linguagens e ferramentas de autoria multimedial 2D, bem como discute trabalhos e ferramentas de autoria com foco em ambientes imersivos 3D.

3.1 Metodologia da Revisão da Literatura

Para esta pesquisa foram utilizadas buscas através de ferramentas especializadas em publicação científica. Baseando-se nas máquinas de busca, foram criadas expressões utilizando a linguagem específica de cada base de dados durante o mês de julho de 2020 para o Instituto de Engenheiros Eletricistas e Eletrônicos (IEEE), *Association for Computing Machinery* (ACM), entre outras de grande relevância.

Critérios de inclusão:

- Relevância de conteúdo;
- Palavras chaves ou sinônimos das mesmas referentes a pesquisa;
- Trabalhos com menos de 10 anos;
- Se é disponível para baixar o arquivo para leitura;
- Outros: Especificados nos comentários.

Critérios de exclusão:

- Idioma de escrita diferente de português ou inglês;

- Trabalhos duplicados;
- Não atender algum critério de inclusão;
- Não disponíveis para download;
- Trabalhos cujo foco não corresponda à questão de pesquisa e
- Outros: Especificados nos comentários.

A busca na base do IEEE apresentou os resultados indicados na Tabela 3.1.

Tabela 3.1: Protocolos de pesquisa na base do IEEE.

ID	Data	Base	Expressão	Quantidade
01	01/07/2020	IEEE	("imersive multimedia" OR "Authoring multimedia" OR "Interactive multimedia")	1397
02	01/07/2020	IEEE	("imersive mulsemedia" OR "Authoring mulsemedia" OR "Interactive mulsemedia" OR "Fully imersive authoring")	123

Através das buscas foram obtidos 1520 resultados, optou-se por descartar o primeiro por ter muitos resultados, após isso foi focado no segundo. Após, optou-se por aplicar os critérios de exclusão citados anteriormente, contudo, na fase de download e leitura alguns não foram possíveis ser baixados ou não tinham ligação com projeto. Por fim, terminamos com aproximadamente 30 artigos relacionados, através dos quais, foram retiradas outras referências para trabalhos não capturados pela pesquisa inicial.

É interessante observar que apesar deste processo ter obtido somente 30 artigos que atendiam aos requisitos, foram incluídos diversas outras referências por outros meios como indicações de grupos científicos e leituras relacionadas posteriormente à busca inicial.

3.2 Autoria Mulsemídia em 2D

3.2.1 Saleme et al.

Em [54], Saleme et al. apresentam uma abordagem orientada a eventos para integrar efeitos sensoriais em ambientes interativos. A abordagem utiliza o padrão MPEG-V para descrever os efeitos sensoriais e aborda as interações do usuário. Além disso, os autores

propõem uma descrição em XML para estabelecer a associação entre as interações do usuário e as ações do aplicativo. Eles também especificam transições entre contextos para determinar qual ação deve ser desencadeada, considerando que eventos semelhantes podem resultar em ações diferentes dependendo do contexto.

Embora a proposta em [54] lide com eventos do usuário, a abordagem não permite que os autores desenvolvam um aplicativo mulsemídia interativo usando apenas uma linguagem declarativa, o que já era previsto, dado que MPEG-V não prevê todas as variedades de conteúdo e evolução de efeitos sensoriais existentes. Por outro lado, esta tese propõe uma abstração dos conteúdos, tratando em nível de autoria todo conteúdo, seja mídia tradicional ou efeito sensorial de maneira semelhante e com execução transparente ao usuário, cabendo ao autor somente incluir o conteúdo e montar o cenário. O ambiente de autoria e execução cuida da administração e entrega do conteúdo ao usuário final.

A proposta apresentada no artigo de Silva et al. [56] sobre livros multissensoriais é extremamente interessante e pode ter um grande impacto no aumento do interesse das crianças pela leitura. A ferramenta MBook consegue separar o conteúdo adicional do livro, tornando mais fácil a sincronização e a renderização, sem necessidade de alterar o conteúdo do livro. Além disso, a ferramenta permite uma sincronização refinada por meio da captura da posição de leitura usando um rastreador ocular. No entanto, é importante ressaltar que as limitações da pesquisa são que a proposta depende muito da imaginação do usuário em conjunto ao efeito sensorial e que a mídia é apresentada em duas dimensões, limitando a experiência sensorial total.

Em Saleme et al. [52], os pesquisadores propõem MulseOnto, uma abordagem para a engenharia de software para o desenvolvimento de sistemas mulsemídia imersivos. Eles argumentam que a proposta pode ser usada como um modelo conceitual consensual para explorar o conhecimento sobre a cadeia completa de sistemas multimídia, prevenindo possíveis equívocos que podem eventualmente levar à construção de soluções que não levam em consideração o reúso, integração e padronização. Essa pesquisa é importante para este trabalho, pois destaca a necessidade de padronizações para o desenvolvimento de conteúdos mulsemídia, que ainda não é padronizado devido a sua relativa novidade.

Em outro trabalho [53], Saleme et al. destacam os desafios fundamentais para fornecer efeitos multissensoriais em sistemas heterogêneos e propõem uma estrutura multimídia interoperável para lidar com esses desafios, atendendo aos requisitos emergentes. Eles relatam que, para isso, foi necessário mudar o núcleo de desenvolvimento seguindo padrões de arquitetura e design para acomodar diferentes perfis de comunicação, conectividade

e padrões de metadados de efeitos sensoriais de acordo com a necessidade de aplicativos e dispositivos multimídia disponíveis no ambiente do usuário. Assim como em [52], os autores demonstram a necessidade de padronização no desenvolvimento de conteúdo mulsemídia e apresentam o PlaySEM como ferramenta para demonstrar a utilização de conteúdo mulsemídia e PlaySEM SER com base em XML para implementação desse conteúdo de forma interoperável.

3.2.2 Guedes et al.

Guedes et al. [34] propõem uma estrutura de programação de alto nível com suporte a interfaces de usuário multimodais em aplicações multimídia. A estrutura integra diferentes tipos de modalidades de entrada e saída, como gestos, reconhecedores de voz, conteúdo audiovisual tradicional, sintetizadores de voz e atuadores. Os autores também destacam a possibilidade de implementar efeitos sensoriais específicos e limitados ao contexto proposto. Isso permite a flexibilidade de criação de aplicações multimídia interativas e imersivas, com suporte a diferentes modalidades de interação e saída.

Em seu trabalho, é utilizada a SRGS (*Speech Recognition Grammar Specification*)¹, que é uma especificação XML para definir qual fala deve ser reconhecida. Estes arquivos definem as âncoras do reconhecedor para especificar a parte do conteúdo a ser executado. Para permitir a sincronização entre reconhecedores e sintetizadores, a estrutura define quatro tipos de comandos de ação: "Start" para iniciar, "Stop" para interromper, "Pause" para pausar e "Resume" para retomar. Além disso, a ferramenta utiliza classes de perfil de usuário com o objetivo de identificar os perfis dos usuários e adaptar os aplicativos de acordo com esses perfis. Outro destaque é que a ferramenta dá suporte à linguagem NCL (*Nested Context Language*)² [5] para prover interoperabilidade entre sistemas.

Como ponto forte, pode-se citar que Guedes et al. [34] propõem como solução de interoperabilidade a linguagem NCL, que é aberta e de fácil manipulação, similarmente à MultiSEL. No entanto, como desvantagem, a proposta apresentada oferece baixa abstração, representando, por exemplo, atuadores como nós dentro de um documento, assim como os demais conteúdos de mídia, o que exige do autor um domínio pleno do documento para ajustes de conteúdo. Isso pode ser um problema para autores menos experientes. Além disso, a proposta de [34] também dificulta a representação de eventos de interação do usuário, pois um componente de software deve ser implementado para cada novo

¹<https://www.w3.org/TR/speech-grammar/>

²<http://www.ncl.org.br/>

evento de interação que um aplicativo suportar.

3.2.3 Mattos et al.

Em 2021 Mattos et al. [25], os autores propõem uma revisão das principais ferramentas de autoria multimídia existentes, visando incentivar a pesquisa e o desenvolvimento de novas soluções para melhorar a criação de conteúdo multimídia. O artigo apresenta uma análise detalhada das características e recursos dessas ferramentas, bem como uma lista de desejos para o desenvolvimento futuro dessas ferramentas. Além disso, os autores discutem o estado atual da arte na área de mulsemídia e apresentam desafios abertos e direcionamentos para a pesquisa futura no campo.

Mattos et al. propuseram alguns trabalhos relevantes para esta pesquisa, dentre os quais pode-se citar a ferramenta STEVE 2.0 [26] (Figura 3.1), que é um editor gráfico que utiliza o seu próprio modelo baseado em eventos, chamado de MultiSEM (*Multimedia Sensory Effect Model*), para representar aplicações multimídia e sincronizar seus nós, sejam eles mídias tradicionais ou efeitos sensoriais. Esse modelo utiliza um paradigma de sincronização temporal baseado em eventos, que permite aos usuários do STEVE definir interatividade e relações temporais entre itens de mídia. Uma vez que a ferramenta fornece uma visão temporal como abordagem, os usuários não precisam ter habilidades de programação para definir uma apresentação multimídia usando o ambiente STEVE 2.0. Ele também fornece feedback aos autores sobre inconsistências de sincronização temporal e fornece extração automática de efeitos sensoriais através da análise de conteúdo audiovisual, aplicando métodos baseados em aprendizado de máquina [1].

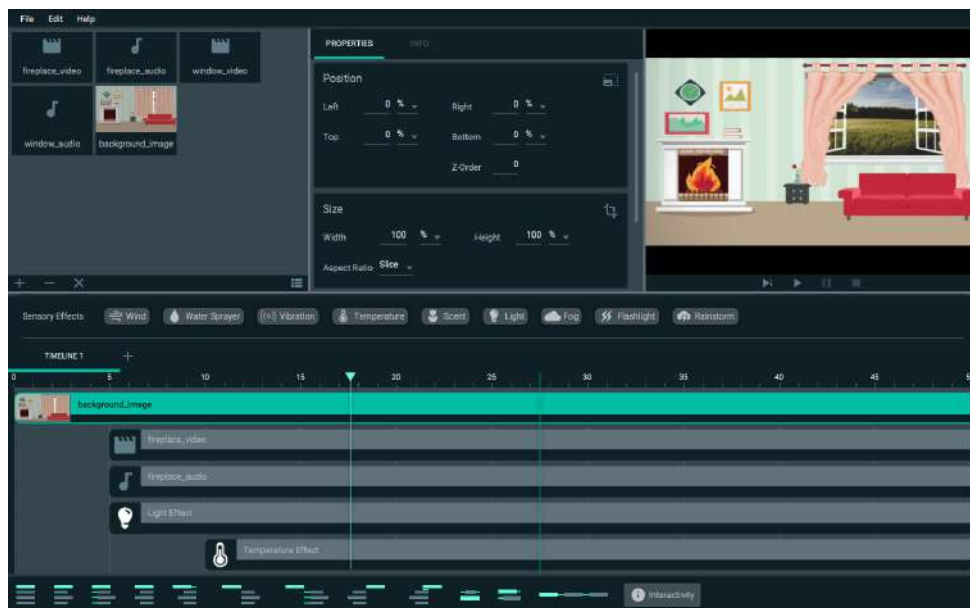


Figura 3.1: STEVE 2.0 de Mattos et al. [26].

O STEVE 2.0 [26] fornece suporte a NCL como formato de exportação, o que permite sua execução em reprodutores de multimídia, no entanto, ainda não é capaz de explorar conteúdo em 360 graus ou em realidade virtual. Em relação ao modelo subjacente do STEVE 2.0, o Modelo MultiSEM [26] representa os efeitos sensoriais como nós, seguindo a abordagem apresentada em [39], que modela os efeitos sensoriais como entidades de primeira classe. Essa abordagem de representação utiliza uma abstração de alto nível, permitindo que a sincronização espacial e temporal dos aplicativos multimídia seja especificada independentemente dos dispositivos usados para implementá-los no mundo real.

Os trabalhos de Mattos et al. (2021) [25], apesar de serem originariamente direcionados para conteúdos 2D, com a utilização do modelo MultiSEM [26], possuem a possibilidade de suportar conteúdos 360° e efeitos sensoriais para realidade virtual, através da linguagem MultiSEL [25], baseada em XML. Como apresentado no Capítulo 2, MultiSEL oferece versatilidade e adaptação a novos tipos de mídia, assim como suporte a efeitos sensoriais, diferentes cenas mulsemídia, hierarquização, interações e até relacionamentos entre mídias. Além disso, é fácil adaptar conteúdos 2D para 3D. Devido a essas características, esta tese optou por adotar a linguagem MultiSEL como formato de interoperabilidade no ambiente de autoria imersivo AMUSEVR.

3.2.4 Linguagem NCL 4.0

Sobre trabalhos relacionados a autoria em 2D que utilizem a linguagem NCL 4.0, destaca-se a tese de doutorado de Josué [38]. Esse trabalho aborda o enriquecimento de aplicações multimídia com efeitos sensoriais para fornecer experiências imersivas de alta qualidade aos usuários. Essas aplicações são chamadas de mulsemídia e permitem a sincronização temporal entre o conteúdo audiovisual e os efeitos sensoriais que estimulam outros sentidos como tato, olfato e paladar. Como exemplo de aplicação mulsemídia utilizando NCL 4.0, o trabalho [38] apresenta o caso de uso ilustrado na Figura 3.2, onde o conteúdo audiovisual 2D é sincronizado com efeitos sensoriais de luz e vento.

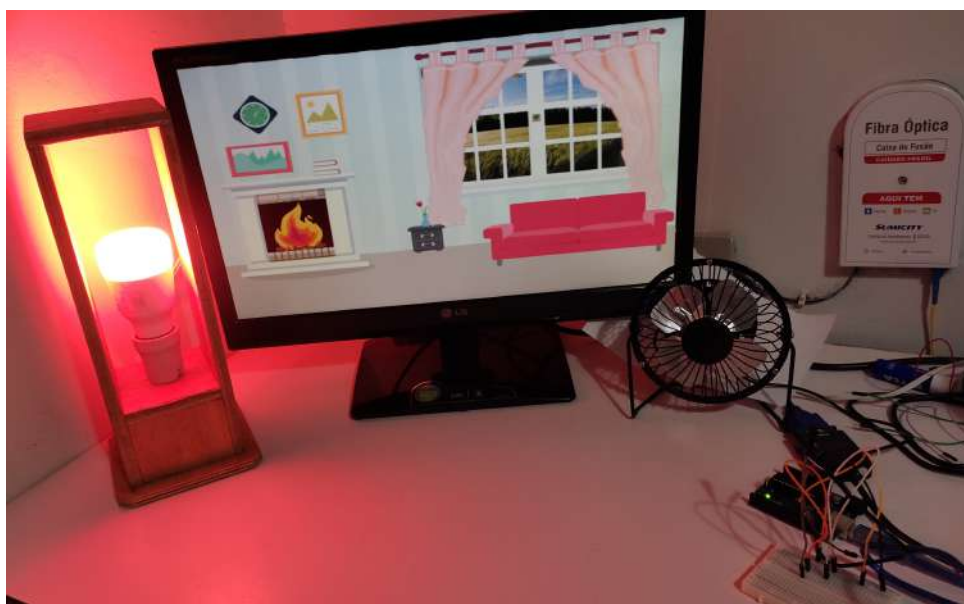


Figura 3.2: Ambiente de execução da aplicação mulsemídia utilizando NCL 4.0 [38].

Para evitar falhas de sincronização durante a apresentação de aplicações mulsemídia, a mesma tese [38] propõe uma operação de preparação para objetos de mídia e efeitos sensoriais. Esta operação pode ser realizada na fase de autoria ou execução do documento mulsemídia e é feita automaticamente pelo formatador mulsemídia baseado nas condições da rede e no comportamento temporal da aplicação. A tese também propõe uma extensão à linguagem NCL e ao modelo NCM para permitir a criação de aplicações NCL com efeitos sensoriais e apresenta alguns casos de uso como a preparação de objetos de mídia para propaganda direcionada e efeitos sensoriais de luz e aroma. Essa tese contribuiu na construção e concepção dos módulos de criação de dos projetos MultiSEL intercambiáveis e execução de efeitos sensoriais do AMUSEVR.

Outro trabalho que aborda a linguagem NCL 4.0 que esta pesquisa teve como estudo

foi [4]. O artigo descreve uma proposta de extensão para o middleware Ginga-NCL 4.0 da TV digital no Brasil para suportar interação multimodal. A interação multimodal se refere ao uso de vários modos de entrada para interagir com o sistema, como fala e gestos, diferentemente do AMUSEVR que trabalha com interação simples por apontamento e clique. O objetivo é permitir que os usuários possam interagir com as aplicações de TV digital de diferentes maneiras, usando dispositivos de entrada que sejam adequados às suas necessidades específicas.

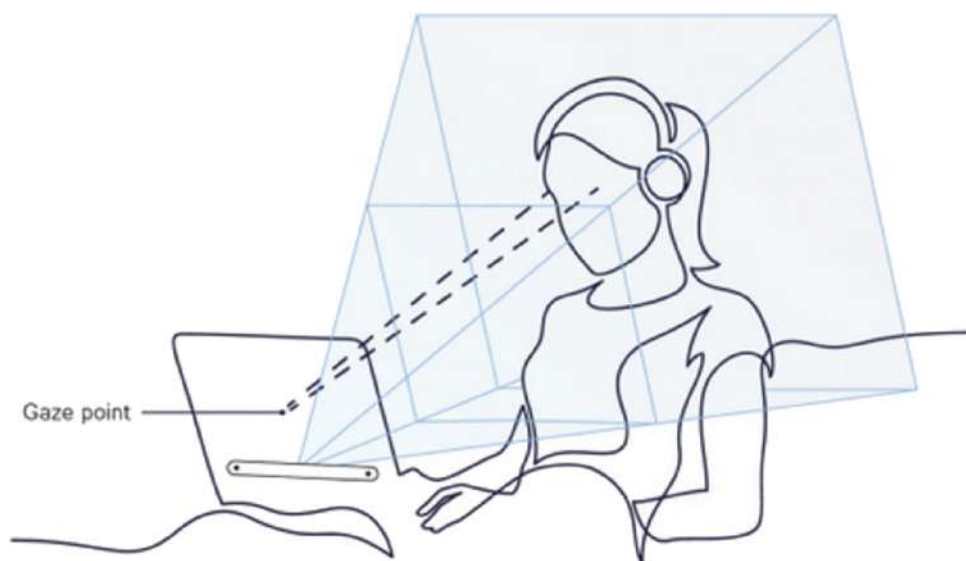


Figura 3.3: Exemplo de interação com os olhos apresentado em [4].

Para alcançar esse objetivo em [4], a linguagem *Nested Context Language* (NCL) na versão 4.0 é estendida para suportar novos eventos de interação e múltiplos usuários. Dois novos módulos são desenvolvidos como prova de conceito, um que suporta interação por voz e outro que suporta interação por olhar, como ilustrado na Figura 3.3. Várias aplicações de TV digital são criadas para avaliar a proposta, e o artigo afirma que a extensão não causou sobrecarga no processamento de eventos do Ginga-NCL. Finalmente, um comparativo é feito entre a versão atual do NCL e a proposta estendida do NCL 4.0, levando em consideração o desempenho e o número de linhas de código da aplicação. Com base nesse experimento comparativo utilizando linhas de código, foi realizada nesta tese também uma comparação similar.

3.3 Outros Trabalhos de Autoria Imersiva em 2D

Como exemplo de ferramenta de autoria que utiliza a Realidade Aumentada (RA) voltada para educação, temos o trabalho de Fraga e Mendes de 2017 [32]. Neste artigo, os autores

propõem a ferramenta autoral A3RA, que trabalha com uma abordagem clássica, sendo utilizada em navegadores web 2D. Com o objetivo de democratizar essa tecnologia e incorporá-la no trabalho pedagógico, a ferramenta adiciona informações educacionais a conteúdos em 360 graus. Esse conteúdo adicional é posicionado como pontos de interesse em meio a um conteúdo real, que é capturado e exibido através de câmeras e posicionado através da geolocalização de equipamentos móveis, funcionando de maneira semelhante aos marcadores presentes na navegação do Google Maps.

O artigo de Reeves et al. de 2004 [49] investiga, sob um enfoque bibliográfico, a temática da interação multimodal, com o objetivo de apresentar o que é, como surgiu e qual é a finalidade deste tipo de interação, bem como os diferentes contextos de uso. O estudo também destaca a importância das interfaces se adaptarem às necessidades e habilidades dos diferentes usuários, além da identificação e otimização de interfaces para o usuário, a fim de proporcionar a aceitação das ferramentas. Como se trata de um guia focado na interação, não apresenta conteúdo específico sobre ferramentas de autoria.

O Google Tour Creator³ é uma ferramenta popular não acadêmica que se concentra na produção de ambientes 360°. Ele tem sido utilizado em diversos trabalhos acadêmicos, como ferramentas educacionais [8] e similares ao Google Street View⁴. Ele foi projetado para fornecer apresentações baseadas em passeios pelo Google Maps. Embora seja versátil e possibilite conteúdo interativo, ele não suporta efeitos sensoriais e a produção e edição via navegadores são limitadas. Além disso, é importante mencionar que o Google anunciou o fim do suporte da ferramenta em 30 de junho de 2021⁵.

3.4 Autoria Imersiva em 3D

3.4.1 Coelho et al.

Em [11], Coelho et al. propuseram uma ferramenta de autoria multissensorial para vídeos 360°, que permite ao autor editar o conteúdo vinculado ao vídeo com um nível básico de personalização. Essa ferramenta de autoria permite a edição e visualização de vídeos 360° com a particularidade de permitir que o vídeo seja complementado com outros estímulos sensoriais, como efeitos táteis e olfativos. Além disso, todo o conteúdo é gerenciado através de uma interface desktop baseada em linhas de tempo, em uma interface de edição

³<https://arvr.google.com/tourcreator/>

⁴<https://www.google.com/maps/>

⁵<https://support.google.com/edu/expeditions/answer/10086878?hl=en>

bidimensional (como visto na Figura 3.4), semelhante às ferramentas de edição comerciais de vídeo no mercado, porém, com visualização do produto através de óculos VR.



Figura 3.4: Ferramenta de autoria bidimensional de Coelho et al. [11].

No trabalho de 2019 de Coelho et al. [12], os autores propõem uma melhoria da ferramenta de autoria, agora se concentrando em três interfaces de autoria diferentes:

- Desktop - a mesma ferramenta de autoria bidimensional mencionada em [11];
- Imersiva - utiliza a interface desktop como base, mas oferece alguns ajustes utilizando uma interface de realidade virtual; e
- Háptica - utiliza a interface desktop como base e parte do conceito da interface imersiva, mas utiliza uma maquete física da interface vista dentro do ambiente de realidade virtual, com o objetivo de oferecer resistência física para alavancas e botões através de equipamentos como servomotores, esperando aumentar a imersão do autor.

Já em [13], os autores atualizaram a ferramenta base de 2018 para fornecer suporte à autoria colaborativa. Nesse trabalho, além das características já mencionadas em [11], os autores se concentraram em prover uma autoria de forma colaborativa, utilizando técnicas como semáforos para controlar as permissões de acesso, gerenciamento de conteúdo e versionamento. É necessário o uso de um servidor em rede. Quanto a essa proposta, pode-se dizer que os autores realizaram experimentos adequados para testar os estudos de caso propostos nos três trabalhos, mas todas as versões de interface propostas dependem de uma interface desktop. Essa interface se baseia nas tradicionais ferramentas de autoria de

vídeo com o paradigma de linha do tempo, que são eficazes para conteúdos bidimensionais, mas como o próprio trabalho deles [12] demonstra, essa abordagem é insuficiente, confusa ou até mesmo incompatível para abordar conteúdos completamente dentro de ambientes imersivos.

3.4.2 Saleme, Covaci, Comsa, Ghinea et al.

Saleme et al. [51] destacam que as empresas estão propondo novas adições ao mundo multissensorial e revelando novos produtos que prometem oferecer experiências incríveis explorando a mídia multissensorial, onde os usuários podem perceber odores, sabores e a sensação do vento soprando em seu rosto. Embora pesquisadores, profissionais e usuários se deparem com uma ampla gama desses novos dispositivos, relativamente pouco trabalho foi realizado para resumir os esforços e iniciativas nesta área. Como proposta de [51], os autores oferecem um tutorial para construir um ambiente multissensorial. Apesar de não apresentar um produto específico, o artigo aponta vários caminhos para criar ferramentas imersivas para conteúdo multissensorial, se tornando um ótimo material de consulta. Além disso, os autores citam a necessidade de evitar a dependência do MPEG-V, devido a *royalties* e muita abstração de hardware que dê suporte a conteúdos multissensoriais.

O artigo de Comsa et al. [14] destaca a dificuldade de entregar Qualidade de Serviço (QoS) enquanto habilita o provisionamento de Qualidade de Experiência (QoE) para os usuários. Para isso, eles propõem uma aplicação multimídia com foco em conteúdo em 360 graus e propõem enriquecer o conteúdo com a adição de efeitos sensoriais. Eles apresentam uma estrutura conceitual para entrega de multimídia 360 graus e um protótipo que permite aos usuários experimentar conteúdo multimídia equirretangulares projetados em 360 graus. Após avaliações de uso com usuários, concluíram que resoluções de vídeo mais altas não levam necessariamente aos níveis mais altos de QoE quando complementado com outros efeitos, confirmando que há benefícios na imersividade ao adicionar efeitos sensoriais ao conteúdo. Apesar de ser interessante, a ferramenta possui algumas limitações como os efeitos sensoriais baseados em conteúdos bidimensionais em um contexto tridimensional com fundo em 360 graus. Nesta situação, é possível que essa limitação seja explorada por seus desenvolvedores, pois a ferramenta foi desenvolvida com apenas 3 graus de liberdade de movimentação (da sigla 3DoF em inglês) (ver Figura 2.7, assim, o usuário não terá problemas em explorar limitações das mídias bidimensionais adicionadas ao projeto, mas que seriam facilmente percebidas em um contexto que o usuário tivesse 6 graus de liberdade de movimentação (da sigla 6DoF em inglês).

Em Covaci et al. [20], os autores se concentram em explicar o que é mulsemídia e descrever o estado da arte do tema, citando diversas aplicações em diferentes categorias de conteúdo. Como se trata de um texto de revisão literária sobre o tema mulsemídia, os autores apontam as limitações encontradas pelas ferramentas existentes e mencionam que há poucos trabalhos com enfoque em conteúdo multissensorial, ou seja, vários tipos de mídia envolvendo múltiplos sentidos. Além disso, o texto aborda a evolução dos desafios identificados anteriormente nesta área e apresenta novas direções de pesquisa. Este trabalho oferece uma boa base para a pesquisa, especialmente porque é um dos poucos artigos que aborda técnicas de produção de conteúdo para efeitos sensoriais, que é o tema abordado nesta tese.

Em Covaci et al. [19] de 2019, os autores continuam a pesquisa descrita em Comsa et al. [14], na qual os resultados relatam que adicionar mídia multissensorial ao conteúdo audiovisual tradicional tem um efeito positivo na Qualidade da Experiência do Usuário (QoE). No entanto, os autores observaram que faltava um estudo com foco em mídias 360 graus. Consequentemente, nesse artigo, foi realizado um estudo de QoE baseado em assistir a um vídeo 360 graus - com e sem efeitos multissensoriais - em uma configuração de VR com 3 graus de liberdade, como exibido na Figura 2.7. Ao final da pesquisa, os autores revelam que, assim como em vídeos tradicionais, para os vídeos 360 graus, quando o conteúdo que era oferecido na rede perdia qualidade, a variação era menos sentida pelo usuário quando o usuário estava recebendo complementação de informações sensoriais diversas, como odores e vento. Esse trabalho, assim como o anterior, também discute a distribuição de conteúdo de efeitos sensoriais baseadas em matrizes bidimensionais, onde cada célula da matriz (*tile*) entrega informações sensoriais independentes como intensidade. Nesta tese, não está sendo usada essa abordagem de matrizes bidimensionais, pois optou-se por aplicar o 6DoF, o que demandaria uma matriz tridimensional, sendo um interessante trabalho futuro.

3.4.3 Horst et al.

Em Horst et al. [35], os autores descrevem um problema comum em apresentações de mídia em 360 graus, exibidas em museus, parques e ambientes similares, que é a dificuldade de guiar o usuário através desses ambientes dinâmicos em realidade virtual. Para abordar este problema, os autores propõem técnicas para destacar e realçar objetos contidos em vídeos em 360 graus. Embora este trabalho se concentre em identificar conteúdo dentro de mídia, ele não contribui diretamente para esta pesquisa de tese, os autores destacam a

independentemente dos outros visualizadores. A solução foi desenvolvida para serviços de entretenimento e apresentação cultural baseados em telepresença. Apesar de a implementação ter sido desenvolvida para equipamentos móveis da geração de *smartphones* Symbia com redes 3G da época e ter foco em distribuição de conteúdo interativo panorâmico, a aplicação compartilha da exibição de conteúdos multimídia dispostos dentro de ambientes em 360 graus, como praticado nesta pesquisa, servindo como uma demonstração de aplicabilidade das técnicas.

No artigo [61] de 2019, Takacs et al. descrevem os quatro desafios fundamentais enfrentados pelas produções complexas de Realidade Virtual (RV): gerenciamento de múltiplas câmeras, controle de qualidade, anotação automática de cinematografia e estimativa de profundidade de mídias em 360 graus. Para resolver esses desafios, os autores propõem uma solução integrada chamada Hyper 360, que utiliza técnicas de inteligência artificial para melhorar o conteúdo. A solução discrimina objetos automaticamente e utiliza diversas técnicas para melhorar o conteúdo em 360 graus, como ilustrado na Figura 3.6. Este artigo se concentra na qualidade da produção multimídia para ambientes de realidade virtual, especificamente no pré-processamento do conteúdo, preocupando-se com a qualidade de mídia em 360 graus, áudio 3D, entre outros. Enquanto isso, esta pesquisa se concentra na fase posterior ao pré-processamento do conteúdo.

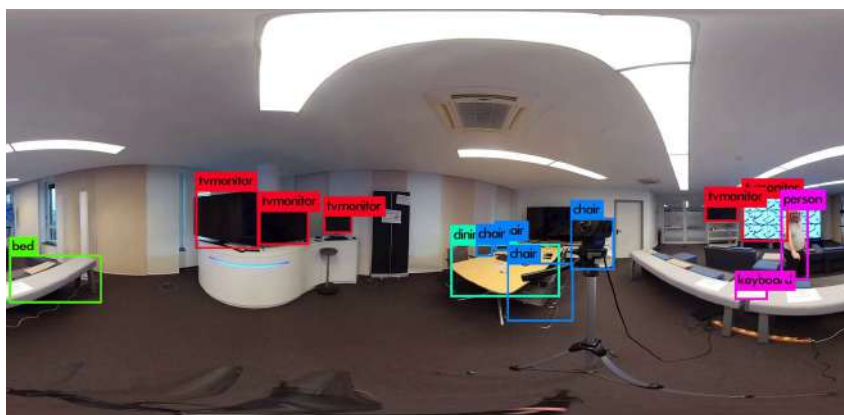


Figura 3.6: Resultado da extração de objetos em mídias de 360 graus de Takas et al. [61].

No artigo [31], Fassold et al. descrevem a necessidade de automatizar várias tarefas repetitivas na produção de vídeo 360 graus, que são tediosas e demoradas quando feitas manualmente. Eles propõem ferramentas automatizadas para cinematografia e para aparelhos de TV convencionais com suporte a conteúdo em 360 graus. Essas ferramentas utilizariam técnicas de inteligência artificial para identificar conteúdos em vídeos 360 graus e vincular recursos automaticamente como texto ou imagens/vídeos 2D. Este artigo

destaca o problema e a dificuldade da autoria de conteúdo baseado em mídias 360 graus. No entanto, toda a parte de autoria é baseada em ferramentas clássicas de interface 2D para desktop e navegadores, além de fazer parte do pré-processamento de conteúdo.

Por fim, no artigo [60], Takacs et al. apresentam sua última versão de um kit de ferramentas de pré-processamento de conteúdo para apresentações em 360 graus, baseado em inteligência artificial. Essa versão inclui melhorias na identificação de conteúdo, proporcionando uma maior qualidade e sensação de profundidade dos objetos inseridos ou identificados no pré-processamento da apresentação não interativa.

3.5 Outros Trabalhos de Autoria Imersiva em 3D

No artigo de Zhang et al. [65], é apresentado o FlowMatic, uma ferramenta de autoria imersiva que aumenta a imersão ao permitir que os programadores especifiquem comportamentos reativos, como ações do usuário, temporizadores do sistema ou colisões. A ferramenta também apresenta primitivas de objetos virtuais para criar e destruir programaticamente novos objetos, para abstrair e reutilizar funcionalidade e para importar modelos 3D. É importante ressaltar que o FlowMatic usa representações visuais para permitir que essas primitivas sejam representadas diretamente em VR. Os resultados de um estudo com usuários também são descritos, ilustrando as vantagens de usabilidade do FlowMatic em relação a uma ferramenta de autoria 2D. Apesar de ser interessante como referência, a ferramenta é um exemplo de autoria para construção de ambientes virtuais focados em ensino e não tem suporte a conteúdos 360 graus, focando em metodologia de desenvolvimentos de fluxogramas como visto na Figura 3.7.

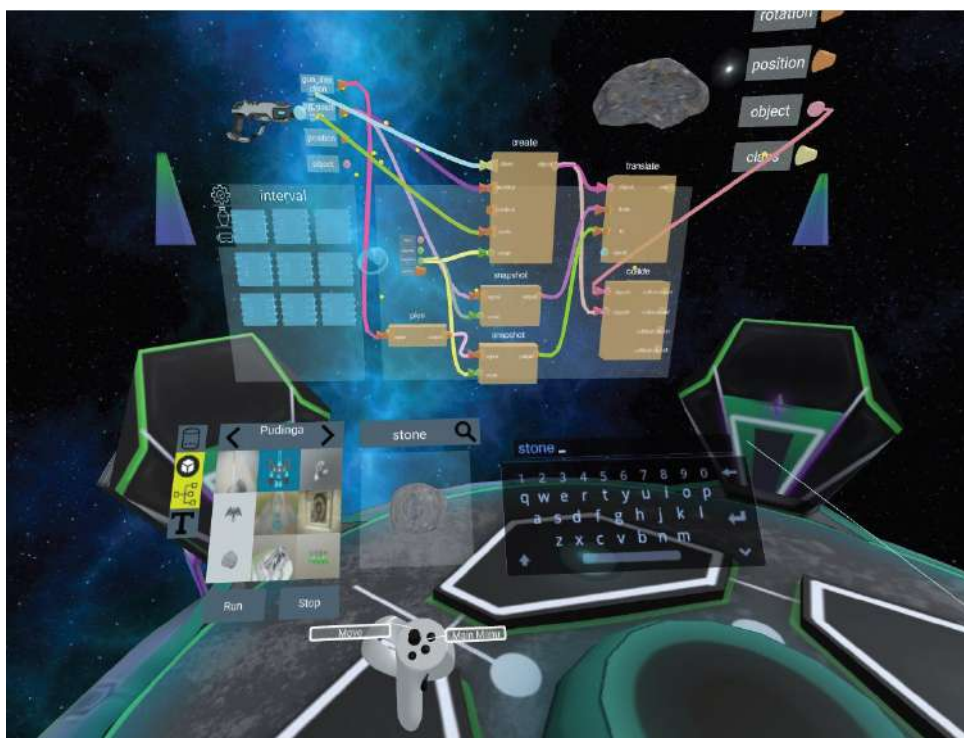


Figura 3.7: Sistema FlowMatic de Zhang et al. [65].

No artigo de Silva et al. [22], é apresentada uma investigação bibliográfica sobre a temática da realidade virtual, com o objetivo de pesquisar o que é a realidade virtual, como surgiu e qual a finalidade desta tecnologia moderna, além de explicar sua importância no mundo globalizado e moderno. A metodologia da pesquisa é descrita inicialmente, seguida de uma análise das teorias envolvendo a realidade virtual e o que ela representa, visando compreender o mundo imersivo e interativo buscado ao utilizar tal tecnologia. É destacado que a realidade virtual pode proporcionar três formas diferentes de interação: passiva, exploratória e interativa. O artigo também apresenta uma seção sobre o valor da realidade virtual em diversas áreas, demonstrando as aplicações dessa ferramenta e as vantagens geradas pela simulação da realidade para o aprimoramento de experiências, além das várias formas de acessá-la e utilizá-la e como é a sua integração com o mundo real. A temática da autoria, mídias em 360 graus ou efeitos sensoriais não são abordadas no texto.

Em relação a outros estudos que suportam vídeos 360°, os autores de [44] propõem um modelo de autoria declarativa que permite aos autores projetarem e visualizarem vídeos 360° interativos (como ilustrado na Figura 3.8), com foco em sua própria linguagem declarativa baseada em XML. Como sua solução foi projetada para conteúdo de mídia com projeção interna em um objeto esférico, o ambiente de RV apresenta limitações de movimentação caso a aplicação possua 6 graus de liberdade. Durante a apresentação, o

usuário é inserido dentro de uma esfera de raio limitado e caso o usuário ultrapasse esse raio, a apresentação pode ser comprometida. Um ponto a ser destacado sobre o trabalho é que a ferramenta de autoria não permite edições de forma imersiva, limitando-as a edições em seu formato proprietário baseado em XML. Quando em modo visualização, permite interações com o conteúdo, assim como nesta tese.



Figura 3.8: Execução da ferramenta autoria em modo visualização de Mendes et al. [44].

No artigo de Chen Li et al. de 2020 [64], é apresentado um estudo que aborda técnicas de pós-processamento em mídias 360°, com o objetivo de obter melhor compressão e manter boa qualidade para o usuário. Para isso, os autores apresentam os diferentes tipos de mídias em 360°, tipos de projeção, resoluções e formatos, sendo um material útil como referência. Outro destaque do trabalho é o estudo sobre o uso da câmera e o melhor aproveitamento do campo preferencial de interesse da visão do usuário. No entanto, o artigo não se concentra nas características da sua ferramenta autoral.

No artigo de Park e Han de 2018 [47], é apresentada uma ferramenta de autoria de conteúdo em realidade virtual com suporte a efeitos sensoriais. Os autores adotaram o padrão MPEG-V como meio para representar e comunicar os efeitos sensoriais no conteúdo de realidade virtual apresentado. A proposta é um complemento (*plug-in*) para Unity, que utiliza os menus do motor gráfico como meio de configurar a apresentação. O complemento oferece suporte a efeitos sensoriais ativos ou passivos, que podem ser ativados por interações do usuário ou automaticamente. Embora o trabalho demonstre a capacidade de comunicação do complemento com os equipamentos de efeitos sensoriais, a parte autoral da ferramenta é bastante limitada e fortemente acoplada a Unity, como mostrado na Figura 3.9.

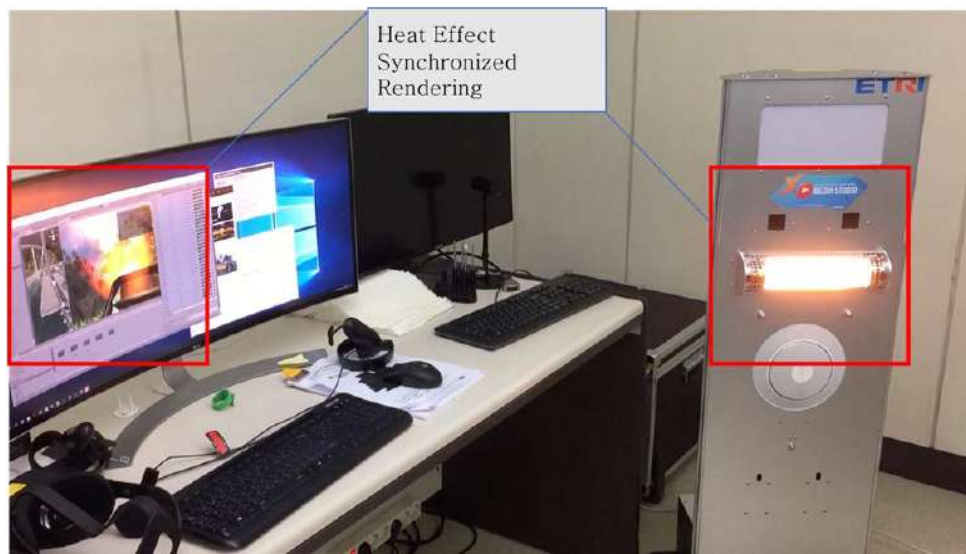


Figura 3.9: Ferramenta de autoria de Park et al. [47].

Em mais um complemento para Unity, a ferramenta de autoria IATK (*Immersive Analytics Toolkit*), de Cordeil et al. de 2019 [17], propõe a criação e exploração interativa de visualização de dados em ambientes imersivos, sendo esses dados obtidos de arquivos em formato *csv* semelhantes aos gerados por ferramentas de planilhas eletrônicas. Com objetivo de favorecer a tomada de decisão e análise de gestão, o usuário imerge em um ambiente virtual onde, através de seus controles, pode manipular a visualização de dados através de rotações, redimensionamentos e troca de gráficos, como ilustrado na Figura 3.10. É possível também criar apresentações para execuções posteriores no modo autoria. A ferramenta não apresenta suporte a conteúdo multimídia além de gráficos de dados.

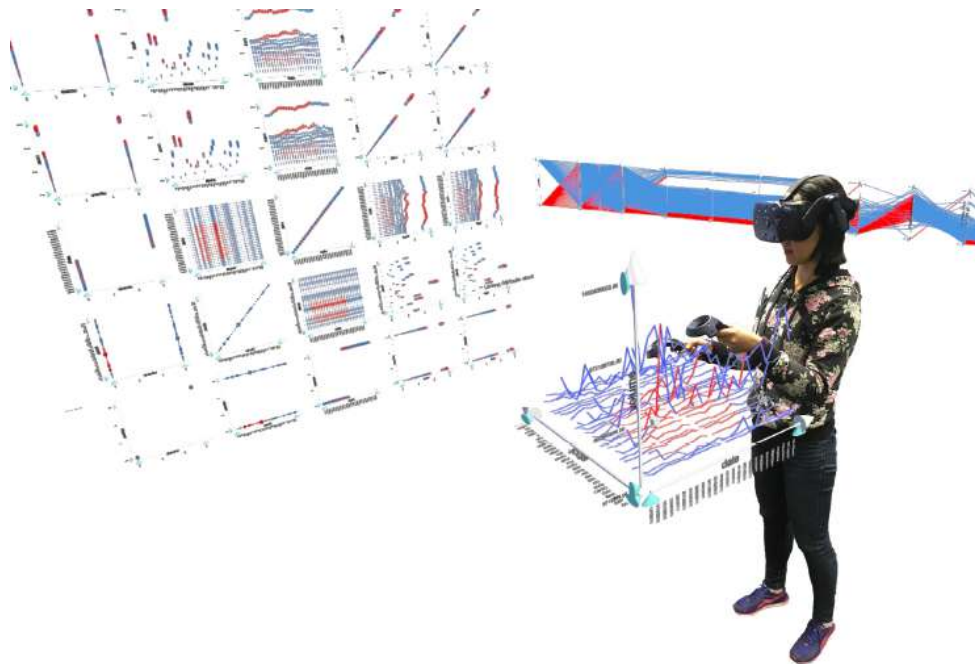


Figura 3.10: Ferramenta de visualização para gestão de Cordeil et al. [17].

Gabriel Souza et al. em [58] apresentam a arquitetura Guaraná, que se baseia no uso de capacetes de realidade virtual (HMDs) como uma "segunda tela". Dessa forma, enquanto uma TV digital apresenta o conteúdo tradicional de uma aplicação, os usuários com HMDs podem assistir a um formato imersivo e até interagir através de controles, com um conteúdo adicional a esta apresentação, como ilustrado na Figura 3.11. O Guaraná provê suporte a mídias bidimensionais, como vídeo, imagens e textos, porém somente suporta uma projeção esférica de vídeo em 360 graus de tamanho de raio limitado, portanto, limita a liberdade de locomoção pelo cenário. Sobre a linguagem adotada, os autores propuseram o formato NCL 360, através de uma adaptação do NCL, a fim de dar suporte ao conteúdo proposto. O modo criador de conteúdo (para autoria) ainda não está disponível.

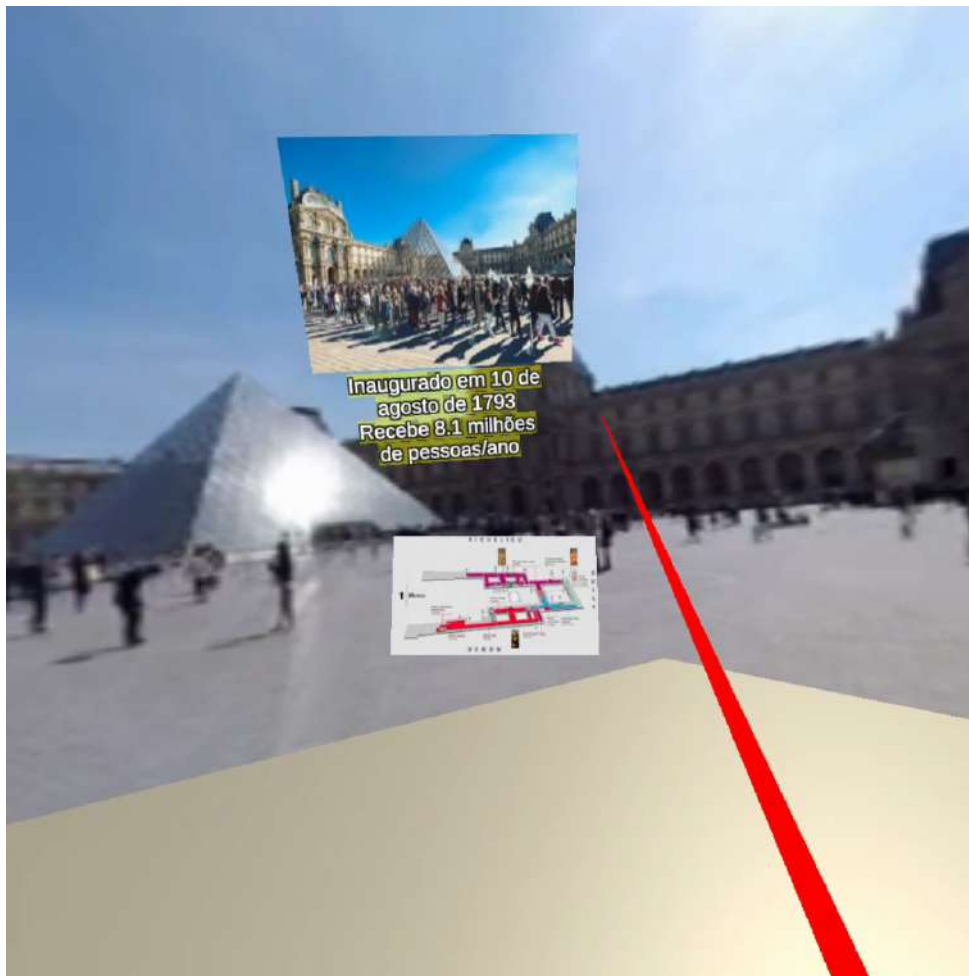


Figura 3.11: Proposta de Souza et al. [58] para segunda tela da TV digital.

3.6 Comparação dos Trabalhos Relacionados

A Tabela 3.2 compara todos os trabalhos apresentados neste capítulo com a proposta desta tese publicada em [27, 28], destacando-a na última linha da tabela. As colunas da tabela significam:

- Conteúdo imersivo em 3D - quando o trabalho possui enfoque em conteúdo imersivo em ambiente 3D;
- Conteúdo interativo - quando o trabalho possui elementos de interatividade em sua proposta;
- Mídia em 360 graus - quando o trabalho aborda apresentações com projeções em pano de fundo com mídias em 360 graus;

- Efeitos sensoriais - quando o trabalho apresenta soluções ou utiliza algum tipo de estímulo a outros sentidos além da visão e audição;
- Ferramenta de autoria - quando o trabalho possui enfoque ou apresenta estudo relacionado a ferramentas de autoria; e
- Edição totalmente imersiva - quando o trabalho aborda alguma ferramenta de autoria, na qual o conteúdo mulsemídia é editado totalmente em modo imersivo; e
- Formato de intercâmbio - quando o trabalho aborda algum arquivo com formato livre a adaptações de contexto mulsemídia, proporcionando intercâmbio com outras ferramentas de autoria.

Tabela 3.2: Resumo e destaque nos trabalhos relacionados

Título	Conteúdo imersivo em 3D	Conteúdo interativo	Mídia em 360 graus	Efeitos Sensoriais	Ferramenta de autoria	Edição totalmente imersiva	Formato de intercâmbio
360° Mulsemmedia: a way to improve subjective QoE in 360° videos [19]	Sim	Não	Sim	Sim	Não	Não	Não
A mulsemmedia framework for delivering sensory effects to heterogeneous systems [53]	Não	Não	Não	Sim	Sim	Não	Não

Continua...

Título	Conteúdo imersivo em 3D	Conteúdo interativo	Mídia em 360 graus	Efeitos Sensoriais	Ferramenta de autoria	Edição totalmente imersiva	Formato de intercâmbio
An Approach for Authoring Mulse-media Documents Based on Events [26]	Não	Sim	Não	Sim	Sim	Não	Sim
An Approach for Authoring Mulse-media Documents Based on Events - Tese [24]	Não	Sim	Sim	Sim	Sim	Não	Sim
An authoring model for interactive 360 videos [44]	Sim	Não	Sim	Não	Sim	Não	Sim
An Event Driven Approach for Integrating Multi-sensory Effects to Interactive Environments [54]	Não	Sim	Não	Sim	Não	Não	Não
Authoring tools for creating 360 multisensory videos—Evaluation of different interfaces [12]	Sim	Sim	Sim	Sim	Sim	Não	Não

Continua...

Título	Conteúdo imersivo em 3D	Conteúdo interativo	Mídia em 360 graus	Efeitos Sensoriais	Ferramenta de autoria	Edição totalmente imersiva	Formato de intercâmbio
Beyond Multi-media Authoring [25]	Não	Sim	Sim	Sim	Sim	Não	Não
Bite-Sized Virtual Reality Learning Applications: A Pattern-Based Immersive Authoring Environment [36]	Sim	Sim	Não	Não	Sim	Sim	Não
Collaborative immersive authoring tool for real-time creation of multisensory VR experiences [13]	Sim	Sim	Sim	Sim	Sim	Não	Não
Deep authoring - an AI Tool set for creating immersive MultiMedia experiences [60]	Sim	Não	Sim	Não	Não	Não	Não
Do I Smell Coffee? The Tale of a 360° Mulsemmedia Experience [14]	Sim	Não	Sim	Sim	Não	Não	Não

Continua...

Título	Conteúdo imersivo em 3D	Conteúdo interativo	Mídia em 360 graus	Efeitos Sensoriais	Ferramenta de autoria	Edição totalmente imersiva	Formato de intercâmbio
Extending multi-media languages to support multimodal user interactions [34]	Não	Sim	Não	Não	Não	Não	Sim
FlowMatic: An immersive authoring tool for creating interactive scenes in virtual reality [65]	Sim	Sim	Não	Não	Sim	Sim	Não
Guidelines for multimodal user interface design [49]	Não	Sim	Não	Não	Não	Não	Não
Highlighting Techniques for 360° Video Virtual Reality and Their Immersive Authoring [35]	Sim	Sim	Não	Não	Sim	Sim	Não
Hyper 360—Towards a Unified Tool Set Supporting Next Generation VR Film and TV Productions [61]	Sim	Não	Sim	Não	Não	Não	Não

Continua...

Título	Conteúdo imersivo em 3D	Conteúdo interativo	Mídia em 360 graus	Efeitos Sensoriais	Ferramenta de autoria	Edição totalmente imersiva	Formato de intercâmbio
IATK: An immersive analytics toolkit [17]	Sim	Sim	Não	Não	Sim	Sim	Não
Immersive edition of multisensory 360 videos [11]	Sim	Sim	Sim	Sim	Sim	Não	Não
Immersive interactive reality: Internet-based on-demand VR for cultural presentation [59]	Não	Sim	Sim	Não	Não	Não	Não
Interactive 360-degree Videos in Ginga-NCL Using Head-Mounted-Displays as Second Screen Devices [58]	Sim	Sim	Sim	Não	Não	Não	Sim
Is multimedia multisensorial? - A review of mulsemedia systems [20]	Sim	Não	Sim	Sim	Sim	Não	Não

Continua...

Título	Conteúdo imersivo em 3D	Conteúdo interativo	Mídia em 360 graus	Efeitos Sensoriais	Ferramenta de autoria	Edição totalmente imersiva	Formato de intercâmbio
Mulsemmedia DIY: A survey of devices and a tutorial for building your own mulsemmedia environment [51]	Sim	Não	Sim	Sim	Sim	Não	Não
REALIDADE VIRTUAL [22]	Sim	Sim	Não	Não	Não	Não	Não
Sensory Effect Authoring Tool for Virtual Reality Simulation [47]	Sim	Sim	Sim	Sim	Sim	Não	Não
State-of-the-art in 360 Video-Image Processing Perception Assessment and Compression [64]	Sim	Não	Sim	Não	Sim	Não	Não
Towards a reference ontology on mulsemmedia systems [52]	Não	Não	Não	Sim	Não	Não	Não

Continua...

Título	Conteúdo imersivo em 3D	Conteúdo interativo	Mídia em 360 graus	Efeitos Sensoriais	Ferramenta de autoria	Edição totalmente imersiva	Formato de intercâmbio
Towards Automatic Cinematography and Annotation for 360° Video [31]	Não	Não	Sim	Não	Não	Não	Não
Um ambiente para autoria e realização de aplicações educacionais com realidade aumentada [32]	Não	Sim	Sim	Não	Sim	Não	Não
Using Multisensory Content to Impact the Quality of Experience of Reading Digital Books [56]	Não	Não	Não	Sim	Não	Não	Não
Virtual reality authoring tools for teachers to create novel and immersive learning scenarios [8]	Sim	Sim	Sim	Não	Sim	Sim	Não
Linguagem NCL 4.0 [4, 38]	Não	Sim	Não	Sim	Sim	Não	Sim

Continua...

Título	Conteúdo imersivo em 3D	Conteúdo interativo	Mídia em 360 graus	Efeitos Sensoriais	Ferramenta de autoria	Edição totalmente imersiva	Formato de intercâmbio
AMUSEVR: A Virtual Reality Authoring Environment for Immersive Multimedia Applications e Immersive Authoring of 360 Degree Interactive Applications [27, 28]	Sim	Sim	Sim	Sim	Sim	Sim	Sim

Considerando os trabalhos apresentados, não há nenhuma solução disponível que ofereça suporte integrado a todas as funcionalidades abordadas nesta tese, tais como: autoria totalmente imersiva em tempo real dentro de um ambiente de realidade virtual, suporte à inclusão de objetos de mídia e efeitos sensoriais em qualquer local do ambiente 3D, suporte à interação com objetos de mídia, transformando a apresentação passiva em interativa, e suporte a um formato que facilita o intercâmbio entre diferentes sistemas, como a linguagem MultiSEL, conforme será discutido nos próximos capítulos.

Capítulo 4

Ambiente de Autoria Totalmente Imersiva

Este capítulo apresenta o ambiente de autoria totalmente imersiva que é a proposta principal desta tese, sendo uma nova abordagem para a criação de conteúdo multimídia baseada em um ambiente imersivo em realidade virtual. Essa abordagem permite que o autor experimente a imersão no conteúdo, com uma sensação semelhante à que será oferecida ao usuário final. A Seção 4.1 apresenta os requisitos do ambiente proposto. A Seção 4.2 apresenta a especificação do ambiente, discutindo sua arquitetura, diagramas de casos de uso e diagramas de classes. A Seção 4.3 discute como o ambiente proposto será avaliado.

4.1 Requisitos do Ambiente de Autoria Totalmente Imersiva

Atualmente, existem diversas ferramentas de autoria multimídia, incluindo algumas que suportam conteúdo multissensorial. No entanto, quando se trata de conteúdo multissensorial imersivo e interativo, tanto a quantidade de opções quanto a variedade de suporte a múltiplos tipos de mídia são reduzidas. Vários fatores contribuem para limitar as poucas propostas existentes. Esses fatores variam desde a dificuldade de implementação de um sistema de gerenciamento e controle de conteúdo funcional até limitações de edição e visualização de conteúdo.

Diante desses desafios, surgiu a motivação para a proposição desta tese. O ambiente de autoria totalmente imersiva apresenta uma abordagem de autoria que supera essas dificuldades, fornecendo uma solução prática para a criação de aplicações multissensoriais imersivas e interativas em ambiente 360.

Levando em consideração alguns trabalhos relacionados citados no Capítulo 3 e outros estudos como [25], bem como nosso objetivo desejado, definimos os seguintes requisitos desejáveis para um ambiente de autoria imersiva e a execução de aplicativos multimedial em 360 graus. Os requisitos relacionados ao conteúdo, layout e cenas 360 graus foram baseados em requisitos propostos em estudos como [44, 13]. No que diz respeito ao requisito de usar um formato intercambiável para importar e exportar projetos, além de ferramentas de sincronização de relacionamentos, fomos inspirados por [1, 44]. Em relação ao suporte a efeitos sensoriais, tivemos algumas ideias com [11, 13]. Com isto, consequentemente, o ambiente deve satisfazer os seguintes requisitos segundo os grupos a seguir:

1. Autoria

- (a) O ambiente imersivo deve oferecer as funcionalidades de autoria e execução de aplicações em 360 graus através de uma plataforma de realidade virtual;
- (b) O ambiente de autoria deve ser oferecido de forma imersiva para que o autor tenha a mesma sensação de imersão que será proporcionada ao usuário final consumidor do conteúdo sendo produzido;
- (c) O ambiente imersivo deve suportar conteúdo de mídias em 360 graus, com opção de exibição;
- (d) O ambiente de autoria deve oferecer suporte para inserir e editar mídias de diversos formatos, incluindo objetos de mídia tradicional (vídeo, áudio, imagem e texto), mídias em 360 graus (vídeo e imagem), efeitos sensoriais (luz, vento, aroma, calor, frio, etc.);
- (e) O ambiente imersivo deve ter suporte a seis graus de liberdade de movimentação do usuário;
- (f) O ambiente deve oferecer opção de pré-visualização do conteúdo produzido (execução de todo conteúdo da cena, sem a necessidade de salvar o projeto), sem que para isso o autor precise sair do ambiente imersivo;
- (g) Oferecer um gerenciador de cenas em 360 graus, para que haja uma melhor divisão e organização de conteúdo segundo a temática da cena e, assim, uma melhor distribuição de conteúdos a serem processados e armazenados na memória do equipamento;
- (h) O ambiente deve oferecer suporte à interatividade nas aplicações a serem criadas;

- (i) Oferecer o conceito de relacionamento entre mídias, a fim de permitir relações temporais entre mídias e cenas;

2. Execução

- (a) O ambiente deve oferecer um modo de visualização das aplicações multissensoriais em 360 graus (equivalente a um *player* do projeto), também em ambiente imersivo;
- (b) Oferecer comandos interativos para prover navegação entre as cenas do projeto e interação com objetos de mídia de uma cena;
- (c) Suportar um gerenciador de conteúdo, permitindo agendar a exibição, execução e ocultação de conteúdo na apresentação;
- (d) Suportar a execução de diversos tipos de mídias simultaneamente;
- (e) Suportar a execução de efeitos sensoriais através de equipamentos atuadores e, complementarmente, a renderização de efeitos animados que simulem visualmente os mesmos efeitos;

3. Interoperabilidade

- (a) Utilizar um formato de intercâmbio do projeto do modo autoria para visualização, baseado em XML (devido à sua flexibilidade e fácil adaptação a novas tecnologias), que possa ser importado de/por outros aplicativos. Este trabalho utiliza a linguagem MultiSEL [24] para esta finalidade; e
- (b) Adaptar e aperfeiçoar o MultiSEL para suportar posicionamento em três dimensões.

Com a combinação destes requisitos, foi possível idealizar um ambiente imersivo, como ilustrado na Figura 4.1. É importante observar que, na ilustração, o usuário (autor ou consumidor do conteúdo) possui liberdade para se movimentar no ambiente virtual, utilizando um dispositivo HMD, e interagir com diferentes objetos de mídia (2D ou 360) e efeitos sensoriais.



Figura 4.1: Cenário ilustrativo.

4.2 Especificação do Ambiente de Autoria Totalmente Imersiva

Seguindo os requisitos apresentados na seção anterior, a Figura 4.2 exibe o diagrama de casos de uso, que contém dois tipos de atores: um para o usuário comum, que está restrito ao uso e visualização do projeto, e o outro ator, o autor-projetista, que possui acesso a todas as funcionalidades necessárias para criar, editar e gerenciar o projeto da aplicação mulsemídia sendo produzida.

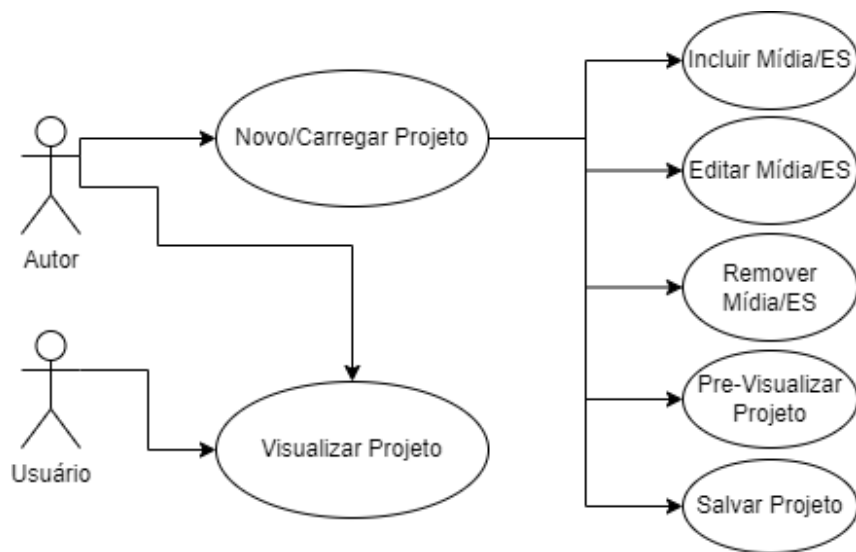


Figura 4.2: Caso de Uso.

De acordo com o requisito de que as mídias possam ser dispostas espacial e temporalmente na apresentação, é proposto o conceito de cenas. Uma cena é uma sequência de ações que acontecem em um lugar específico, enquanto o cenário é o espaço virtual/físico onde essas ações estão acontecendo. Neste projeto, vamos considerar ambos como o ambiente virtual onde acontecem as ações. Cada cena é independente das demais e possui seus próprios objetos de mídia, permitindo ao autor criar múltiplas cenas com diferentes temas, o que permite criar apresentações com diversos cenários no mesmo projeto. Neste contexto, o usuário pode navegar entre as cenas e interagir com os objetos de mídia presentes nelas, além de poder transitar livremente pelos cenários, devido ao projeto ter sido planejado e baseado na tecnologia de 6 graus de liberdade de movimento do usuário.

Para melhor compreensão, a Figura 4.3 ilustra uma cena de 360 graus contendo diferentes objetos de mídia. Na parte central superior, é mostrado um objeto de mídia do tipo imagem 2D, na parte inferior da mesma figura, há um outro objeto de mídia do tipo vídeo 2D, por fim, um terceiro objeto de mídia do tipo vídeo em 360 graus, que é exibido ao fundo da cena.

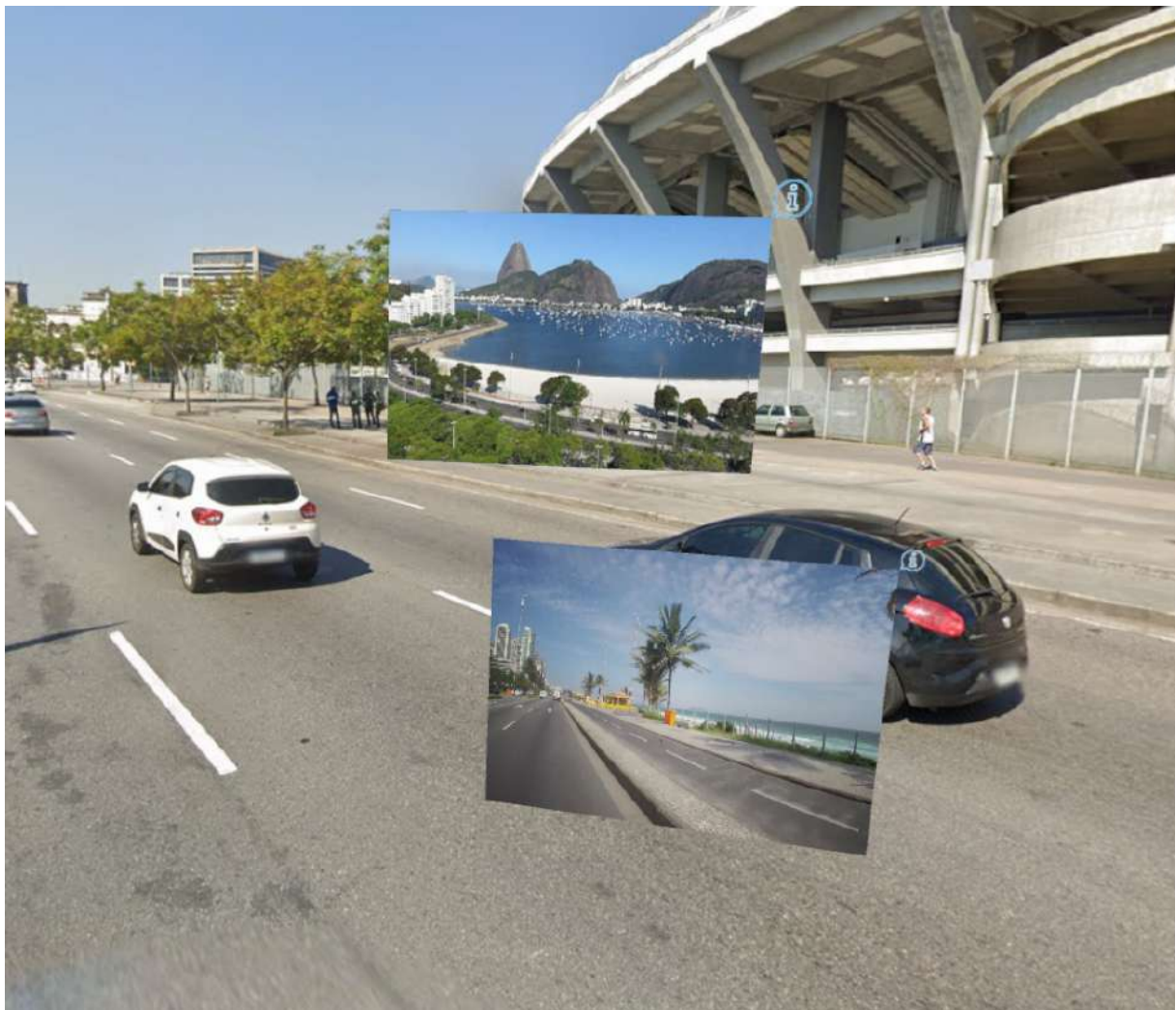


Figura 4.3: Cena 360 graus com diferentes objetos de mídia.

A Figura 4.4 ilustra o diagrama de classes proposto para o modelo de dados a ser utilizado no ambiente de autoria e execução. Um projeto mulsemídia é composto por uma ou mais cenas, que por sua vez são compostas por diferentes tipos de objetos de mídia 2D, 3D e efeitos sensoriais, além de relações (*links*) entre eles. Os *links* permitem especificar relações de sincronização entre dois objetos seguindo um modelo temporal baseado em eventos, como descrito em [26].

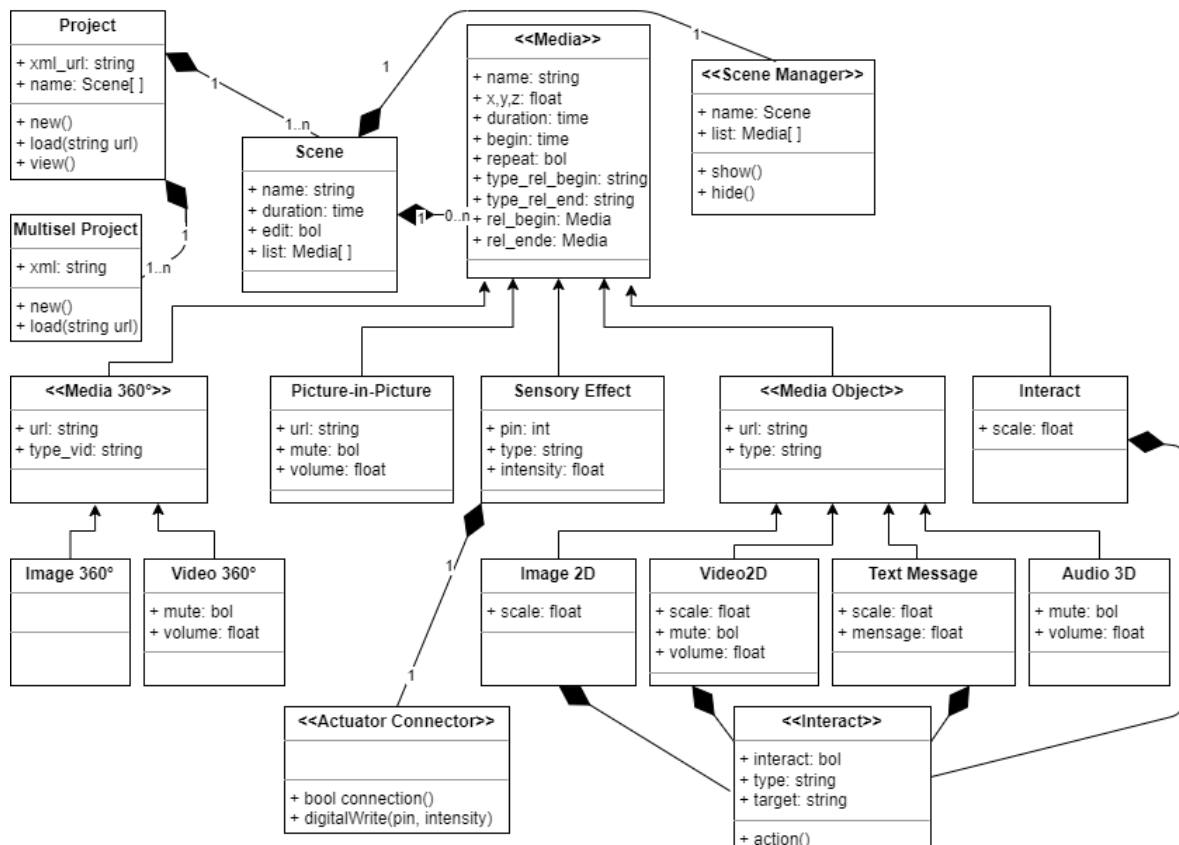


Figura 4.4: Diagrama de classes.

O diagrama de classes possui uma Classe Project que foi projetada para gerenciar, entre outras coisas, a criação e o carregamento do projeto, seja esse projeto novo ou carregado de um projeto em MultiSEL. A Classe Project é responsável por criar as cenas, onde a relação da Classe Scene com Project, representa que cada projeto deve ter uma ou mais cenas, além das Classes Scene e Multisel Project serem dependentes por composição da Classe Project. Relação semelhante acontece com a Classe Media, em que Scene pode ter nenhuma ou muitas mídias, onde a Classe Media também é dependente por composição da Classe Scene.

A classe Media é uma classe abstrata, desenvolvida com o objetivo de proporcionar reusabilidade através do conceito de hierarquia, onde outras classes podem estender suas propriedades. Além da classe Media, outras classes também são abstratas neste projeto, como as classes Media 360, Scene Manager, Media Object, Actuator Connector e Interact.

A Classe Scene Manager utiliza as informações de uma cena e também a lista de todas as mídias contidas nela, com isso, cruzando as informações e montando o cronograma de apresentação de cada mídia em cena, assim como a retirada de cada objeto da cena. Esta classe também é responsável pelo gerenciamento do modelo temporal criado pelos

relacionamentos de todas as mídias e efeitos sensoriais da cena, sincronizando-o com o cronograma existente.

A classe *Interact* é responsável por estender métodos e atributos, através da composição, às outras classes, permitindo a interação nas classes que a herdam. Outras classes, como *Picture-in-Picture*, *Image 360*, *Video 360*, *Sensory Effect*, *Image 2D*, *Video 2D*, *Link*, *Text Message* e *Audio 3D*, herdam métodos e atributos, pois são subclasses. A Classe *Sensory Effect* é um caso especial, pois invoca a Classe *Actuator Connector*, que é responsável pela comunicação com equipamentos físicos, os quais serão responsáveis por renderizar os efeitos sensoriais da apresentação.

A Figura 4.5 apresenta o diagrama de componentes do sistema proposto. Este diagrama é subdividido em três camadas virtuais: a camada de Execução, responsável pela execução do projeto; a camada de Edição e Criação, responsável pela edição e criação de projetos; e a camada de Conexão com Atuadores, responsável pela comunicação com os atuadores. O módulo *Player de Projeto*, contido na camada de Execução, foi desenvolvido com o objetivo de invocar o módulo *Gerenciador de Cenas*, contido na camada de Edição em Modo Visualização de projeto, informando os dados carregados na base de informações em MultiSEL [24] e gerenciando a execução de todas as cenas do projeto. No entanto, quando invocado pelo *Gerenciador de Cenas* no Modo de Edição, o módulo *Player de Projeto* se comporta em Modo de Pré-visualização, não consultando a base de informações do projeto em MultiSEL, usando somente as informações editadas em cena em tempo de execução disponibilizadas pelo *Gerenciador de Cenas*.

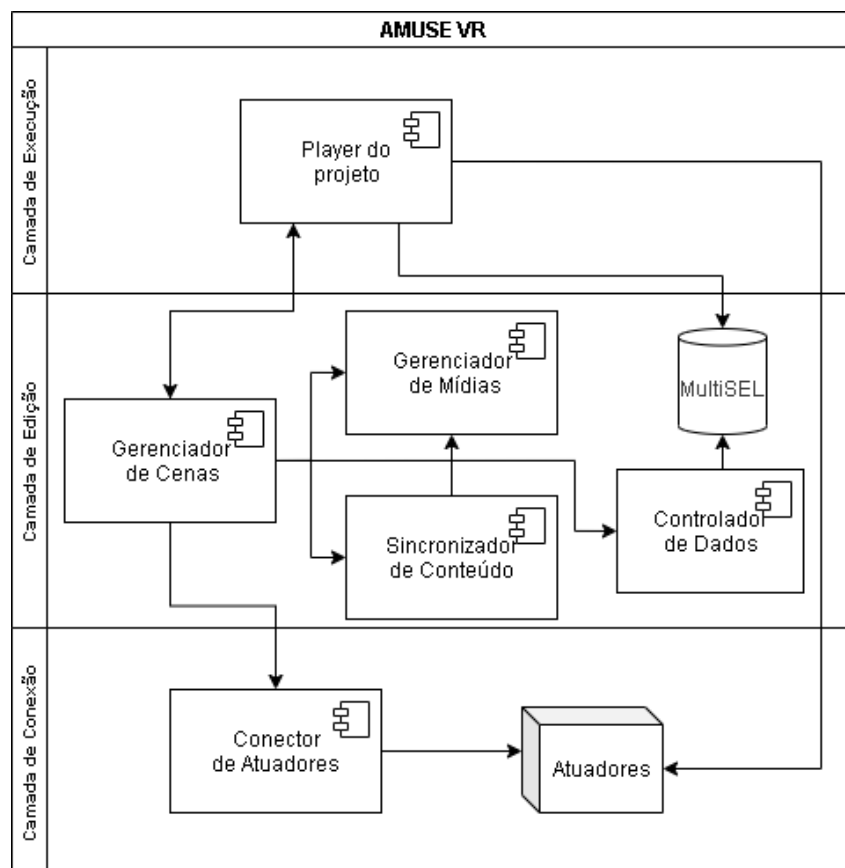


Figura 4.5: Diagrama de componentes.

O Gerenciador de Cenas presente no diagrama de componentes tem o papel principal de gerenciar todo o conteúdo autoral no Modo Edição. Ele tem acesso ao módulo Gerenciador de Mídias, que é responsável por inserir, gerenciar propriedades e conteúdo, além de permitir o posicionamento e localização de cada mídia na cena. O Gerenciador de Cenas também controla o módulo Sincronizador de Conteúdo. Este módulo também acessa o módulo Gerenciador de Mídias para obter todas as propriedades temporais de todas as mídias e, com isso, montar um plano (cronograma) de execução para entregar ao Gerenciador de Cenas. Além disso, o Sincronizador de Conteúdo considera todas as informações de relacionamentos entre as mídias e atualiza a sincronização do cronograma. Outra função do Gerenciador de Cenas é escrever e atualizar as informações no arquivo MultiSEL através do módulo Controlador de Dados.

Por fim, o Gerenciador de Cenas tem acesso ao módulo Conector de Atuadores, presente na camada de conexão. Este módulo é responsável por receber as informações das mídias sensoriais presentes na cena e executá-las, conforme o plano de execução fornecido pelo Gerenciador de Cenas, semelhante ao esquema centralizado de controle de atuadores multissensoriais de Saleme et al. [51] apresentado na Figura 2.8. O elemento Atuadores é

um módulo responsável por conectar equipamentos atuadores, e é considerado uma caixa-preta (encapsulado), sendo desenvolvido assim devido à grande variedade de componentes responsáveis pela execução. O módulo basicamente se concentra em enviar informações a cada frame da execução da aplicação para um servidor externo, que se encarrega de comunicar-se com os equipamentos, como mostrado na Figura 2.8. Todos os módulos foram desenvolvidos buscando maior independência, favorecendo possíveis substituições e extensibilidade.

4.3 Avaliação do Ambiente de Autoria

Para avaliar o ambiente de autoria, uma solução foi implementada atendendo a todos os requisitos listados na Seção 4.1, como será apresentado no Capítulo 5. Além da implementação do ambiente proposto, foram realizadas as seguintes avaliações com usuários:

- Teste de usabilidade do sistema por usuários, a fim de verificar se o ambiente de autoria atende a um nível de usabilidade aceitável;
- Teste de experiência de uso do sistema proposto por usuários, a fim de medir o grau de satisfação ao utilizar;
- Teste de facilidade de uso, a fim de verificar se os usuários leigos em produção autoral conseguem produzir conteúdo básico; e
- Teste comparativo com outra ferramenta autoral, a fim de avaliar as diferenças entre elas e as vantagens do ambiente proposto. Como não foi encontrada nenhuma outra ferramenta de autoria similar gráfica disponível, esse teste foi feito com a linguagem declarativa MultiSEL.

Com o objetivo de oferecer uma melhor qualidade avaliativa e respaldar os resultados, as avaliações foram planejadas utilizando o modelo GQM (*Goal Question Metric*) [6]. Foram utilizados modelos de questionários avaliativos bem conhecidos na literatura, como o questionário da Escala de Usabilidade do Sistema (*System Usability Scale* - SUS) [10] e o questionário de experiência do usuário (*User Experience Questionnaire* - UEQ) [40, 55] e o Framework de Dimensões Cognitivas de Notações (*Cognitive Dimensions of Notations* - CDN) [9]. As avaliações serão detalhadas no Capítulo 6.

Capítulo 5

AMUSEVR

Este capítulo apresenta o ambiente AMUSEVR (registrado como: *AMUSEVR - Authoring 360° Multimedia and Sensory Effects in VR* no INPI - Instituto Nacional da Propriedade Industrial), que implementa a abordagem de autoria imersiva proposta nesta tese, com o objetivo de atender aos objetivos da pesquisa e oferecer uma solução que avance o estado-da-arte em autoria multimídia.

AMUSEVR [27, 28] é uma ferramenta de autoria imersiva em realidade virtual, com enfoque em mídias em 360 graus e suporte a apresentações multissensoriais interativas, nas quais as mídias podem ser dispostas espacialmente em três dimensões no cenário.

A ferramenta foi implementada em Unity com uma primeira versão para o HTV Vive e uma segunda versão para o Oculus Quest 2, como será detalhado nas seções a seguir.

5.1 Desenvolvimento com Unity

A implementação do ambiente AMUSE VR foi desenvolvida utilizando o motor Unity na versão 2019, com compilação voltada para plataforma Windows. Durante o desenvolvimento, foi utilizado o pacote complementar de interface com usuário ViveMenuToolkit¹, responsável por servir como base para as janelas, menus e interações básicas. Além disso, foi utilizado o VRKeys², que fornece uma interface de teclado QWERT com foco em ambientes virtuais. Todos os pacotes adicionais são de código aberto. Para a execução dos efeitos sensoriais fora do mundo virtual, foi utilizado, nos experimentos, um Arduino Uno com Arduino IDE³ para compilação dos códigos específicos. Para as imagens demonstra-

¹<https://github.com/davrj012/ViveMenuToolkit>

²<https://github.com/campfireunion/VRKeys>

³<https://www.arduino.cc/en/software>

tivas e experimentos, foi utilizado conteúdo de mídia disponível em sites gratuitos⁴⁵.

Durante a idealização, foi cogitada a utilização do motor gráfico Unreal Engine, até por possuir algumas funções de edição em modo imersivo, mas devido a melhor expertise do autor em desenvolver com a Unity e devido a boa quantidade de material de apoio foi optado por esse motor gráfico. Além disso, dentro da grande variedade de equipamentos de realidade virtual para os quais a Unity oferece suporte, é necessário implementar especificações para algumas famílias de equipamentos de alguns fabricantes. Portanto, na primeira versão do AMUSEVR (AMUSEVR Vive), decidiu-se adotar e desenvolver para a família de equipamentos HTC Vive⁶. Isso se deve, em parte, ao fato de, na época, cerca de 2019, a plataforma oferecer mais poder computacional do que seus concorrentes, favorecendo projetos que apresentam muitas mídias simultaneamente, como foi o caso da primeira versão do AMUSEVR (AMUSEVR Vive), utilizada nos primeiros experimentos publicados em [27, 28].

A primeira versão do ambiente AMUSE VR foi desenvolvida para o kit de realidade virtual HTC Vive, que inclui o capacete de realidade virtual, um par de fones de ouvido com suporte a áudio multidirecional (áudio 3D) e um par de controles gêmeos, ambidestros e com rastreamento posicional, giroscópios e acelerômetro inclusos. O equipamento requeria um ambiente de execução preparado previamente, pois os sensores eram fixados em paredes, e o capacete era ligado ao computador por cabos, diminuindo a mobilidade. Os controles exigiam programação específica, pois cada marca e modelo possuía um layout específico de botões, gatilhos e comandos. O computador utilizado para o desenvolvimento possuía um processador I7 de nona geração, uma placa gráfica NVIDIA GTX 1080 TI, 16 GB DDR4 de memória e SSDs NVMe M. 2. No entanto, foi observado que essa configuração era mais que suficiente para a execução da ferramenta.

Nesta primeira implementação, foi planejado que todos os comandos, tanto na fase de autoria quanto na fase de visualização, fossem realizados somente com um botão de gatilho. Esses comandos incluem um clique simples com o gatilho e um conjunto sequencial de pressionar, mover e soltar, com o objetivo de proporcionar familiaridade e semelhanças com os comandos de um apontador de mouse em plataformas PC.

Além disso, no modo de autoria da ferramenta, um dos controles é responsável pela função de apontamento de objetos dentro do ambiente virtual. Essa função é ativada quando um objeto está dentro do raio em formato de laser de apontamento do controle, o

⁴<https://www.pexels.com/>

⁵<https://www.mettle.com/360vr-master-series-free-360-downloads-page/>

⁶<https://www.vive.com/us/>

que ativa uma animação de apontamento laser e destaque luminoso de seleção no objeto apontado. Quando o controle é desviado do objeto, a animação do apontador laser e o destaque de seleção são desativados. O apontamento laser também é ativado quando o controle é apontado para menus e interfaces do usuário, como mostrado na Figura 5.1.



Figura 5.1: Menu e controles.

A decisão de projeto da interface foi tomada com o objetivo de informar ao usuário quando ele tem foco em um menu ou objeto editável. Quando não há foco, a informação de apontamento laser do controle é retirada, diminuindo a poluição visual e melhorando a informação fornecida ao usuário, conforme mostrado na Figura 5.1. Além disso, o controle com a função de apontador exibe uma mensagem textual em verde, informando o modo de autoria ativo no qual o autor está trabalhando no momento. As mensagens podem variar entre "modo autor", "modo inclusão", "modo edição" e "modo exclusão", exemplos disso podem ser vistos nas Figuras 5.1 e 5.16.

Como se trata de uma ferramenta de autoria em ambiente virtual, foi necessário considerar uma abordagem diferente da tradicionalmente utilizada em ferramentas bidimensionais. Diante disso, decidimos exibir o menu de forma resumida e discreta, a fim

de evitar poluição visual. Além disso, para garantir compatibilidade com um ambiente tridimensional, adotamos a estratégia de fixar o menu a um dos controles. O controle responsável por exibir o menu principal foi posicionado sempre acima do usuário, com a face voltada para ele, e poderia ser escondido ou exibido ao posicionar ou retirar o controle do campo de visão do capacete virtual.

No entanto, através do botão "Switch Menu", mostrado na Figura 5.1, é possível vincular o posicionamento do menu principal ao centro da visão do capacete do usuário. Essa opção foi incluída para permitir que o autor possa esconder o menu/controla de sua visão, concentrando-se apenas no conteúdo trabalhado, e, quando necessário, rapidamente elevar o controle à frente de sua visão para visualizar novamente o menu principal e editar o projeto.

Como mencionado anteriormente, devido ao fato de o HTC Vive possuir um conjunto de controles gêmeos, nesta versão do AMUSEVR Vive, é possível utilizar e alternar a posição do controle de menu principal e de apontamento livremente, sem necessidade de configurações prévias, atendendo tanto pessoas destros quanto canhotos.

O AMUSEVR foi desenvolvido utilizando o motor gráfico Unity, cuja IDE de desenvolvimento suporta principalmente a linguagem de programação C#. Grande parte do desenvolvimento é realizado através de comandos de alto nível fornecidos pelas ferramentas da Unity, mas para o tratamento e relacionamento dos "Objetos de Jogo" (*Game Objects*) é possível complementar a lógica e funcionamento desses objetos com scripts em C#.

A estrutura básica do desenvolvimento no Unity consiste em projetos, que contêm uma ou mais cenas de jogo, cada uma contendo zero ou vários objetos de jogo, que podem assumir formas 1D, 2D e 3D de acordo com a necessidade do desenvolvedor. No desenvolvimento do AMUSEVR, foi seguida essa hierarquia padrão do Motor Gráfico, aplicando a capacidade morfológica dos objetos de jogo às mídias e, assim, concentrando-se nas especializações dos "Objetos de Mídia", como ilustrado na Figura 5.2.

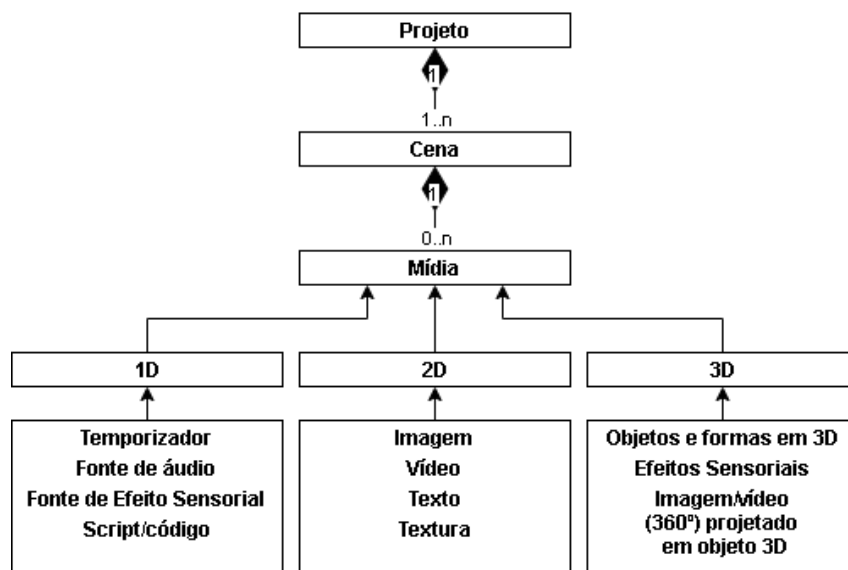


Figura 5.2: Hierarquia no projeto.

A abordagem apresentada na Figura 5.2 possui a desvantagem de dificultar a documentação do projeto. Isso ocorre devido à falta de ferramentas de documentação automatizada fornecidas pelo Unity, além da mistura de scripts criados pelo desenvolvedor com outros scripts da ferramenta e complementos de terceiros no projeto.

O AMUSEVR implementa uma estrutura de classes, conforme ilustrado na Figura 4.4. De acordo com este diagrama, as classes *Mídia*, *Mídia 360*, *Objeto de mídia* e *Interação* são classes abstratas, não podendo ser instanciadas diretamente, mas somente acessíveis por suas classes herdeiras, favorecendo a reutilização de código.

A Figura 5.3 apresenta um fluxograma que descreve a lógica sequencial das principais ações executadas pelo autor no modo de autoria para criar ou editar uma cena.

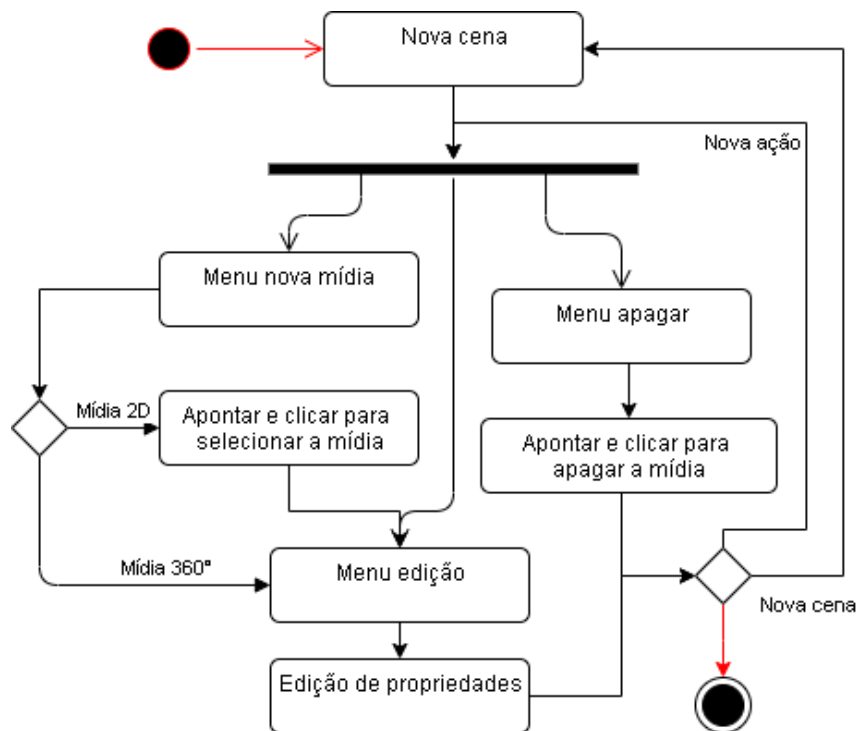


Figura 5.3: Fluxograma para edição de cenas.

5.2 Características da ferramenta

O AMUSEVR foi projetado para usar tecnologias de equipamentos vestíveis de realidade virtual e, com elas, visa suportar, inicialmente, mídias 2D, 360 graus e efeitos sensoriais em seu ambiente imersivo. Para isso, foi projetado que todos os procedimentos autorais seriam executados dentro do ambiente virtual. Destacando-se opcionalmente ambientes com fundo animado ou estático de mídias em 360 graus, com suporte a múltiplas mídias executadas em paralelo, conforme a especificação feita pelo autor. A seguir, serão discutidas diversas de suas características.

5.2.1 Modos de uso da ferramenta

Conforme os casos de uso já apresentados anteriormente na Figura 4.2, a ferramenta AMUSEVR foi implementada com dois modos principais: o modo de carregamento de projetos, conhecido como modo Visualização, e o modo de criação de projetos, conhecido como modo Autoria, visando seu uso pelos usuários finais e autores respectivamente. O modo Autoria é responsável por permitir a criação de projetos imersivos, incluindo a visualização e a edição dos projetos em tempo de execução. No modo autoria, ainda é possível acessar via menu principal a opção de pré-visualização (*preview*), essa opção

simula todo o cronograma de execução e interação da cena atualmente editada, sem a necessidade de salvar o projeto.

Dentro do modo Autoria no menu principal, é disponibilizada a criação de cenários, também conhecidos como cenas (*Scenes*), além dos comandos Criar, Editar e Apagar, utilizáveis em todos os objetos de mídia. Todos os comandos seguem a sequência ilustrada na Figura 5.4. Além disso, é possível executar um modo de Pré-Visualização da cena, que simula a execução da cena ativa no projeto, com a execução de todos os objetos de mídia segundo o cronograma de apresentação estabelecido pelo autor. Os objetos de mídia são iniciados quando são exibidos e interrompidos após serem retirados da cena, obedecendo os relacionamentos criados pelo autor no projeto. O modo Pré-Visualização da Cena apresenta um cronômetro referencial para apresentação, que serve de referência para as mídias quando não possuem um relacionamento prévio especificado pelo autor.

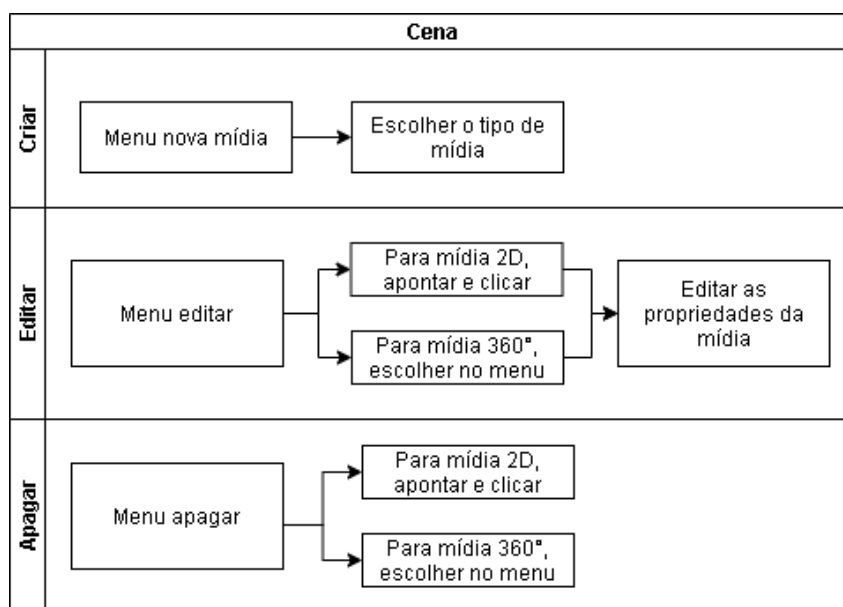


Figura 5.4: Fluxograma de criação, edição ou remoção de mídias.

5.2.2 Objetos de Mídias

Como mencionado anteriormente, a ferramenta de autoria faz uso de objetos de mídia que dão suporte a diversos tipos de mídia, como ilustrado na Figura 5.2. Todas as mídias são do tipo objeto de mídia, mas o que diferencia um objeto de outro é o tipo de especialização do objeto de mídia, que é melhor descrito na Tabela 5.1.

Tabela 5.1: Características da mídia segundo seu tipo.

Tipo de mídia	Características
Áudio 3D	Repetição, volume e mudo
Imagem 2D	Escala
Imagem 360°	Repetição
Objetos interativos	Alvo ou relação eventual de início/fim com objeto/cena
<i>Picture in picture</i>	Repetição, volume e mudo
Efeito sensorial	Repetição, duração e intensidade
Mensagem de texto	Repetição e conteúdo textual
Vídeo 2D	Repetição, volume, mudo e escala
Vídeo 360°	Repetição, volume e mudo

Durante o processo de autoria, o usuário tem a opção de editar as características dos objetos de mídia. Para isso, é exibida uma janela flutuante de edição de propriedades, que é individual para cada objeto de mídia criado em cena. A Figura 5.5 ilustra um exemplo dessa tela de edição para vídeo em 2D, com suas configurações editáveis. Em relação ao reprodutor de vídeo, a documentação associada completa com todo o suporte aos formatos pode ser encontrada nas páginas de desenvolvimento da Unity⁷⁸.

⁷<https://docs.unity3d.com/Manual/AudioFiles.html>

⁸<https://docs.unity3d.com/Manual/VideoSources-FileCompatibility.html>

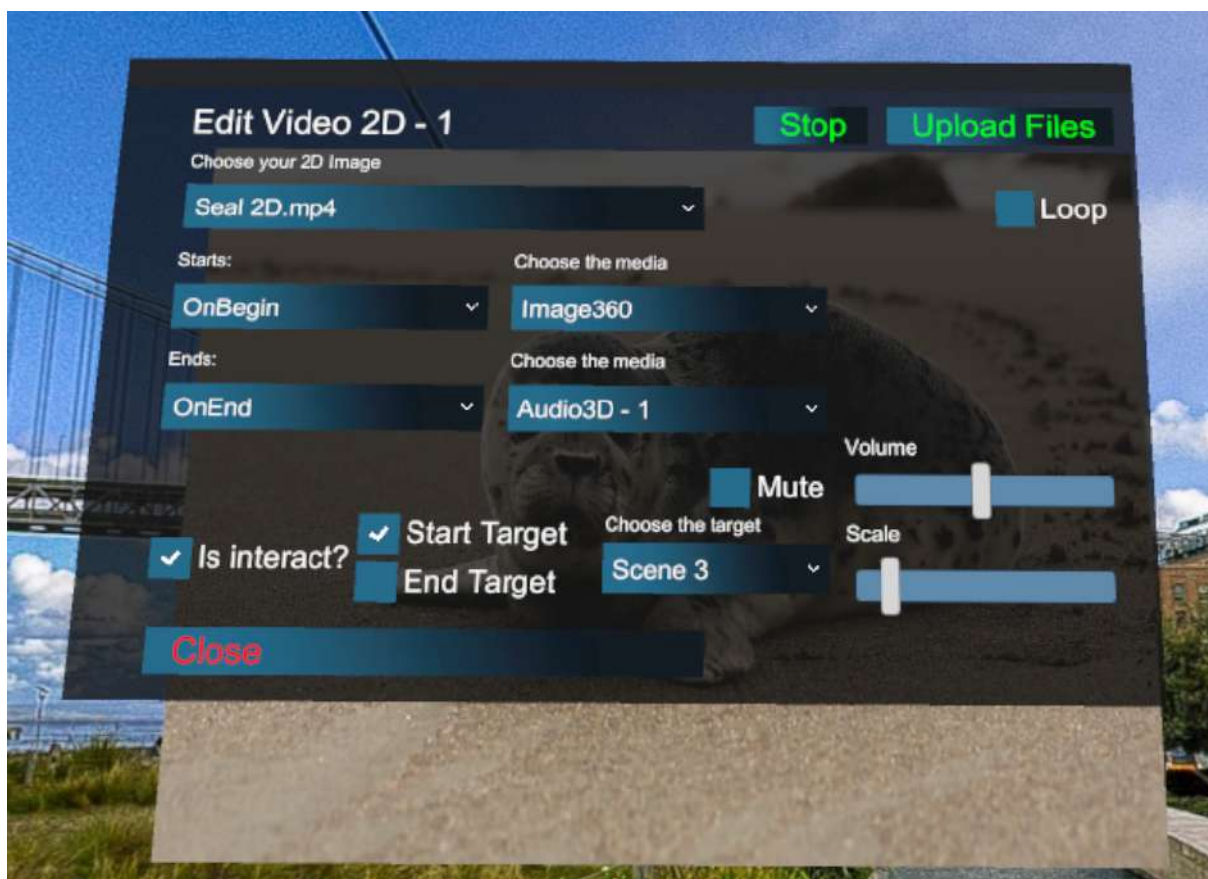


Figura 5.5: Tela de edição.

Para deixar mais claro como é composto um menu de um objeto de mídia, na Figura 5.6, é possível ver uma janela de edição de vídeo 2D de amostra. Suas propriedades são as seguintes:

1. Exibe o nome de identificação (ID) do objeto de mídia;
2. Uma lista suspensa (*dropdown*) de arquivos com conteúdo compatível com o objeto de mídia, presentes na pasta de mídias da instalação.
3. Relações baseadas em eventos especificando como este o objeto de mídia começa ou termina na apresentação. Dentro da caixa de seleção "*Starts*", podemos encontrar as opções "*OnBegin This Scene*", "*OnBegin*", "*OnEnd*" e "*Not defined*", cada uma pode estar relacionada a outro objeto de mídia dentro da mesma cena. As opções da caixa de seleção "*Ends*" são "*End of This Media Time*", "*OnBegin*" e "*OnEnd*", que também podem ser relacionadas a outro objeto de mídia dentro da cena;
4. Opções de interação do usuário. Ao habilitar o objeto de mídia para ser interativo, ele pode iniciar ou terminar outra cena ou um objeto de mídia quando o usuário

- final interagir com ele durante a apresentação;
5. Barra deslizante de escala para mudar o tamanho do objeto de vídeo;
 6. Grupo de opções de áudio relacionadas ao objeto;
 7. Define um atraso para o tempo de início do objeto;
 8. Controla a duração dos itens de mídia estáticos e/ou efeitos sensoriais. Controla o looping de vídeo/áudio;
 9. Importa para o projeto novos arquivos de mídia na pasta de instalação da ferramenta para preencher o menu de lista suspensa (*dropdown*) do item 2;
 10. Inicia ou para a mostra de execução da mídia durante o modo de autoria;
 11. A opção "*Look at*" mantém a mídia com a face de exibição orientada para o usuário durante a execução; e
 12. Uma faixa que permite ao usuário arrastar, soltar, distanciar ou aproximar o objeto para outra posição dos eixos x , y e z do projeto.



Figura 5.6: Janela de edição de um objeto de vídeo em detalhes.

A seguir, é descrito de forma detalhada o comportamento das mídias suportadas pela ferramenta, acompanhado de imagens ilustrativas:

Áudio 3D - Com suporte aos principais tipos de áudios populares, como MP3 e WAV, possui suporte ao posicionamento espacial, possibilitando a fixação de emissores de som independentemente da posição do usuário, o que resulta em uma imersão aumentada. Devido ao uso de recursos contidos na Unity, é possível entregar um áudio em 3D. Além disso, oferece a possibilidade de pré-execução do áudio, controle de volume e habilidade de mutar o som (Figura 5.7).

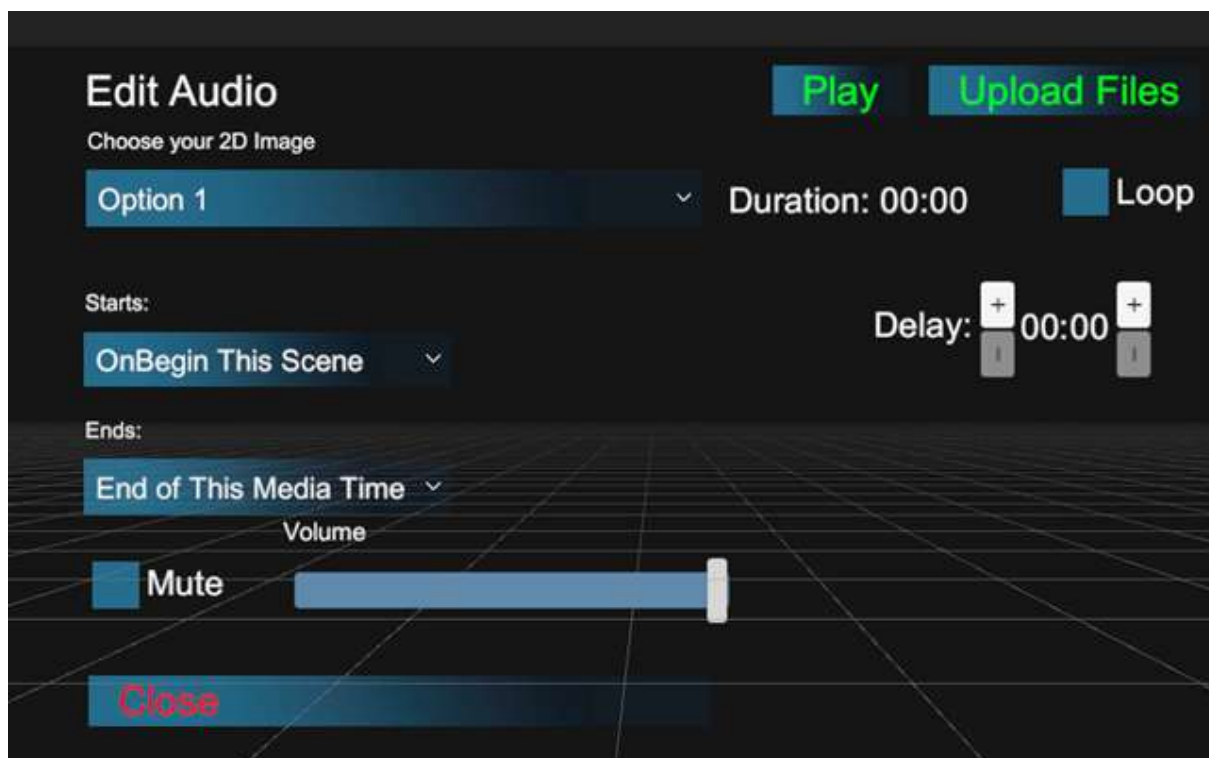


Figura 5.7: Edição de Áudio 3D.

Imagem estática em 2D - Com suporte aos principais tipos de imagens populares, como PNG e JPG, além de suporte ao posicionamento espacial, com função de ajuste e reposicionamento automático da face da imagem, de forma que ela sempre esteja voltada para o autor. Essa funcionalidade permite a alteração do tamanho escalar da imagem no cenário (Figura 5.8).

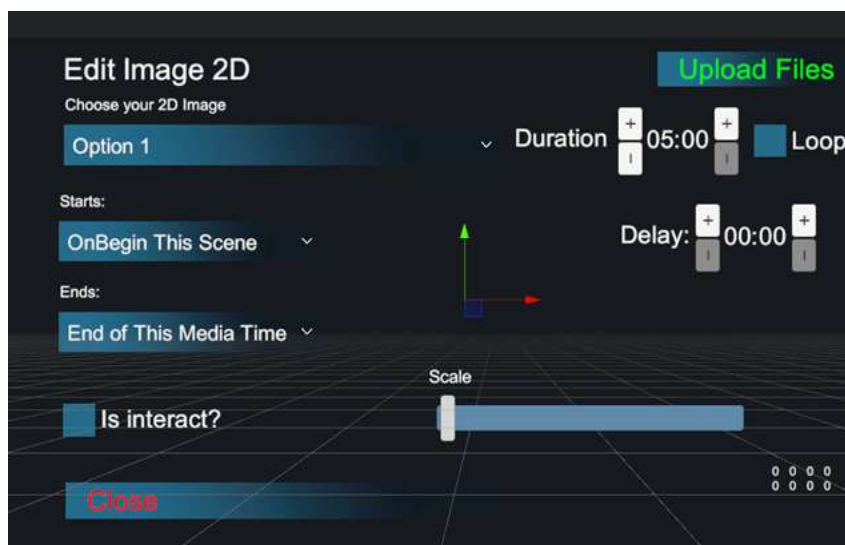


Figura 5.8: Edição de Imagem em 2D.

Imagem em 360 graus - Com suporte aos principais tipos de imagens populares, como PNG e JPG, recomenda-se o uso de imagens no formato monoscópico, panorâmico e equirrectangular em 360 graus, para melhor aproveitamento do campo de visão (Figura 5.9).

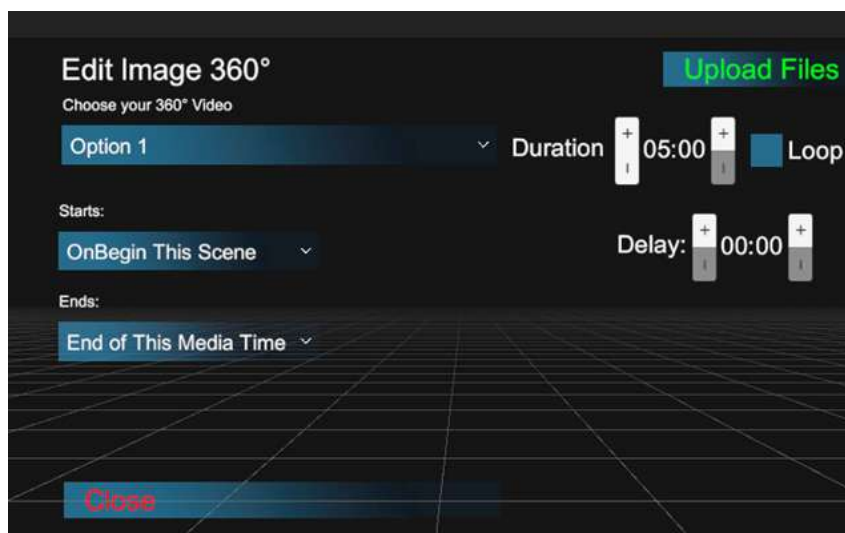


Figura 5.9: Edição de Imagem em 360 graus.

Objetos interativos - São objetos criados durante o projeto para funcionar como origem de um elo (*link*), um portal para ir para outras cenas ou como botões de interação com outros objetos de mídia relacionados. Visualmente, são apresentados como um ícones de porta (Figura 5.10). Outros objetos em 2D como imagens, vídeos e textos podem assumir essa função do objeto interativo se assim forem configurados pelo autor.

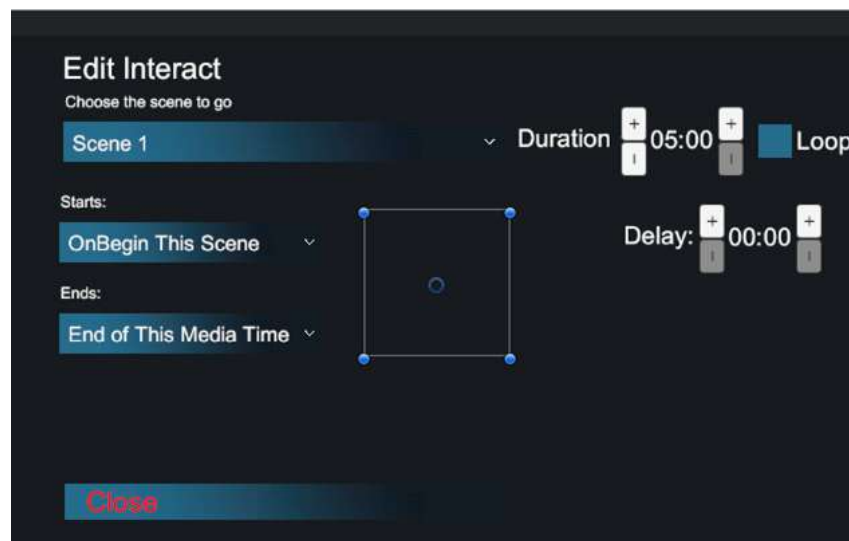
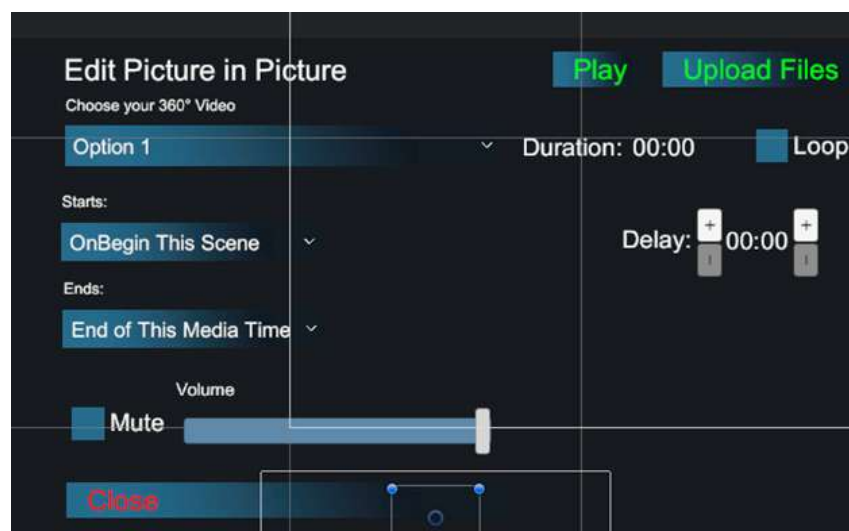


Figura 5.10: Edição de Objetos interativos.

Picture-in-Picture (PIP) - Com suporte aos principais tipos de vídeos populares, como MP4 e AVI, possui a característica de ficar sempre fixo na região inferior direita, à frente do campo de visão do usuário, isto é, independentemente dos movimentos realizados pelo usuário no cenário ou com a cabeça. Tem como objetivo apresentar informações secundárias e opcionais à apresentação, como uma legendas, vídeos em libras ou replays de uma cena esportiva. Disponibiliza a pré-visualização do vídeo, alteração de volume e controle de mutagem binária (Figura 5.11).

Figura 5.11: Edição de *Piture-in-Picture* - PIP.

Efeito Sensorial (ES) - São objetos criados durante o projeto para simular visualmente efeitos e acionar equipamentos, com três tipos de efeitos sensoriais pré-definidos disponíveis na implementação atual (calor, vento e vibração). A simulação visual, por

exemplo, apresenta partículas de vento, somadas ao acionamento externo de equipamentos sincronizados com o objeto na apresentação. Além disso, é possível utilizar, assim como o áudio 3D, um ponto de emissão visual correspondente ao efeito sensorial escolhido. Como propriedade, é possível alterar a intensidade de emissão do efeito (Figura 5.12).

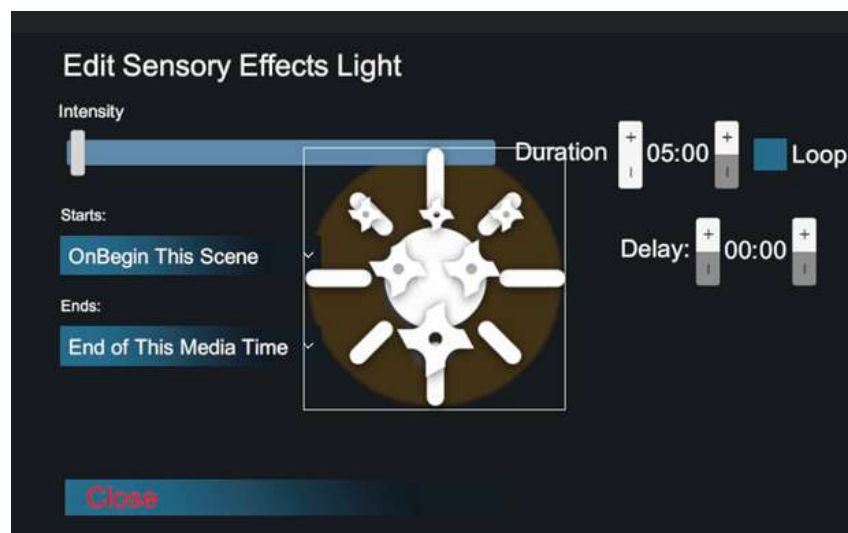


Figura 5.12: Edição de Efeito Sensorial - ES.

Mensagens de Texto - São objetos criados durante o projeto para exibir mensagens textuais. As mensagens são elaboradas no modo de autoria, utilizando um teclado virtual oculto, invocado através das propriedades do objeto da ferramenta, conforme ilustrado na Figura 5.13. Além disso, é possível ajustar o tamanho escalar do texto no cenário.

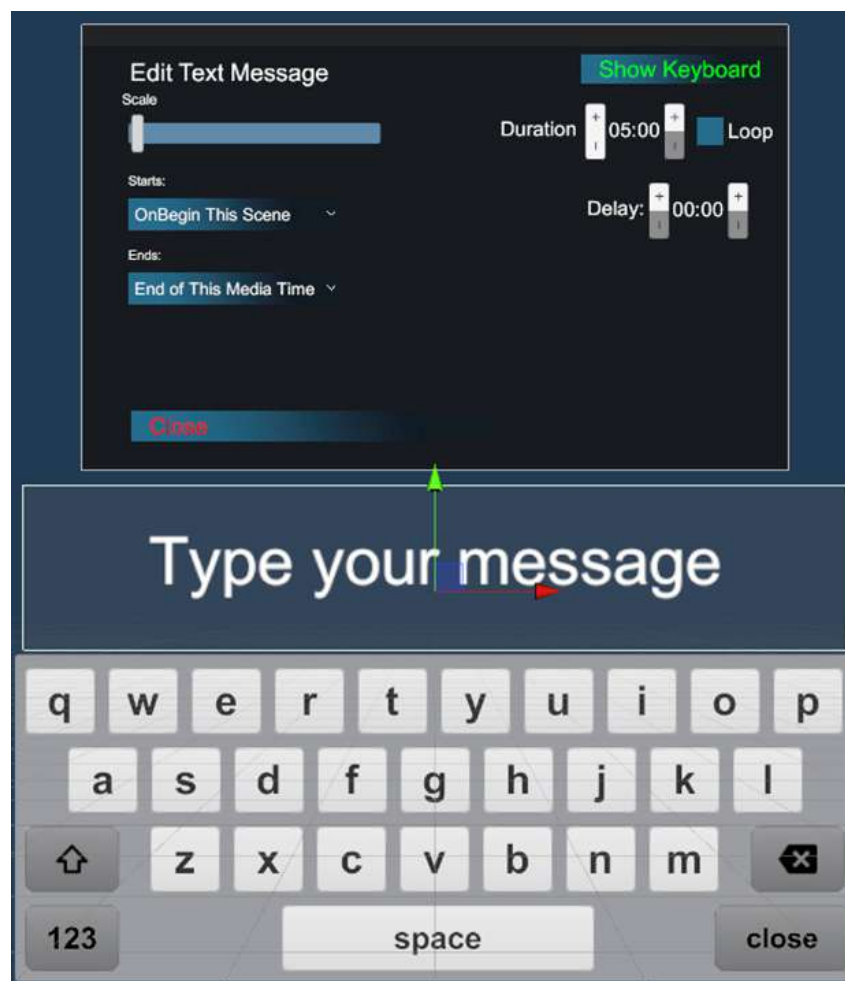


Figura 5.13: Edição de Mensagem Textual.

Vídeo em 2D - Com suporte aos principais tipos de vídeos populares, como MP4 e AVI, disponibiliza a pré-visualização do vídeo durante a edição, alteração do volume, controle de mutagem binária, interatividade e possibilidade de ajustar o tamanho escalar do vídeo no cenário (Figura 5.14).

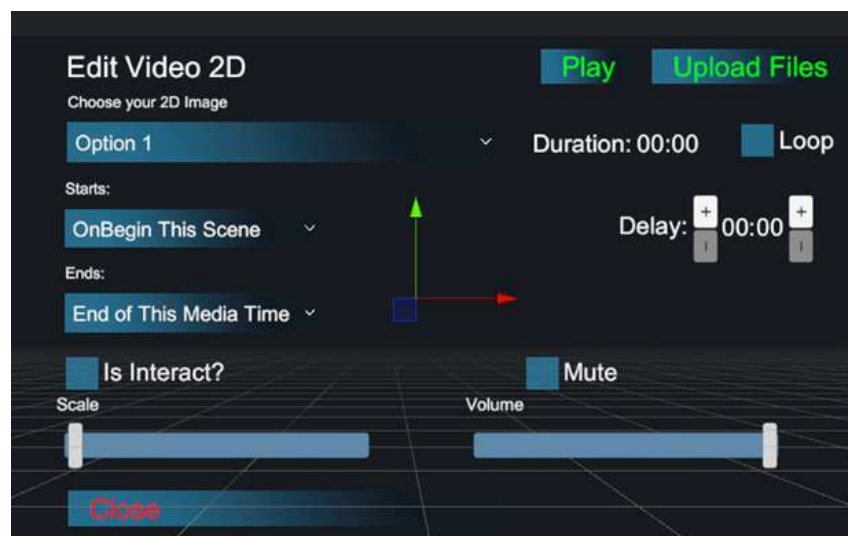


Figura 5.14: Edição de Vídeo em 2D.

Vídeo em 360 graus - Com suporte aos principais tipos de vídeos populares, como MP4 e AVI, recomenda-se o uso de vídeos no formato monoscópio, panorâmico e equirrectangular em 360 graus para obter uma melhor imersão e aproveitamento do campo de visão. Estes vídeos podem ser pré-visualizados, e possuem opções de controle de volume e mutagem binária (Figura 5.15).

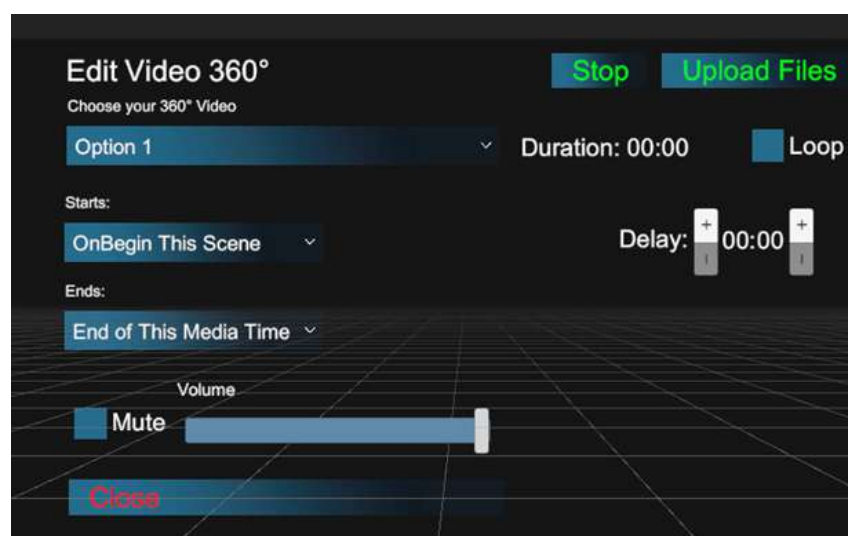


Figura 5.15: Edição de Vídeo em 360 graus.

Outro aspecto importante a ser mencionado é que todos os objetos compartilham características comuns em seus menus de edição:

- Alteração da duração de exibição da mídia na cena;

- Tempo de atraso de início da exibição;
- Opção de controle binário de repetição da duração do tempo; e
- Vinculação do relacionamento temporal inicial e final com outros objetos ou com a cena em si, através de *links*.

Por fim, alguns tipos de objeto de mídia possuem restrições para criação (instanciação), como as mídias em 360 graus, que podem ser criadas apenas uma por cena, devido à ocupação da mesma textura de projeção no pano de fundo. Como decisão de projeto, optou-se por não disponibilizar a alteração de mídia em 360 graus na mesma cena, para incentivar a utilização de várias cenas e evitar a superconcentração de mídias em uma única cena, favorecendo o melhor uso de memória RAM do equipamento. Além disso, os vídeos em *Picture-in-Picture* também são restritos a um por cena, pois são exibidos sempre à frente do campo de visão do usuário. Como estratégia de usabilidade, acredita-se que inúmeros vídeos posicionados fixamente à frente do usuário poderiam prejudicar a experiência do usuário, limitando o campo de visão. Para edição destas mídias específicas, deve-se acessar diretamente o menu de edição, visível e acessível por padrão no menu principal, acima do controle, como visto na Figura 5.16



Figura 5.16: Modo de edição.

5.2.3 Posicionamento Espacial dos Objetos

A Figura 5.17 ilustra um aspecto importante do conceito espacial 3D proposto, que é a capacidade de posicionar objetos de mídia 2D em qualquer lugar nos eixos x , y e z dentro

do ambiente em 3D, simplesmente utilizando a ação de arrastar e soltar. Adicionalmente, na Figura 5.17, é mostrada à esquerda a imagem de um ponto turístico (Pão de Açúcar) na cidade do Rio de Janeiro, que é um objeto de mídia do tipo imagem 2D. No lado direito da mesma figura, há um outro objeto de mídia do tipo vídeo 2D, que retrata uma pequena foca em uma praia. Na mesma figura, abaixo da imagem do Pão de Açúcar, o cenário ainda contém um terceiro objeto de mídia do tipo mensagem de texto, contendo a frase "*YOUR MESSAGE HERE*" em formato textual na cor branca. Por fim, um quarto objeto de mídia do tipo vídeo em 360 graus, que retrata uma cena de praia ao fundo do cenário. Todos os objetos de mídia têm posições e atributos editáveis dentro da cena, permitindo que os autores ajustem suas características e posições.



Figura 5.17: Objetos de mídia 2D em uma cena de 360 graus.

Dentro do ambiente virtual, todos os elementos na cena possuem um posicionamento espacial tridimensional, dado pela tupla de coordenadas de valores reais x , y e z , para objetos bidimensionais. Essas coordenadas se referem à ancoragem dos pontos que estão no centro das dimensões x e y de cada objeto. Cada objeto e menu ancorado em uma apresentação possui uma propriedade que mantém a face de exibição do objeto de mídia sempre voltada para a visão do usuário. Isso foi idealizado devido ao projeto prevê 6 graus de liberdade de movimentação do usuário, e, dessa forma, quando o usuário contorna um objeto, andando em volta dele ou até quando se abaixar, sempre terá todos os objetos de duas dimensões voltados para a sua visão. Outra característica a ser destacada é a preferência de exibição de menus dentro do campo de interesse do usuário. Por padrão, conforme descrito em [64], é considerado campo preferencial de interesse quando a câmera (visão do usuário) está na posição x igual a 0, z igual a 0 e voltada na direção do eixo y .

positivo.

Com exceção da cena, que serve como referência para todos os objetos contidos nela, todos os demais objetos têm a posição de marco inicial $(x, y, z) = (0, 0, 0)$ da cena como referência. A partir desse marco inicial, cada objeto é posicionado e tem sua posição individual no cenário. Cada objeto na cena pode servir como referência e representar um novo marco inicial para outros objetos, o que é considerado uma nova referência posicional.

Como exemplo, suponha a ilustração apresentada na Figura 4.1, que representa um ambiente de autoria imersiva. Essa figura descreve um ambiente com uma imagem de-sértica em 360 graus como pano de fundo, um usuário autor usando um capacete de realidade virtual que representa o ponto de vista do autor, um par de controles que são visíveis dentro do ambiente virtual, dentro do campo de interesse preferencial do usuário, uma tela de *Picture in Picture* vinculada à visão do capacete, um áudio à direita do autor, um efeito sensorial mais à frente do autor e, por fim, um vídeo posicionado à esquerda do campo de interesse preferencial da visão do autor.

Descrito o exemplo e usando a Figura 5.18, temos um quadro com a descrição do referencial de cada objeto em cena. Neste quadro, representando uma árvore, de baixo para cima, todas as folhas são dependentes e utilizam como referencial o posicionamento do galho acima. Como destaque, o nó do ponto de visão do usuário utiliza o cenário como referência.

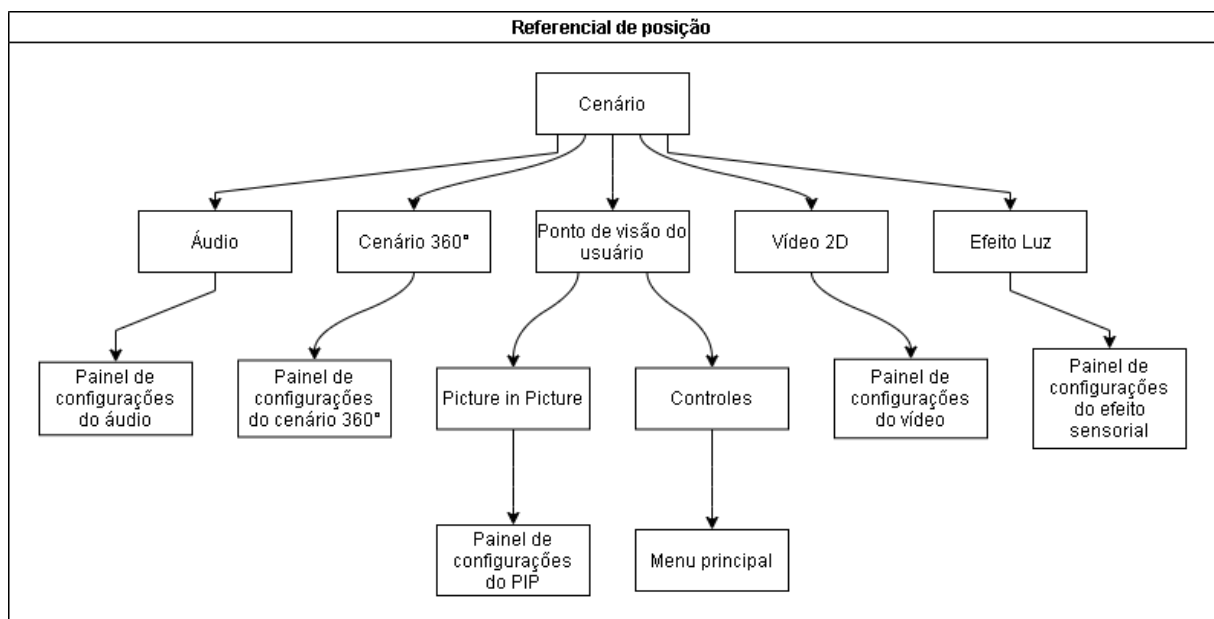


Figura 5.18: Dependência posicional dos objetos em cena.

Durante a execução da aplicação do projeto, a posição (x, y, z) do usuário é atualizada

constantemente conforme ele se move pelo cenário. Como resultado, todos os objetos vinculados também são automaticamente atualizados, incluindo o *Picture in Picture* e os controles com seus respectivos menus e informações. Isso significa que, por exemplo, se o autor se move 5 pontos no eixo y , todos os objetos dependentes de sua referência se moverão também 5 pontos no eixo y . Na prática, isso significa que o *Picture in Picture* na ilustração 4.1 permanecerá sempre fixo à frente da visão do autor, independentemente de ele olhar para cima, para baixo ou em torno do próprio eixo.

5.2.4 Seis graus de liberdade

A ferramenta AMUSEVR possui modos de autoria e visualização que suportam movimento em seis graus de liberdade (*six degrees of freedom* - 6DoF), visando favorecer a liberdade criativa e flexibilidade para os autores e maior imersão para os participantes em uma apresentação. Isso é possível devido à capacidade do usuário de andar livremente pelo cenário virtual, limitado apenas pelo ambiente físico no mundo real.

No modo de autoria da AMUSEVR, os efeitos de mídia e efeitos sensoriais quando em modo de edição, oferecem uma opção chamada "*Look at*", mostrada como item 11 na Figura 5.6. Sempre que ativada, ela faz com que a mídia 2D, áudio ou até mesmo um efeito sensorial gire em no próprio eixo central e focalize sua face frontal ou a direção de emissão para o centro do HMD de realidade virtual do autor (ou usuário final), ainda mantendo seu ancoramento nas posições x , y , z definidas pelo autor. Quando a opção é desativada, a mídia 2D, áudio ou até mesmo efeito sensorial, para de atualizar o foco no HMD de realidade virtual e mantém sua direção de frente mais recentemente atualizada.

Por meio da propriedade de mídia "*Look at*", juntamente com a permissão para os participantes terem completa liberdade de movimento dentro da apresentação, o autor pode projetar cenários imersivos onde a posicionamento de mídia é coordenado com a participação dos usuários.

Considerando que os objetos de áudio na AMUSEVR suportam áudio 3D, o som é espalhado espacialmente a partir de aproximadamente 1 metro do ponto de emissão (distância aproximada dentro do mundo virtual), com perda de volume linear até alcançar aproximadamente 10 metros de raio a partir do ponto de emissão. Quando os objetos de áudio 3D produzem sons gravados em formato mono, eles distribuem o som igualmente em todas as direções; enquanto que, quando o conteúdo é gravado no modo multicanal, ele usa o ícone visual como referência central para a emissão de som.

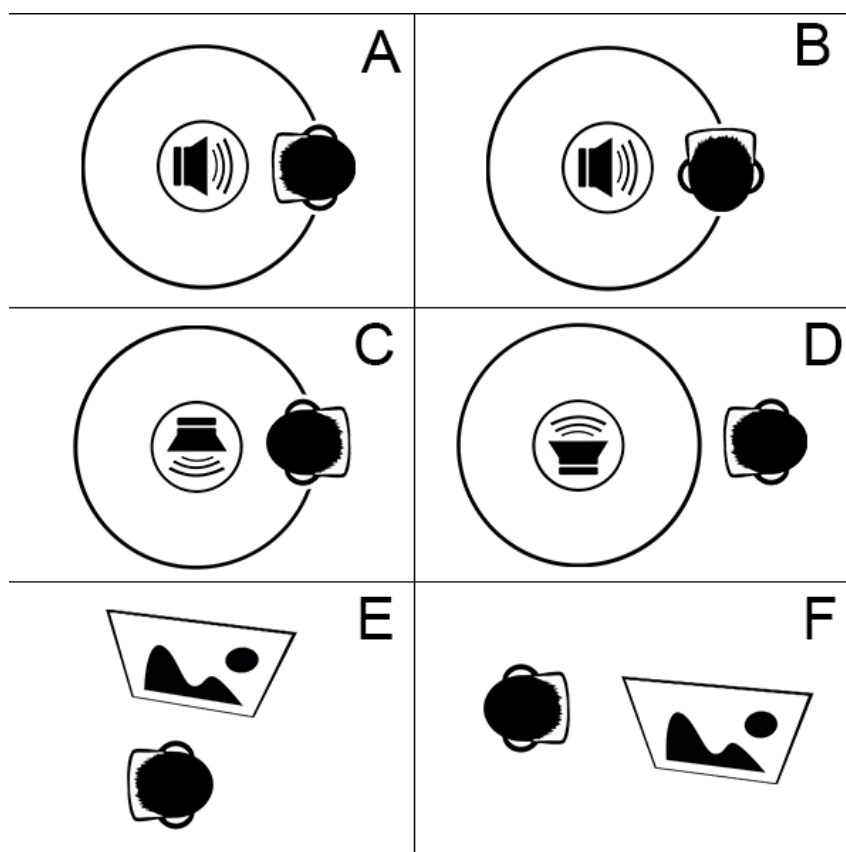


Figura 5.19: Exemplo de 6DoF.

Considerando também imagens e vídeos 2D, o conteúdo será apresentado através da frente do quadro (*layer*). Isso significa que um participante observando lateralmente a 90 graus da frente ou diretamente atrás do objeto de mídia de imagem ou vídeo 2D não será capaz de ver o conteúdo sendo exibido. Seguindo as considerações acima, a Figura 5.19 descreve cenários personalizados que exemplificam o uso de mídias, como áudio 3D e imagens 2D, levando em conta a posição do participante:

- No quadro A - temos áudio que possivelmente tem a função "*Look at*" habilitada e o participante está no alcance de som do áudio, ele perceberá o áudio como sendo transmitido à sua frente.
- No quadro B - temos áudio que possivelmente tem a função "*Look at*" habilitada e apenas o lado esquerdo do participante está no raio de som do áudio. Ele perceberá o áudio como sendo transmitido à sua esquerda.
- No quadro C - temos áudio que tem a função "*Look at*" desabilitada e o participante está no raio de som do áudio, ele perceberá o áudio como sendo transmitido atrás dele.

- No quadro D - temos áudio que tem a função "*Look at*" desabilitada e o participante está fora do alcance de som do áudio, ele não ouvirá esse objeto de áudio específico.
- No quadro E - temos uma imagem que possivelmente tem a função "*Look at*" habilitada e o participante será capaz de visualizar a imagem assim que o campo de visão do participante for girado para o lado direito.
- No quadro F - temos uma imagem que tem a função "*Look at*" desabilitada e o participante poderá ter dificuldade em visualizar o objeto de imagem 2D, a menos que ele/ela se posicione mais à frente da imagem e olhe para ela, ele não conseguirá visualizar a imagem.

Considerando a coleção de cenários na Figura 5.19, podemos ver como o autor, tendo a liberdade de colocar diferentes conteúdos para serem exibidos em diferentes momentos e diferentes locais, aproveitaria a diversidade de opções durante o processo criativo de seus projetos imersivos.

5.2.5 Efeitos sensoriais

AMUSEVR propõe-se a suportar apresentações multimídia com a integração de efeitos sensoriais, como os mencionados em [51]. De acordo com autores como [47, 52, 11], é possível aumentar a imersão, expandindo as sensações obtidas com a tecnologia de realidade virtual, através da adição de efeitos sensoriais sincronizados.

Inicialmente, baseado em [20], há duas formas de integrar efeitos sensoriais em uma apresentação multimídia: capturando e armazenando dados dos efeitos sensoriais previamente e reproduzindo-os posteriormente com equipamento adequado, ou criando-os autoralmente com geração automática, sintética e programada. Esta última é a metodologia utilizada pela ferramenta AMUSEVR, que pode ser executada diretamente em modo de autoria e armazenada no projeto, se necessário.

No sistema, os efeitos sensoriais são representados por ícones relacionados ao tipo de efeito reproduzido, como ícone de sol para calor, ícone de vento para vento e assim por diante. Cada efeito sensorial é tratado como um objeto de mídia, com propriedades semelhantes a outros tipos de mídia, mas com a propriedade de alterar a intensidade de emissão do efeito, semelhante a um controle deslizante de volume de áudio, como mostrado na Figura 5.12.

Quanto ao comportamento dos efeitos sensoriais, eles possuem propriedades seme-

lhantes às mídias de áudio, tanto no aspecto visual, pois são representados por um ícone, quanto na emissão e dispersão do efeito. Com emissão a partir de uma localização inicial x , y e z e dissipando em direção ao usuário, dentro do ambiente imersivo é adicionado um efeito visual ao ícone, que pode ser representado por emissões de partículas amarela de calor ou até meio opacas para vento. Isso tem como objetivo complementar a sensação externa que já está sendo enviada ao usuário através de dispositivos atuadores no mundo físico real.

No que diz respeito à comunicação da ferramenta com os dispositivos externos, para a execução dos efeitos sensoriais, foi cogitado adotar inicialmente como em [27], o uso do protocolo MQTT (*Message Queue Telemetry Transport*)⁹, para comunicação entre os dispositivos, bem como o servidor Mosquitto¹⁰. Assim como proposto por projetos relacionados a aplicativos de TV Digital, como [5], e aplicativos voltados para sistemas multimídia que utilizam a tecnologia de Internet das Coisas (IoT) [7]. Saleme et al. [51] escolheram centralizar os controles de atuadores individuais em torno de um servidor. Outro método seria como usada por [44], na qual controlam diretamente o equipamento atuador via sua API (Interface de Programação de Aplicativos), mas nesse caso seria uma programação específica para cada equipamento.

Por fim, nessa implementação do AMUSEVR, cada objeto de efeito sensorial deve acessar um módulo Conector Atuador, conforme mostrado na Figura 4.4. Por sua vez, esse módulo comunica uma mensagem digital, com intensidade variando de 0 a 255 ou ligado e desligado, via seu pino digital para ser interpretado por uma placa Arduino. A Figura 5.20 ilustra as conexões físicas entre atuadores e placas digitais. Na implementação atual, os atuadores disponíveis incluem um ventilador para renderizar o efeito de vento e uma lâmpada infravermelha para renderizar efeitos de calor. O equipamento adicional na Figura 5.20 consiste em um módulo de relés múltiplos, uma placa Arduino UNO, jumpers usados para conexões entre módulos, sem mencionar plugues e fios necessários para fornecer energia elétrica ao equipamento.

⁹<https://mqtt.org/>

¹⁰<https://mosquitto.org/>

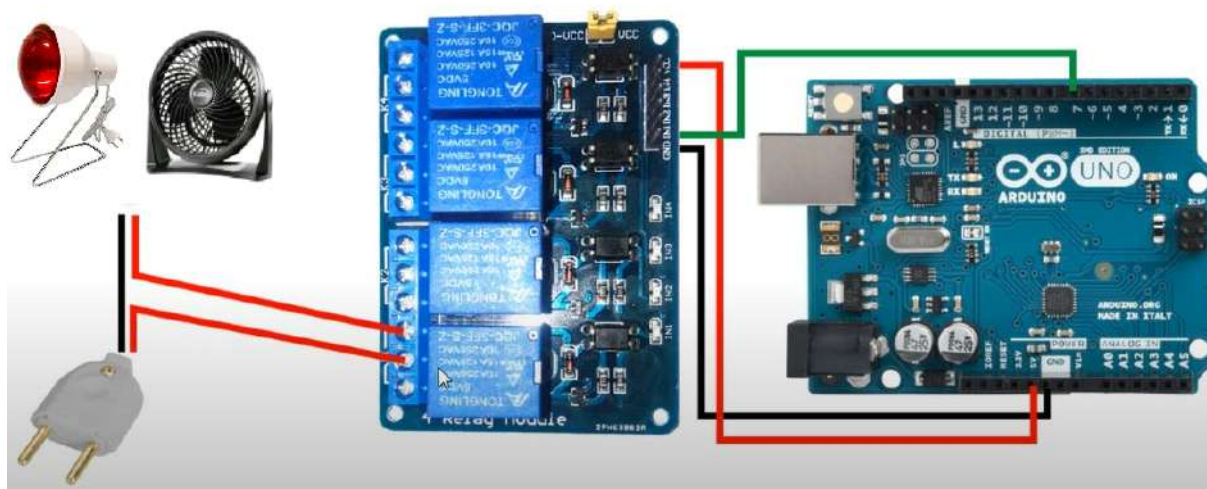


Figura 5.20: Esquema do Arduino.

Na implementação atual, a comunicação com os atuadores é limitada a ligá-los ou desligá-los através do uso de relés, o código para isto para o Arduino pode ser visto no Apêndice A.1. No entanto, diferentemente de sinais binários, as especificações do AMUSEVR podem ser implementadas usando uma comunicação mais sofisticada com os atuadores, por exemplo, controlando a intensidade dos efeitos sensoriais, necessitando pouca modificação física na ponta.

Um vídeo que demonstra o uso do AMUSEVR com efeitos sensoriais está disponível¹¹.

5.2.6 Gerenciamento de mídias e sincronização

Em modo autoria, durante o desenvolvimento do projeto, é possível manipular múltiplas mídias e efeitos sensoriais ao mesmo tempo, com suas execuções em paralelo. Para isso, foi necessário implementar um sistema regente (formatador) de mídias (módulo Sincronizador de Conteúdo da Figura 4.5), que cuida da sincronização de todos os objetos de mídia a cada *frame* na execução.

Para sincronizar as atividades dos objetos de mídia em uma cena do AMUSEVR, adotamos uma metodologia híbrida, agregando e adaptando os conceitos de linhas de tempo e paradigmas de relacionamento baseado em evento, para os objetos de mídia do projeto. Isso porque, em um ambiente tridimensional, com um modelo de autoria imersiva, não há o mesmo contexto de aplicações desktop bidimensionais. Portanto, a utilização de técnicas de formatação multimídia consolidadas em ferramentas exclusivamente bidimensionais

¹¹<https://youtu.be/RIEbnJBqCeo>

precisa ser adaptada ao novo contexto.

Ao descrever o gerenciador de conteúdo aplicado à AATI, foi implementado um gerenciador de cenas, que atua como regente do cenário. Esse gerenciador reconhece cada cena como uma unidade autônoma, com seu próprio relógio discreto e independente, que é acionado quando invocado pelo projeto. Além disso, cada cena não tem permissão de acesso ao conteúdo de outras cenas. O relógio da cena serve como referência de tempo para todos os objetos de mídia contidos nela. Pode-se dizer que essa característica foi inspirada no consolidado paradigma das linhas de tempo.

No entanto, os objetos de mídia também respondem a eventos de interação ou relacionamento entre outras mídias da mesma cena. Esses eventos são capazes de interromper ou alterar a linha de tempo dos objetos da cena. Porém, devido a problemas de usabilidade, a gestão das linhas de tempo não é acessada diretamente pelo autor em modo autoria. Problemas semelhantes podem ser vistos em trabalhos como [11, 47, 58], que, por não considerarem adequado implementar interfaces clássicas de linha de tempo em ambientes imersivos, propuseram a gestão da orquestração das linhas de tempo em ambiente desktop, para posteriormente visualizar em ambiente de realidade virtual.

Para o AMUSEVR, caso o autor deseje orquestrar sua apresentação de mídias, ele tem à sua disposição a edição individual de cada objeto de mídia. Cada um possui suas próprias propriedades de início, término e atraso, que são exibidas e configuradas ou não dependendo da preferência do autor, como visto na Figura 5.7.

Outra forma de controlar a apresentação de objetos de mídia em cena é por meio de uma adaptação do paradigma de relacionamento baseado em eventos, na qual é possível associar o início e o fim da exibição de um objeto de mídia a outro objeto de mídia na mesma cena. Na Figura 5.5 e também localizado no item 3 na Figura 5.6, o menu “*Starts*” define uma relação inicial com o objeto de mídia “Image360” através do comando “*OnBegin*”. Ou seja, o objeto em edição, denominado “Video 2D - 1”, só iniciará quando o objeto “Image360” for iniciado. Para o mesmo objeto “Video 2D - 1”, temos outra propriedade chamada “*Ends*”, que relaciona o fim da exibição do “Video 2D - 1” em (“*OnEnd*”), com o fim da exibição do objeto “Audio3D - 1” da mesma cena. Assim, o objeto de mídia identificado como “Video 2D - 1” somente terminará quando o objeto “Audio3D - 1” terminar.

Outro exemplo da aplicabilidade da adaptação do paradigma de relacionamento para o AMUSEVR pode ser visto na Figura 5.7. Onde há a vinculação de relacionamento do objeto de mídia identificado como "Audio". Neste mesmo exemplo, no menu "*Starts*", o

objeto de áudio será iniciado juntamente com a cena, quando a cena for iniciada. Esse relacionamento é descrito através da opção "*OnBegin This Scene*". Já no menu "*Ends*", o objeto de áudio será finalizado, assim que a cena for finalizada. Esse relacionamento é descrito através da opção "*End of This Media Time*". Em resumo, foi vinculado e atrelado o início e o fim do objeto de mídia "Audio" à duração do relógio discreto da linha do tempo da cena, o que poderá ser interrompido prematuramente somente através do acionamento de uma interação em outro objeto da mesma cena. Essas propriedades de relacionamento estão disponíveis a todas as mídias em cena, inclusive as mídias em 360 graus.

5.2.7 Interatividade

A ferramenta foi projetada para fornecer interatividade programada pelos autores nas apresentações. Com essa função, é possível usar mídias interativas, como textos, imagens, vídeos e botões, simulando subsistemas com ações nas apresentações, onde os usuários podem interagir e até mesmo alterar o conteúdo da apresentação em tempo de execução. O fluxograma dos comandos interativos são executados segundo o diagrama da Figura 5.21.



Figura 5.21: Fluxo de interatividade.

No exemplo da Figura 5.16 e também localizado no item 4 na Figura 5.6, é exibido o modo de edição para o objeto identificado como "Video 2D - 1", que apresenta o conteúdo de um vídeo de uma pequena foca em uma praia que está presente atrás de seu próprio menu de edição semitransparente de título *Edit Video 2D - 1*. Nesta mesma figura, há outra mídia do tipo áudio 3D, representada por um ícone localizado à esquerda do

vídeo da foca e também uma imagem em 360° panorâmica ao fundo, representando uma cidade. Nesta mesma cena, no menu de edição do vídeo em destaque, é possível observar que a ferramenta fornece métodos interativos através do menu de comando binário "*Is interactive?*". Com a ativação desse atributo, o objeto configurado ganha propriedade semelhante a um botão, além de um ícone visual em formato de "*i*" no canto superior direito, como visto no vídeo na Figura 5.22 ou na Figura 4.3.



Figura 5.22: Cena com vídeo interativo.

Esta função interativa pode ser executada a qualquer momento, com um clique no objeto empregado em modo de visualização ou pré-visualização. Como as ações que reagem à execução de funções de interação, é possível exibir ou ocultar uma mídia ou cena quando isso é definido em modo de autoria. Com isso, por exemplo, é possível transitar entre cenas ao clicar em objetos de mídia, exibir ou ocultar uma legenda, ativar um PIP de tradução em LIBRAS, alterar o áudio ou versão de um vídeo, tudo isso, a depender das decisões interativas do usuário. Essas funções interativas garantem liberdade criativa e funcionalidades diversas às apresentações disponíveis ao autor do projeto.

5.2.8 Gerenciamento de Arquivos

O gerenciamento de arquivos da ferramenta é dividido em Gerenciador de Mídias e Gerenciador de Arquivos.

O gerenciador de mídias, que pode ser visto no diagrama de componentes da Figura 4.5, trabalha carregando um vetor de endereços de todo conteúdo de mídia do diretório de mídias do projeto e, em seguida, oferecendo visualmente uma listagem em um menu

de lista suspensa (*dropdown*) (item 2 da Figura 5.6). Para cada tipo de objeto de mídia, o vetor de endereços é filtrado buscando compatibilidade, por exemplo, para um objeto de mídia áudio, somente os endereços com extensões de áudio compatíveis como MP3 e WAV estarão disponíveis para o autor.

Na versão do AMUSEVR para Vive, para a inclusão/importação de novas mídias ao projeto, é possível adicioná-las a qualquer momento diretamente (manualmente) nas pastas do projeto, ou via gerenciador de arquivo do Windows, somente em tempo de execução. Neste segundo caso, o comando é invocado pelo sistema através de uma nova janela do Windows Explorer ao clicar no botão "*Upload Files*" (item 9 da Figura 5.6), presente nos menus de edição de alguns objetos.

O Gerenciador de arquivos é responsável por gerenciar os arquivos gerados e lidos externamente pelo sistema que não são propriamente mídias. A maior parte do trabalho dele consiste em gerenciar os arquivos do projeto e manipular suas cenas. A Unity tem como característica limpar todo o conteúdo da memória ao transitar entre as cenas de qualquer tipo de projeto. Portanto, é necessário salvar os dados antes de sair da cena (salvar o estado da cena), para que, quando for necessário retornar novamente, a cena possa ser reconstruída com todo o conteúdo em edição (carregar o estado antigo da cena).

Para tal, existem vários métodos para salvar estados, como a gravação dos dados com tecnologias JSON¹², mas foi decidido usar a técnica de representação com arquivo usando a metalinguagem XML própria. Portanto, a cada transição de cena é criado um estado da cena com todos os dados necessários em um arquivo temporário em XML e, se necessário, é carregado um estado para nova cena de outro arquivo XML referente ao início da cena de destino.

5.2.9 Projetos utilizando MultiSEL

MultiSEL (*Multimedia Sensory Effect Language*) é uma linguagem baseada em XML para especificar aplicações multimídia com conteúdo visual de 360° para ambientes de realidade virtual, linguagem essa já abordada na Seção 2.5.

Em nossa AATI, AMUSEVR usa MultiSEL como formato de intercâmbio entre o Modo Autor e o Modo Visualizador. Portanto, um autor pode criar ou editar uma aplicação imersiva de 360° usando o Modo Autor e distribuí-la posteriormente para os usuários finais experimentarem usando o Modo Visualizador.

¹²<https://www.json.org/json-en.html>

Escolhemos essa linguagem pois ela demonstrou ser adequada para autoria imersiva em ambientes 3D, além disto, é possível se beneficiar da clareza fornecida pela estrutura XML. MultiSEL foi projetada como base para futuro suporte para ferramentas que trabalham com múltiplas cenas 3D, mídias e efeitos sensoriais simultaneamente, além de fornecer conceitos como relações de interação. Como nosso AATI é uma abordagem de autoria e edição usando controles dentro de um ambiente imersivo, torna-se possível para os autores começarem a desenvolver sua aplicação em um editor 2D, como STEVE [25], exportando um projeto intermediário em MultiSEL e, em seguida, importando o projeto MultiSEL em AMUSEVR. Alternativamente, autores especialistas podem simplesmente criar seu projeto MultiSEL diretamente em XML e, em seguida, aprimorar e finalizar seu projeto em AMUSEVR. As possibilidades fornecidas podem ser vistas na Figura 5.23.

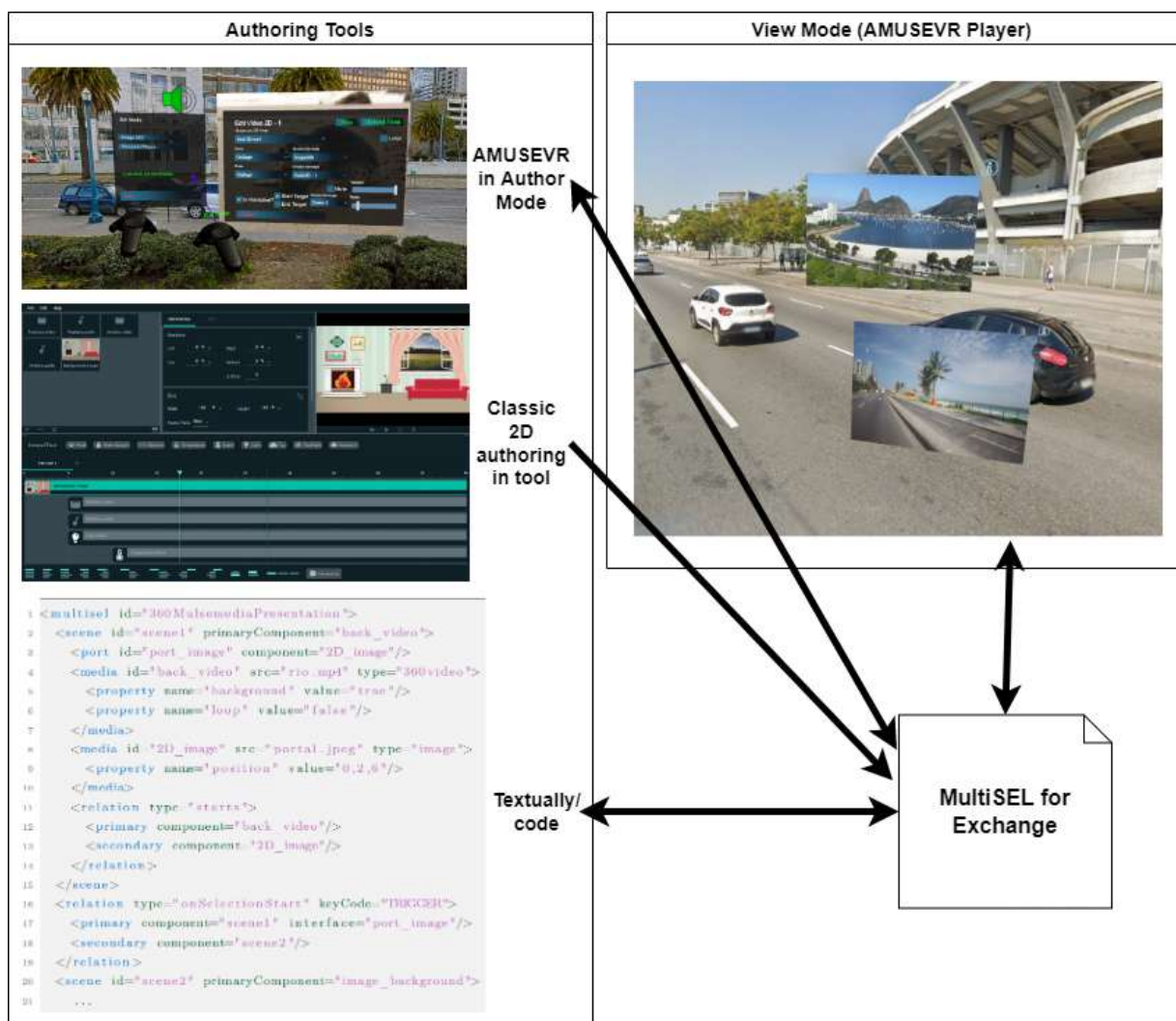


Figura 5.23: Exemplo de portabilidade de um documento MultiSEL.

Um exemplo de código MultiSEL para AMUSEVR é dado na Listagem 5.1. Este

documento MultiSEL define duas cenas: *scene1* e *scene2*. *scene1* começa com a mídia *video_background* como seu pano de fundo 360°. Ele também contém um elemento de relação cujo tipo é *starts*. Essa relação define que o componente secundário (*2D_image*) começa quando o principal (*video_background*) começa. Incluído no elemento *body*, o documento especifica uma relação do tipo *onSelectionStart* entre ambas as cenas. Ele define que *scene2* começa quando a mídia *2D_image* é acessada através do elemento *port* em *scene1* e selecionada usando a tecla de gatilho. *scene2* contém um efeito sensorial de vento que começa com um atraso de 10s quando *image_background* começa. *image_background* é um objeto de mídia contido em *scene2*, mas não é mostrado na Listagem 5.1.

```

1 <?xml version="1.0" encoding="utf-8"?>
2 <multisel id="360MulsemmediaPresentation" xmlns="DefaultProfile">
3   <head>
4     <meta name="author" value="anonymous" />
5     <meta name="year" value="2022" />
6   </head>
7   <body primaryComponent="scene1">
8     <scene id="scene1" primaryComponent="video_background">
9       <port id="port_image" component="2D_image"/>
10      <media id="video_background" src="rio.mp4" type="360video">
11        <property name="background" value="true"/>
12        <property name="loop" value="false"/>
13      </media>
14      <media id="2D_image" src="portal.jpeg" type="image">
15        <property name="x" value="0" />
16        <property name="y" value="2" />
17        <property name="z" value="6" />
18      </media>
19      <relation type="starts">
20        <primary component="video_background"/>
21        <secondary component="2D_image"/>
22      </relation>
23    </scene>
24    <scene id="scene2" primaryComponent="image_background">
25      ...
26      <effect id="wind_effect" type="windEffect">
27        <property name="intensity" value="50"/>
28        <property name="x" value="0" />
29        <property name="y" value="6" />
30        <property name="z" value="12" />
31      </effect>
32      <relation type="starts" delay="10">
33        <primary component="image_background"/>
34        <secondary component="wind_effect"/>
35      </relation>
36    </scene>
37    <relation type="onSelectionStart" keyCode="trigger">
38      <primary component="scene1" interface="port_image"/>
39      <secondary component="scene2"/>
40    </relation>

```

```

41 </body>
42 </multisel>

```

Listing 5.1: Exemplo de documento MultiSEL para o AMUSEVR.

MultiSEL trata-se de uma linguagem criada para comportar um escopo muito maior que o necessário para o AMUSEVR, mas durante o desenvolvimento da ferramenta, foi necessário criar adaptações na linguagem original de [24] para o contexto do ambiente imersivo do AMUSEVR. A seguir algumas mudanças necessárias, sendo que mais detalhes podem ser vistos no Apêndice B.

- Para identificação da localização no plano cartesiano foram criadas três *tags* `<property>` separadas com nome x, y e z, como mostrado nas Linhas 15 a 17 da Listagem 5.1;
- As *tags* `<body>` e `<scene>` foram refinadas com um atributo chamado *primary-Component*, como mostrado na Linha 7 da Listagem 5.1, sendo que em `<scene>` é opcional;
- Valores numéricos, como os de coordenadas, são identificados como valores em números reais.

Em demonstração da equivalência entre AMUSEVR e MultiSEL, usando diferentes cenários como exemplos, a equivalência da categoria de relação "*Starts*" do AMUSEVR mostrada no item 3 da Figura 5.6 pode ser vista nas linhas 19-22 na Listagem 5.1. Outro caso, envolvendo interatividade, pode ser encontrado no item 4 da Figura 5.6, que representa a relação encontrada nas linhas 37-40 na Listagem 5.1.

5.3 Porte para versão do Oculus Quest 2

Atualmente, o AMUSEVR encontra-se muito mais robusto que na primeira versão AMUSEVR para Vive publicado em [27, 28]. Após essa versão, foi decidida a migração de todo conteúdo que antes foi projetado para o HMD da HTC Vive, para uma nova versão de HMD, o Oculus Quest 2¹³ da empresa Meta (antiga Facebook), portanto definido como versão AMUSEVR Oculus. Esta decisão foi tomada, pois havia dificuldade em transportar todo o equipamento da antiga versão. Com isto, foi verificado que o HMD Quest 2 estava tornando-se mais popular tanto pelo quesito preço, quanto pela praticidade de não

¹³<https://www.meta.com/quest/products/quest-2/>

necessitar de ancoramentos ou cabos como o HTC Vive. Como desvantagem em realizar o porte, perderíamos poder computacional e partiríamos de uma plataforma final Windows para Android. Isto gerou um grande trabalho de reconstrução de interface, recriação de códigos, adaptação de ativos e principalmente muitos testes. Um vídeo de demonstração da nova versão está disponível¹⁴ e demonstra a criação de um projeto sobre um passeio turístico no Museu de Arte Contemporânea (MAC) em Niterói no Rio de Janeiro.

Após a migração ter sido terminada com considerada equivalência, começou o processo de inclusão de novos módulos não presentes na versão anterior, como os módulos de Modo Visualização, módulo de gerenciamento de projetos em MultiSEL e o módulo de Efeitos Sensoriais. Algumas funções foram perdidas como a de importar arquivos de mídias diretamente em tempo de execução por dificuldades implantadas pela plataforma Android do próprio Oculus Quest. Mas foram incorporadas funções adicionais relacionadas à usabilidade, como possibilidade de agarrar os objetos de mídias e janelas presentes no projeto e girá-los em qualquer eixo. Além disso, também é possível afastar ou aproximar os menus de configuração flutuantes em formato de janelas através das alavancas do controle favorecendo pessoas com problemas de visão ao ler conteúdos textuais.

O próximo capítulo aborda a avaliação da abordagem de autoria imersiva desta tese, relatando todos os experimentos com usuários e resultados obtidos.

¹⁴<https://youtu.be/80hBuyKGK7o>

Capítulo 6

Avaliação

Para avaliar o ambiente AMUSEVR, foram realizados, ao todo, três experimentos com usuários. O primeiro foi referente à primeira versão do AMUSEVR Vive [27]. O segundo experimento foi relacionado à linguagem MultiSEL para o AMUSEVR, já em sua versão Oculus. Por fim, um último experimento mais completo em relação ao primeiro, envolvendo efeitos sensoriais e MultiSEL para o AMUSEVR, também em sua versão Oculus. Os experimentos foram planejados com a metodologia *Goal Question Metric* (GQM), que facilita sua reprodução e foi em parte publicada em [28]. A Tabela 6.1 mostra um resumo comparativo entre os três experimentos.

Tabela 6.1: Resumo comparativo dos experimentos.

Experimento	Metodologia e Ferramentas	Participantes	Realizado
AMUSEVR Vive	<i>Goal Question Metric</i> (GQM); Escala de Usabilidade do Sistema (SUS); Questionário de Experiência do usuário (UEQ); Questionário Personalizado.	10	05/2021
MultiSEL	<i>Goal Question Metric</i> (GQM); <i>Cognitive Dimensions of Notations</i> (CDN) Questionário Personalizado; Método estatístico Mann-Whitney.	20	05/2022
AMUSEVR Oculus	<i>Goal Question Metric</i> (GQM); Escala de Usabilidade do Sistema (SUS); Questionário de Experiência do usuário (UEQ); Comparativo Quantitativo; Questionário Personalizado; Método estatístico Mann-Whitney.	42	06/2022

6.1 Experimento com AMUSEVR Vive

O primeiro experimento foi realizado com o AMUSEVR Vive e teve como foco avaliar a viabilidade da ferramenta como um todo, para tal visando investigar a pergunta principal de pesquisa desta tese: "*Qual é a experiência do autor na criação de aplicações multissensoriais interativas em 360 graus usando um ambiente de autoria imersivo?*", foi utilizada a metodologia *Goal Question Metric* (GQM) [6] para organizar os objetivos do experimento, foram utilizados o Questionário de Experiência do Usuário (UEQ - *User Experience Questionnaire*)¹ [40, 55] e o questionário Escala de Usabilidade do Sistema (SUS - *System Usability Scale*)² [10]. A seguir, serão discutidos a metodologia aplicada e os resultados obtidos.

¹<https://www.ueq-online.org/>

²<https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html>

Com o objetivo de descrever o processo, a metodologia usada para estruturar e elaborar a avaliação e experimentos é discutida na Seção 6.1.2. Na Seção 6.1.3, apresentam-se os resultados obtidos e finalmente, na Seção 6.1.6, realiza-se uma análise e discussão das conclusões e das limitações do experimento.

6.1.1 Metodologia

O método *Goal Question Metric* (GQM) [6] foi escolhido para estruturar esta avaliação. O mecanismo GQM pode ser visualizado como um fluxograma no qual, seguindo o conceito de *top-down*, o fluxo começa nos nós de objetivos, continua para os nós de questões e termina nos nós de métricas. Além disso, as diretrizes GQM visam a definir a proposta e a perspectiva dos objetivos. A proposta define o objeto sendo estudado e os motivos para o seu estudo. A perspectiva define o aspecto pessoal ou ponto de vista dado pelo avaliador.

A aplicação do método GQM resulta em um modelo de hierarquia de três camadas: conceitual (objetivos), operacional (questões) e quantitativo (métricas). A definição do modelo tem o estabelecimento de objetivos como um passo inicial. Cada objetivo definido é desdobrado em questões que visam a definir um elemento de experiência, produzindo resultados quantificáveis. As questões, por sua vez, são aprimoradas em métricas objetivas ou subjetivas.

A Tabela 6.2 mostra os objetivos definidos para a avaliação da ferramenta AMUSEVR Vive. Para cada objetivo G1-G3, a coluna "Descrição" apresenta o objeto sendo analisado.

Tabela 6.2: Os objetivos definidos para os experimentos AMUSEVR Vive

Objetivo	Descrição
G1	Analisar a ferramenta com o objetivo de avaliar sua usabilidade a partir da perspectiva do usuário.
G2	Analisar a ferramenta com o objetivo de avaliar sua qualidade de experiência a partir da perspectiva do usuário.
G3	Analisar a ferramenta com o objetivo de avaliar o esforço de autoria exigido tanto para usuários experientes quanto para iniciantes.

6.1.1.1 G1 - Questões e Métricas

Para atingir o objetivo G1, ou seja, analisar a ferramenta com o objetivo de avaliar sua usabilidade do ponto de vista do usuário, foi utilizado o questionário *System Usability Scale* (SUS) [10]. O SUS especifica um questionário de dez itens mostrados na Tabela 6.3, onde cada um aceita respostas na escala Likert, variando de "Discordo fortemente" classificado como 1 até "Concordo fortemente" classificado como 5 [42]. Usamos a pontuação do SUS como métrica para as perguntas de G1.

Tabela 6.3: Questionário do SUS.

Q#	Questões
SUS-Q1	Eu acho que eu gostaria de usar este sistema frequentemente.
SUS-Q2	Eu achei que o sistema era desnecessariamente complexo.
SUS-Q3	Eu achei que o sistema era fácil de usar.
SUS-Q4	Eu acho que eu precisaria do apoio de uma pessoa técnica para poder usar este sistema.
SUS-Q5	Eu achei que as várias funções deste sistema estavam bem integradas.
SUS-Q6	Eu achei que havia muita inconsistência neste sistema.
SUS-Q7	Eu imagino que a maioria das pessoas aprenderia a usar este sistema muito rapidamente.
SUS-Q8	Eu achei o sistema muito incômodo de usar.
SUS-Q9	Eu me senti muito confiante ao usar o sistema.
SUS-Q10	Eu precisava aprender muitas coisas antes de conseguir começar a usar este sistema.

6.1.1.2 G2 - Questões e Métricas

Para atingir o objetivo G2, isto é, analisar a ferramenta com o objetivo de avaliar sua qualidade de experiência do ponto de vista do usuário, foi utilizado o Questionário de Experiência do Usuário (UEQ-S) [40, 55]. Como pode ser visto na Tabela 6.4, o questionário contém oito itens (UEQ-Q1 a UEQ-Q8) com feedback negativo ou positivo em resposta à frase "A experiência com AMUSEVR foi". As respostas dos participantes foram numa

escala de (-3) (feedback negativo) a (+3) (feedback positivo). Foram usados os scores pragmáticos e hedonistas do UEQ como métricas para as perguntas de G2.

Tabela 6.4: Questionário UEQ-S.

Itens	A experiência com AMUSEVR foi:	
	Negativo	Positivo
UEQ-Q1	Obstrutivo	Autoexplicativo
UEQ-Q2	Complicado	Fácil
UEQ-Q3	Ineficiente	Eficiente
UEQ-Q4	Confuso	Claro
UEQ-Q5	Entediante	Excitante
UEQ-Q6	Desinteressante	Interessante
UEQ-Q7	Convencional	Inovador
UEQ-Q8	Comum	Vanguardista

6.1.1.3 G3 - Questões e Métricas

Para atender ao objetivo G3, ou seja, analisar a ferramenta com o objetivo de avaliar o esforço de autoria necessário tanto para usuários experientes quanto iniciantes, foi utilizado um conjunto de perguntas específicas, como pode ser visto na Tabela 6.5. A Q1 verificou se o participante conseguiu completar uma tarefa e, caso positivo, o tempo necessário para o completamento da tarefa foi anotado na Q4. As perguntas Q2 e Q3 foram destinadas a identificar os usuários que possuíam experiência anterior em realidade virtual e ferramentas de autoria multimídia, separando os experientes dos iniciantes.

Tabela 6.5: Perguntas personalizadas para G3.

Q#	Perguntas
Q1	Você conseguiu finalizar a tarefa?
Q2	Indique seu nível de experiência na utilização de realidade virtual ou ambientes imersivos.
Q3	Indique seu nível de experiência na utilização de ferramentas de autoria multimídia.
Q4	Quantos minutos você levou para completar a tarefa?

As questões Q2 e Q3 foram projetadas para separar os participantes em dois grupos,

aqueles que já se consideravam com alguma experiência em ambientes imersivos e os que se consideravam com alguma experiência com ferramentas de autoria. Cada questão aceita respostas na escala Likert, variando de "Nenhuma experiência" classificada como 1 até "Muita experiência" classificada como 5 [42]. A hipótese é que os participantes com experiência em ambientes imersivos e com ferramentas de autoria multimídia teriam um desempenho melhor do que aqueles sem experiência.

O tempo médio de edição para concluir a tarefa solicitado na questão Q4 foi usado como métrica para comparar o esforço de autoria entre usuários experientes e novatos/iniciantes.

6.1.1.4 Grafo GQM

A Figura 6.1 resume e ilustra a estrutura do modelo GQM construído. As experimentações realizadas para a coleta de dados serão apresentadas nas seções subsequentes. As experimentações foram criadas de forma a atender aos objetivos G1, G2 e G3.

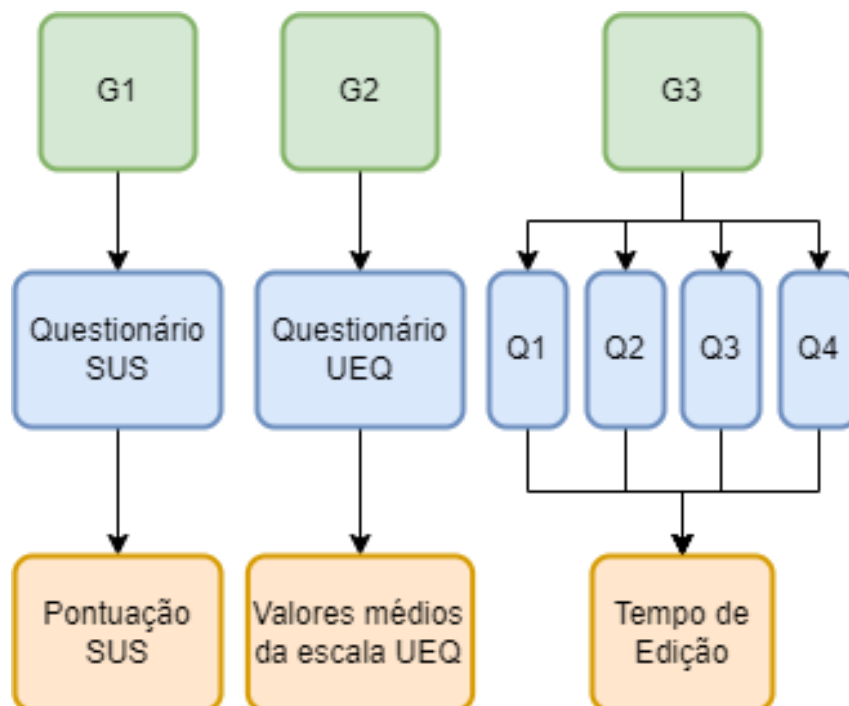


Figura 6.1: Diagrama direcionado com Objetivos, Questões e Métricas (GQM) do experimento com AMUSEVR Vive.

6.1.2 Procedimentos com Participantes

Os procedimentos experimentais com usuários foram realizados considerando os objetivos definidos na Seção 6.1 para investigar a pergunta principal de pesquisa desta tese: *"Qual é a experiência do autor na criação de aplicações multissensoriais interativas em 360 graus usando um ambiente de autoria imersivo?"*. Devido à pandemia de COVID-19 e às restrições sanitárias relacionadas ao uso e compartilhamento de ambientes e equipamentos, não foi possível realizar o primeiro experimento com um grande número de participantes. Portanto, foram realizados neste experimento testes com dez (10) usuários, sendo oito homens e duas mulheres com idades entre 23 e 47 anos, com escolaridade variando entre graduação e pós-graduação. Metade dos participantes (cinco) não havia utilizado equipamento de realidade virtual anteriormente, enquanto quatro dos dez não tinham experiência alguma com sistemas de autoria multimídia. Além disso, um usuário era canhoto e um possuía estrabismo.

Os participantes tiveram a opção de escolher entre realizar o experimento em pé ou sentado, conforme ilustrado na Figura 6.2. Todos os experimentos foram realizados entre dezembro de 2020 e maio de 2021, no Laboratório MídiaCom da Universidade Federal Fluminense, em Niterói, Rio de Janeiro, utilizando os equipamentos fixos da HTC Vive. Durante os experimentos, havia sempre a presença de um supervisor responsável por coletar informações sobre a experiência do usuário e possíveis melhorias no software. Além disso, havia um auxiliar ao lado do participante, responsável por auxiliar na colocação do capacete de realidade virtual e uso dos demais acessórios, evitando qualquer tipo de acidente. Durante os testes, era exibida uma tela adicional para monitoramento em tempo real, replicando que o participante visualizava no capacete de RV, como também pode ser observado na Figura 6.2. Um vídeo sobre o experimento está disponível no Youtube³.

³https://www.youtube.com/watch?v=KPFfSj-5_qk



Figura 6.2: Participante no primeiro experimento.

Todos os participantes tiveram um período de tempo para se familiarizar com o ambiente e as ferramentas utilizadas (em média, cinco minutos). Após cada período de familiarização, eles iniciaram as tarefas. Durante as tarefas, alguns dados foram registrados pelo supervisor através de anotações. Ao final de cada experimento, os participantes submeteram suas respostas e observações individuais de forma reservada, através da plataforma do Google Forms⁴.

⁴<https://docs.google.com/forms/>

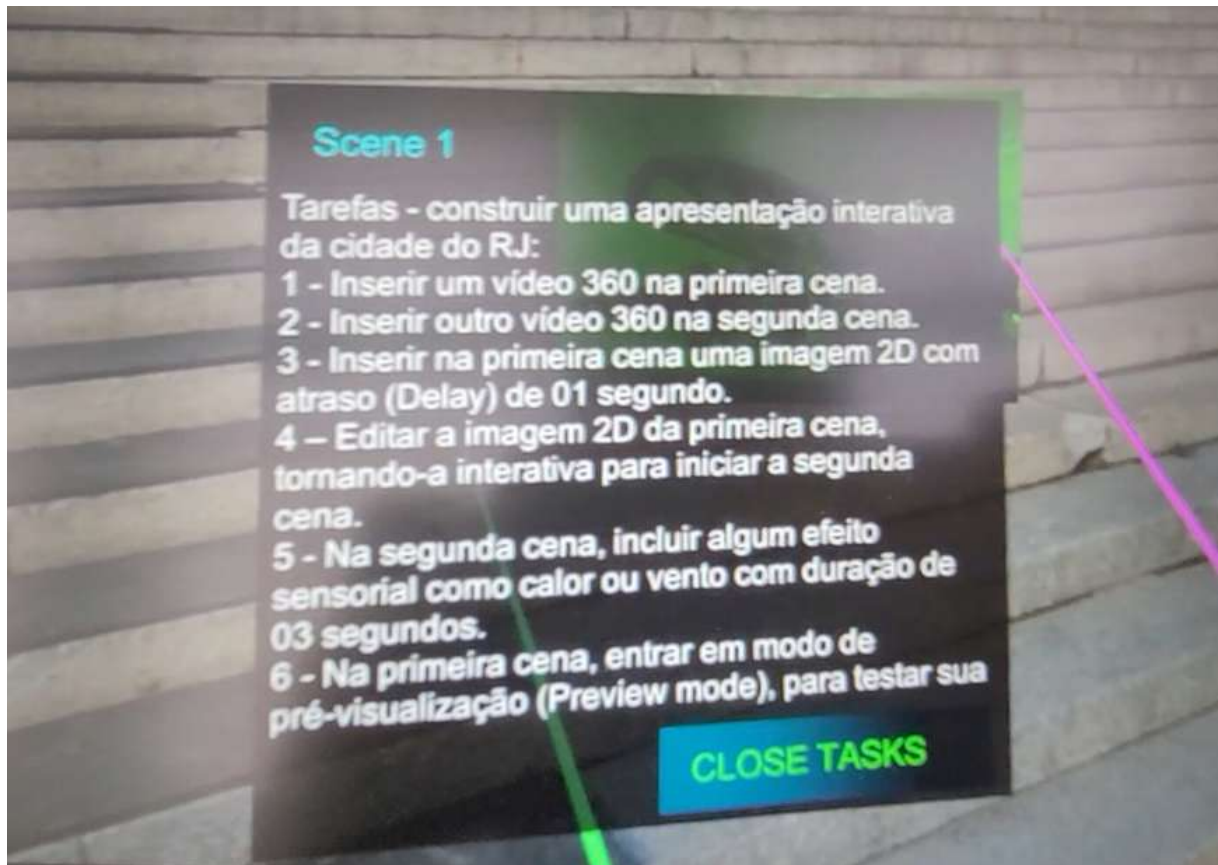


Figura 6.3: Exibição das tarefas no menu virtual.

As tarefas solicitadas aos participantes visavam avaliar a usabilidade e viabilidade do sistema AMUSEVR, através da criação de uma apresentação interativa multimídia em 360°, com a temática de propaganda turística sobre a cidade do Rio de Janeiro. Para orientar os participantes, havia uma lista de tarefas disponíveis no menu virtual do controle em "*Tarefas*" (como visto na Figura 6.3), descrevendo as atividades a serem realizadas, tais como:

1. inserir um vídeo ou imagem 360° na primeira cena;
2. inserir um vídeo ou imagem 360° na segunda cena;
3. inserir na primeira cena uma imagem 2D com atraso (*Delay*) de 1 segundo;
4. editar a imagem 2D da primeira cena, tornando-a interativa para iniciar a segunda cena;
5. Na segunda cena, incluir algum efeito sensorial como calor ou vento com duração de 3 segundos; e

6. Na primeira cena, entrar em modo pré-visualização (*Preview mode*), para testar a apresentação criada.

Após os testes, os participantes foram solicitados a preencher os questionários de Escala de Usabilidade do Sistema (SUS) [10] e o questionário de experiência do usuário (UEQ) [40, 55], bem como responder outras perguntas adicionais sobre idade, nível de escolaridade, experiência com ferramentas de autoria e sugestões para o trabalho. Com isso, os resultados obtidos foram analisados.

6.1.3 Resultados

Nesta subseção, são analisados e discutidos os dados obtidos no experimento e nos questionários, que respondem aos objetivos definidos anteriormente.

6.1.3.1 G1 - Análise dos Resultados

Considerando o objetivo G1, conforme apresentado anteriormente, foi usado o questionário da Escala de Usabilidade do Sistema (SUS) [10], composto por dez questões de escolha única, conforme Tabela 6.3, que utiliza a escala Likert [42] para a pontuação, dividindo as respostas em uma escala de cinco pontos, em relação a uma afirmação. A discordância máxima é representada por um (1) ponto e a concordância máxima por cinco (5) pontos.

Com a aplicação do SUS, obteve-se uma pontuação de aceitabilidade média de 82,25 pontos para o SUS, conforme Figura 6.4, que contabiliza o total de respostas para cada um dos cinco níveis das dez questões para os 10 participantes. Com esta pontuação SUS, pode-se concluir que o sistema é considerado aceitável em relação à sua usabilidade, já que é maior que 68 pontos, que representam a média de base aceitável como citado em [10]. A Figura 6.5 exibe a escala de aceitabilidade e adjetivos do SUS.

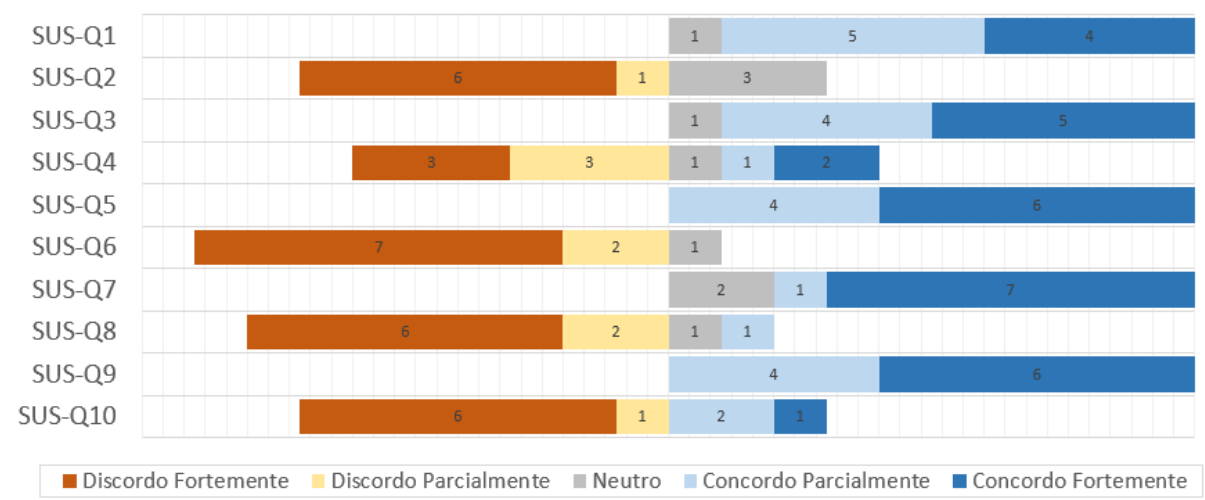


Figura 6.4: Número de respostas às afirmações do questionário do SUS (1 - discordo totalmente; 5 - concordo totalmente).

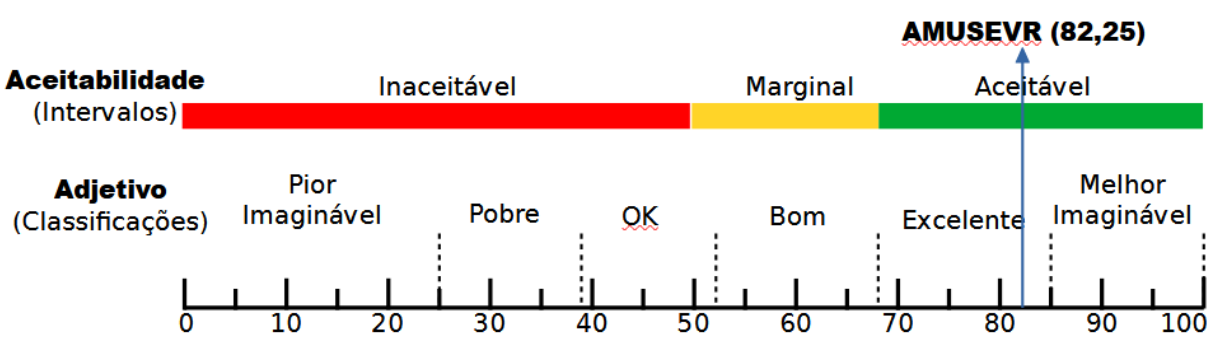


Figura 6.5: Escalas de Aceitabilidade e Adjetivos (pontuação SUS).

Além da escala da relativa à pontuação SUS como aceitável, é possível classificar a usabilidade do AMUSEVR Vive utilizando-se uma escala de adjetivos [3], como também pode ser visto na Figura 6.5. Essa escala é dividida em seis categorias associadas à pontuação SUS. Baseado na escala adjetiva, pode-se classificar o AMUSEVR Vive com o adjetivo "Excelente".

6.1.3.2 G2 - Análise dos Resultados

Para alcançar o objetivo G2, o questionário de Experiência do Usuário simplificado (UEQ-S) foi aplicado de forma complementar. Este questionário utiliza uma escala de 7 níveis de pontuação, variando de (-3) para o pior resultado a (+3) para o melhor resultado. O questionário UEQ-S possui 8 itens com uma escala entre adjetivos negativos e positivos, conforme apresentado na Tabela 6.6. A mesma tabela também indica a pontuação média e a variância das respostas correspondentes. Na Tabela 6.7, é possível observar a pontu-

ação relacionada à resposta de cada um dos dez participantes, para cada uma das oito perguntas.

Tabela 6.6: Resultado UEQ-S obtido.

Qualidade UEQ	Negativo	Positivo	Pont.(-3 à +3)	Variância
Pragmática	Obstrutivo	Autoexplicativo	+1,8	0,8
	Complicado	Fácil	+1,5	1,2
	Ineficiente	Eficiente	+2,4	0,5
	Confuso	Claro	+2,0	1,1
Hedônica	Entediante	Excitante	+2,5	0,5
	Desinteressante	Interessante	+2,6	0,5
	Convencional	Original	+2,6	0,3
	Comum	Vanguardista	+2,5	0,3

Tabela 6.7: Dados UEQ-S obtidos.

UEQ-S	Questões e respostas							
Participantes	1	2	3	4	5	6	7	8
1	1	1	2	1	2	3	2	2
2	2	2	2	2	2	2	2	2
3	0	0	3	3	3	3	3	3
4	3	1	1	2	3	3	3	3
5	3	3	3	3	3	3	3	3
6	2	0	2	0	2	2	2	3
7	2	2	3	3	3	3	3	2
8	1	1	3	1	3	3	3	2
9	2	3	2	2	1	1	2	2
10	2	2	3	3	3	3	3	3

Os resultados apresentados na Tabela 6.6 indicam que metade das questões obteve classificação de "Qualidade Pragmática" (de acordo com a classificação UEQ, média em torno de 2 pontos positivos ou mais) e o restante obteve classificação de "Qualidade Hedônica" (de acordo com a classificação UEQ, média em torno de 2,5 pontos positivos ou mais). Na Figura 6.6, é exibido o gráfico de *benchmark* gerado pelo UEQ, que mostra que a ferramenta alcançou um resultado classificado como excelente, de acordo com os requisitos de avaliação descritos em [40, 55].

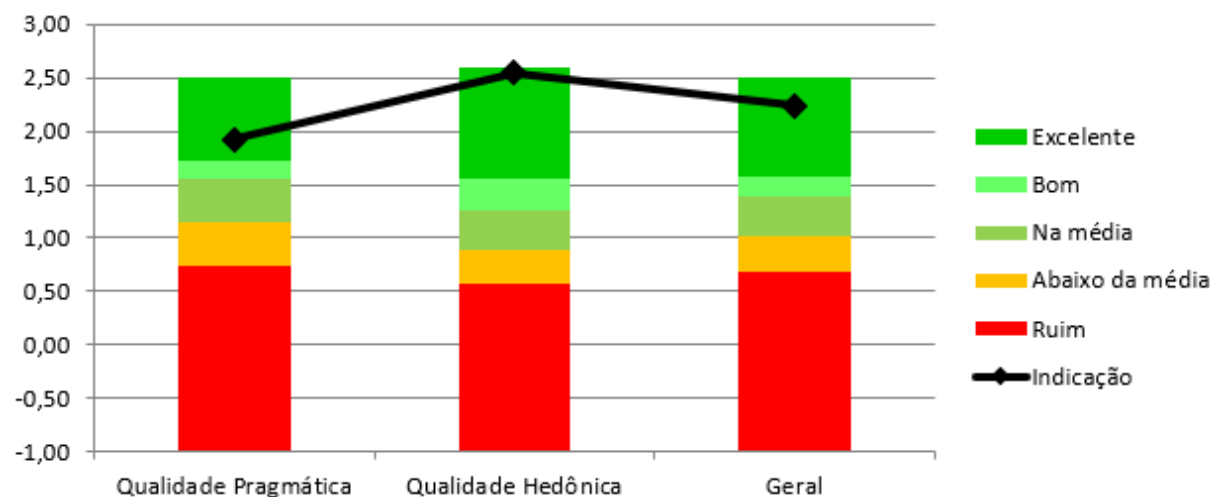


Figura 6.6: Avaliação de desempenho no UEQ (*benchmark*).

6.1.3.3 G3 - Análise dos resultados

As questões Q1, Q2, Q3 e Q4 foram usadas para analisar o objetivo de analisar a ferramenta com o objetivo de avaliar a autoria esforço necessário para usuários experientes e novatos. As respostas dos participantes a essas perguntas são mostradas na Tabela 6.8. Podemos observar que todos os participantes concluíram as tarefas propostas (Q1); 50% dos participantes já haviam utilizado tecnologia de realidade virtual e os outros 50% não (Q2); 60% dos participantes haviam utilizado anteriormente ferramentas de autoria e os outros 40% não (Q3); e o tempo médio para concluir as tarefas com o AMUSEVR (Q4) foi de 7,6 minutos, com um desvio padrão de 2,87 minutos para todo o grupo. O tempo necessário para os diferentes grupos de usuários experientes e novatos em realidade virtual é apresentado na Tabela 6.9. A Tabela 6.10 apresenta o tempo necessário considerando usuários experientes e novatos em ferramentas de autoria multimídia.

Tabela 6.8: Respostas dos participantes.

Participante	Q1	Q2	Q3	Q4 (minutos)
1	Sim	Sim	Sim	14
2	Sim	Não	Não	6
3	Sim	Sim	Não	7
4	Sim	Sim	Sim	6
5	Sim	Não	Sim	7
6	Sim	Sim	Sim	5
7	Sim	Não	Não	10
8	Sim	Não	Não	4
9	Sim	Não	Sim	9
10	Sim	Sim	Sim	8

Tabela 6.9: Tempo dos usuários experientes e novatos em VR para concluir o experimento.

Usuários de VR	Tempo em minutos	
	Média	Desvio padrão
experientes	7,27	4,83
Novatos	6,54	4,11

Tabela 6.10: Tempo para usuários experientes e novatos em ferramentas de autoria multimídia para concluir o experimento.

Usuário da ferramenta de autoria	Tempo em minutos	
	Média	Desvio padrão
experientes	8,90	4,84
Novatos	4,90	3,77

6.1.4 Conclusões

Neste estudo, nosso primeiro experimento exploratório resultou em alguns resultados promissores. Portanto, podemos concluir que:

- Considerando G1, o objetivo foi alcançado, pois a pontuação SUS de 82,25 pontos mostra que a usabilidade da ferramenta AMUSEVR Vive foi excelente do ponto de vista do usuário, como mostrado na Figura 6.5;

- Quando consideramos G2, temos o objetivo alcançado pelos resultados do UEQ, que indicaram que a experiência do usuário com a ferramenta AMUSEVR Vive foi excelente considerando o benchmark UEQ, como mostrado na Figura 6.6;
- E considerando G3, concluímos que o AMUSEVR pode ser usado por usuários iniciantes, pois eles foram capazes de criar uma experiência interativa de 360 graus. Por outro lado, ao comparar usuários experientes e iniciantes, nossos resultados foram diferentes do que esperávamos, pois os participantes com mais experiência em VR e ferramentas de autoria levaram em média mais tempo para concluir as tarefas de autoria. Acreditamos que os participantes experientes estavam mais interessados em explorar as instalações do AMUSEVR Vive e conhecer o ambiente imersivo em mais detalhes do que os usuários iniciantes. Portanto, isso explica por que eles levaram mais tempo para usar o sistema de autoria.

6.1.5 Análise Subjetiva

Neste primeiro experimento foi observado que o tempo de ambientação livre antes da atividade foi muito útil para que as pessoas que nunca tiveram contato com as tecnologias apresentadas se sentissem mais confortáveis, perdendo possíveis receios.

Durante a atividade alguns participantes, principalmente os mais experientes, dedicavam mais tempo em escolher melhores mídias e explorar melhor as opções, de forma a ter uma apresentação mais bonita, o que normalmente influenciava no tempo final da tarefa.

Mesmo considerando que foram poucos participantes no primeiro experimento, havia uma possibilidade de ocorrer algum desconforto com algum participante, o que felizmente não aconteceu.

6.1.6 Discussão

Este primeiro experimento serviu para demonstrar que a abordagem empregada em uma autoria realizada totalmente de forma imersiva é viável. Com isto foi possível expandir o projeto e realizar novos experimentos com mais participantes avaliando mais características da abordagem proposta.

Em relação à pergunta adicional que solicitava sugestões e comentários sobre a ferramenta, foram obtidas sugestões que inclusive já foram atendidas na nova versão do AMUSEVR Oculus, tais como:

- Portabilidade – foi sugerido um dispositivo portátil, semelhante ao Google Cardboard⁵, com o objetivo de melhorar o acesso através de dispositivos móveis;
- Interface do usuário - foram sugeridas melhorias na interface, como aumento de algumas fontes e janelas de aviso para ações como a exclusão de objetos;
- Adições – foram sugeridas algumas adições à ferramenta, como a inclusão de objetos em 3D, um menu de tutorial e mais opções de edição de objetos de mídia;
- Elogios – foram relatados alguns elogios em relação à imersão, ao vanguardismo e satisfação de uso.

Sobre as sugestões apresentadas pelos participantes do experimento com AMUSEVR Vive, as seguintes foram implantadas no AMUSEVR Oculus:

- Portabilidade – a plataforma Oculus Quest 2, que é uma plataforma portátil, foi escolhida para migrar o sistema devido a um bom suporte da comunidade de desenvolvimento, dentre outros fatores.
- Interface do usuário – através da sugestão dos primeiros participantes, observamos que seria interessante adotar métodos que melhorassem a visualização dos textos dos menus e objetos em cena, para isso foi incluindo um comando para agarrar e com isso, aproximar ou afastar objetos do usuário, como janelas de menu ou objetos do projeto. Além disso, algumas mensagens pedindo confirmação de ação foram incluídas.
- Adições – Foram adicionadas diversas funções de edição na nova versão, mas a inclusão de edição de mídias não foi realizada por fugir do escopo do projeto. Sobre objetos 3D, apesar de possível, ainda não foram adotados por demandar um grande esforço de implementação e correção de possíveis erros.

6.2 Experimento com a Linguagem MultiSEL para AMUSEVR Oculus

O segundo experimento teve como objetivo avaliar a linguagem MultiSEL, detalhada na Seção 2.5, aplicada ao AMUSEVR como formato de intercâmbio. Foram realizados experimentos em que os usuários criaram seus códigos MultiSEL usando uma ferramenta de

⁵<https://arvr.google.com/cardboard/>

edição regular XML com o esquema XML MultiSEL (disponível no Apêndice B.2). A ferramenta de edição *Visual Studio Code* foi sugerida, mas os participantes podiam escolher outra conforme desejavam. Foi utilizada a metodologia *Goal Question Metric* (GQM) [6] para planejar o experimento e o framework de Dimensões Cognitivas de Notações (*Cognitive Dimensions of Notations* - CDN) [9] para avaliar a usabilidade da linguagem-MultiSEL.

A fim de descrever o processo, a metodologia utilizada para estruturar e elaborar nossa avaliação é discutida na Seção 6.2.1. O experimento com usuários é discutido na Seção 6.2.2. A Seção 6.2.3 apresenta os resultados obtidos e finalmente, na Seção 6.2.5, é feita uma análise e discussão das conclusões e limitações atuais.

6.2.1 Metodologia

Assim como no primeiro experimento o método *Goal Question Metric* (GQM) [6] foi escolhido para estruturar esta avaliação.

A Tabela 6.11 mostra os objetivos definidos para a avaliação da linguagem MultiSEL. Para cada objetivo G1 e G2, a coluna "Descrição" apresenta o objeto sendo analisado.

Tabela 6.11: Objetivos do experimento MultiSEL

Objetivo	Descrição
G1	Analisar a linguagem MultiSEL com o objetivo de avaliar o entendimento sobre a linguagem e o esforço de autoria exigido tanto para usuários experientes quanto para novatos.
G2	Analisar a linguagem MultiSEL com o objetivo de avaliar sua usabilidade do ponto de vista dos usuários.

6.2.1.1 G1 - Questões e Métricas

Para atingir o objetivo G1, ou seja, analisar a linguagem MultiSEL com o objetivo de avaliar o entendimento sobre a linguagem e o esforço de autoria necessário tanto para usuários experientes quanto novatos, foi usado um conjunto de questões específicas, como pode ser visto na Tabela 6.12. A Q1 verifica se o participante conseguiu concluir a tarefa de autoria e, se sim, o tempo necessário para a conclusão da tarefa foi anotado na Q2. A questão Q3 é destinada a gerar correlação com os usuários que tinham experiência em linguagens baseadas em XML, separando os experientes dos iniciantes. A questão Q4

verifica se, na prática, realmente os participantes conseguiram compreender a linguagem. A hipótese é que os participantes com experiência em XML teriam melhor desempenho no teste comparados àqueles que não têm experiência.

Tabela 6.12: Perguntas do G1

Q#	Perguntas
Q1	Você conseguiu finalizar a tarefa?
Q2	Quantos minutos você levou para completar a tarefa?
Q3	Indique sua experiência anterior com linguagens baseadas em XML.
Q4	A linguagem MultiSEL foi assimilada pelos participantes?

As métricas para as questões Q1 e Q2 são as seguintes:

- Para Q1, apenas respostas afirmativas ou negativas foram consideradas, portanto, a métrica é a porcentagem de usuários que concluíram a tarefa de autoria;
- Para Q2, a métrica é o tempo médio gasto na tarefa reportado em minutos;
- Q3 foi utilizada para conhecer a experiência prévia dos participantes com XML. Ela especifica uma escala Likert [42], que vai de "Muito novato" classificado como 1 até "Muito experiente" classificado como 5. Os participantes que declararem 1 a 2 serão classificados como iniciantes, enquanto que os participantes que declararem 3 a 5 serão classificados como experientes, assim separando em dois grupos de usuários.
- Para avaliar a questão Q4 sobre entendimento da linguagem, foi realizada a correção dos códigos MultiSEL de projetos criados pelos participantes. Foi estipulado que médias abaixo de 40% seriam classificadas como insuficientes, ou seja, representariam incompreensão da linguagem, mais de 40% e menos de 70% uma certa assimilação da linguagem, já superior a 70% foi considerado que a linguagem foi assimilada pelo participantes do experimento.

6.2.1.2 G2 - Questões e Métricas

Para alcançar o objetivo G2, isto é, analisar a linguagem MultiSEL com o objetivo de avaliar sua usabilidade do ponto de vista dos usuários, foi utilizado o framework CDN (*Cognitive Dimensions of Notations*) [9], pois é um modelo conhecido e utilizado na literatura [2] para avaliar a usabilidade de linguagens. Foi utilizado um questionário de

nove perguntas mostrado na Tabela 6.13, onde cada uma aceita respostas na escala Likert [42], que vai de "Discordo fortemente" classificado como 1 a "Concordo fortemente" classificado como 5. Além das perguntas, a tabela também mostra a dimensão cognitiva correspondente analisada pela pergunta. Usamos a mediana das respostas como métrica para as perguntas do G2.

Tabela 6.13: Perguntas G2 - Questionário CDN

Q#	Perguntas	Dimensão Cognitiva	Fator
CDN-Q1	Você consegue visualizar claramente os componentes do seu aplicativo?	Visibilidade	Positivo
CDN-Q2	Você compreendeu a função dos elementos da linguagem MultiSEL?	Expressividade de Papel	Positivo
CDN-Q3	Quão difícil foi implementar o projeto com os elementos da linguagem?	Aproximação de Mapeamento	Positivo
CDN-Q4	Você acha que esta linguagem precisa de muitas palavras para se expressar?	Difusão	Negativo
CDN-Q5	Você considera que a linguagem restringe muito a ordem na qual os elementos XML são criados no seu projeto?	Comprometimento Prematuro	Negativo
CDN-Q6	Você considera que a linguagem tem muitas dependências ocultas entre seus elementos?	Dependências Ocultas	Negativo
CDN-Q7	Você acha que a sintaxe da linguagem o induz a cometer erros no uso de seus elementos?	Susceptibilidade a Erros	Negativo
CDN-Q8	Você acha que elementos da linguagem com semânticas similares são expressos por sintaxes similares?	Consistência	Positivo
CDN-Q9	Modificar um elemento do seu código XML requereu modificar outros elementos?	Viscosidade	Negativo

É importante destacar que, como mostrado na Tabela 6.13, algumas dimensões cognitivas estão relacionadas a fatores positivos da linguagem, como: Visibilidade (CDN-Q1), Expressividade de papéis (CDN-Q2), Proximidade de mapeamento (CDN-Q3) e Consistência (CDN-Q8). Enquanto outras dimensões cognitivas estão ligadas a fatores negativos da linguagem, como: Difusão (CDN-Q4), Comprometimento prematuro (CDN-Q5), Dependências escondidas (CDN-Q6), Propensão a erros (CDN-Q7) e Viscosidade (CDN-Q9). Uma boa avaliação de usabilidade de uma linguagem é aquela em que as dimensões cognitivas relacionadas a fatores positivos da linguagem são melhor avaliadas e as dimensões cognitivas relacionadas a fatores negativos da linguagem recebem valores mais baixos no experimento [9].

6.2.1.3 Grafo GQM

A Figura 6.7 resume e ilustra a estrutura do modelo GQM construído. Os experimentos realizados para coleta de dados serão apresentadas nas seções a seguir.

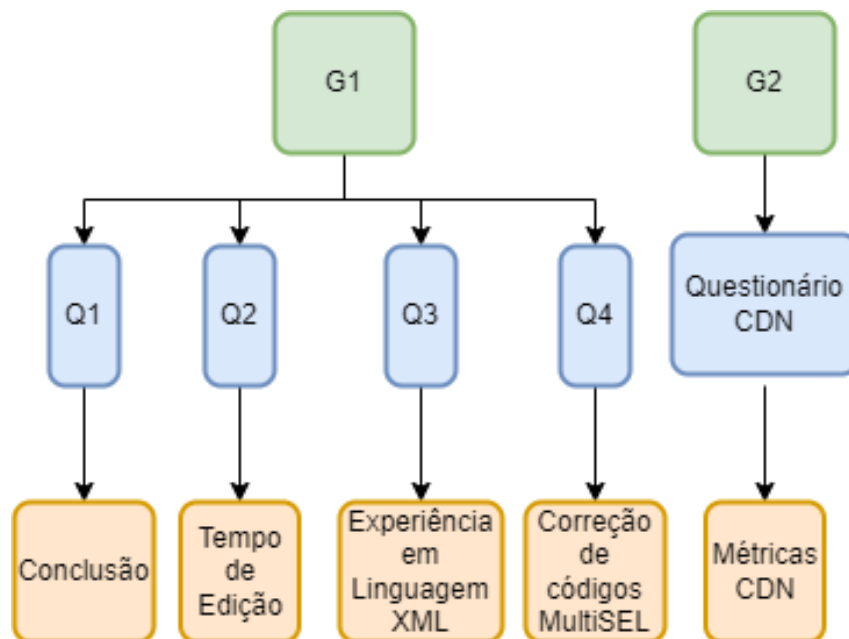


Figura 6.7: Grafo Direcionado para o modelo GQM definido.

6.2.2 Procedimentos com Participantes

Os procedimentos experimentais com usuários foram realizados durante o mês de maio de 2022 com vinte (20) participantes, um número suficiente para destacar a maioria dos problemas de usabilidade [37]. Eles eram 16 homens e 4 mulheres. As idades dos participantes variaram de 21 a 48 anos, com níveis de educação do ensino médio à pós-graduação.

O experimento completo foi realizado remotamente, usando um portal web hospedado no Github⁶, no qual os participantes tiveram acesso a:

- Um site com um tutorial para aprender a usar a linguagem MultiSEL;
- Um site com uma tarefa prática sobre MultiSEL;
- Um arquivo MultiSEL inicialmente preenchido (disponível no Apêndice B.1) e dicas de como usar IDEs com o esquema XML do MultiSEL (disponível no Apêndice B.2) foram disponibilizadas para facilitar a criação do código MultiSEL; e
- Um formulário do Google com as perguntas que respondem G1 e G2 e solicitando o *upload* da aplicação MultiSEL desenvolvida.

O objetivo das tarefas era criar uma apresentação interativa em 360° de uma mostra turística do Rio de Janeiro, tarefa equivalente ao primeiro experimento apresentado na Seção 6.1. O projeto deveria conter duas cenas com conteúdo de vídeo 360°. Na primeira cena, era necessário incluir uma imagem 2D a pelo menos 1 metro de distância da visão, exibida 1 (um) segundo depois de iniciar a cena. Essa imagem deveria ser interativa, ou seja, ao ser clicada pelo usuário, deveria iniciar a segunda cena. Na segunda cena, deveria ser incluído 1 (um) efeito sensorial de vento (*WindType*) com duração de 3 segundos, com este efeito iniciando quando a cena é carregada.

Após concluir as tarefas e fazer *upload* do código do projeto MultiSEL, os participantes foram solicitados a preencher uma pesquisa que abrangia as questões G1 e G2, bem como questões adicionais sobre idade, nível de escolaridade e experiência com autoria multimídia.

6.2.3 Resultados

Esta seção analisa e discute os dados obtidos dos experimentos e dos questionários para avaliar os objetivos definidos anteriormente.

6.2.3.1 G1 - Análise dos resultados

Para alcançar o objetivo G1, ou seja, analisar a linguagem MultiSEL com o objetivo de avaliar o entendimento sobre a linguagem e o esforço de autoria necessário tanto para

⁶https://fmflavio.github.io/experimento_multisel/home.html

usuários experientes quanto novatos, foram coletados os dados obtidos para as questões Q1, Q2, Q3 e Q4, como pode ser visto na Tabela 6.14.

Tabela 6.14: Respostas das perguntas Q1, Q2, Q3 e Q4.

Participante	Q1 (Sim/Não)	Q2 (Min.)	Q3 (Exp. 1-5)	Q4 (Sucesso)
1	Sim	56	4	67%
2	Sim	50	5	87%
3	Sim	35	4	100%
4	Sim	35	3	67%
5	Sim	45	4	80%
6	Sim	37	4	93%
7	Sim	35	5	73%
8	Sim	35	1	93%
9	Sim	30	3	60%
10	Sim	47	2	100%
11	Sim	40	2	100%
12	Sim	29	1	93%
13	Sim	24	3	87%
14	Sim	35	2	93%
15	Sim	46	1	100%
16	Sim	66	3	60%
17	Sim	20	5	100%
18	Sim	45	2	100%
19	Sim	22	4	73%
20	Sim	43	2	87%

As questões Q1, Q2, Q3 e Q4, apresentadas na Tabela 6.12, foram usadas para analisar o objetivo G1. De acordo com as respostas mostradas na Tabela 6.14, pode-se concluir que:

- Q1 - todos os participantes (100%) terminaram as tarefas propostas;
- Q2 - O tempo médio para completar as tarefas de autoria com o MultiSEL foi de 38,8 minutos com um desvio padrão de 11,45 minutos para todo o grupo de participantes;
- Q3 - 12 participantes (60%) indicaram ter experiência anterior com linguagens

baseadas em XML, com respostas iguais ou superiores a 3 pontos em 5 pontos (classificando-se como experientes), enquanto os outros 8 participantes (40%) não tinham muita experiência anterior (classificando-se como novatos); e

- Q4 - O grupo de participantes experientes conseguiu uma média de 79% de sucesso nos códigos, enquanto o grupo de novatos conseguiu uma média de 96% de sucesso nos códigos, para obter a pontuação da correção foi usado o código em forma de gabarito que pode ser visto no Apêndice B.3.

Os resultados de Q1, Q2 e Q4 serão discutidos com base na comparação entre ambos os grupos, experientes e novatos em XML.

Para avaliar a métrica tempo (Q2) para ambos os grupos, utilizamos o método estatístico de Mann-Whitney [43]. Para o teste, o valor de U foi 40,5 e o valor de p-exato foi 0,2856. Como $p\text{-exato} > 0,05$, não há diferença significativa entre o tempo gasto para ambos os grupos, considerando Q2. Na Figura 6.8, é possível observar que o tempo médio para o grupo de experientes em XML foi de 38 minutos e o desvio padrão foi de 13,49, enquanto para o grupo de novatos em XML foi de 40 minutos e o desvio padrão de 6,02.

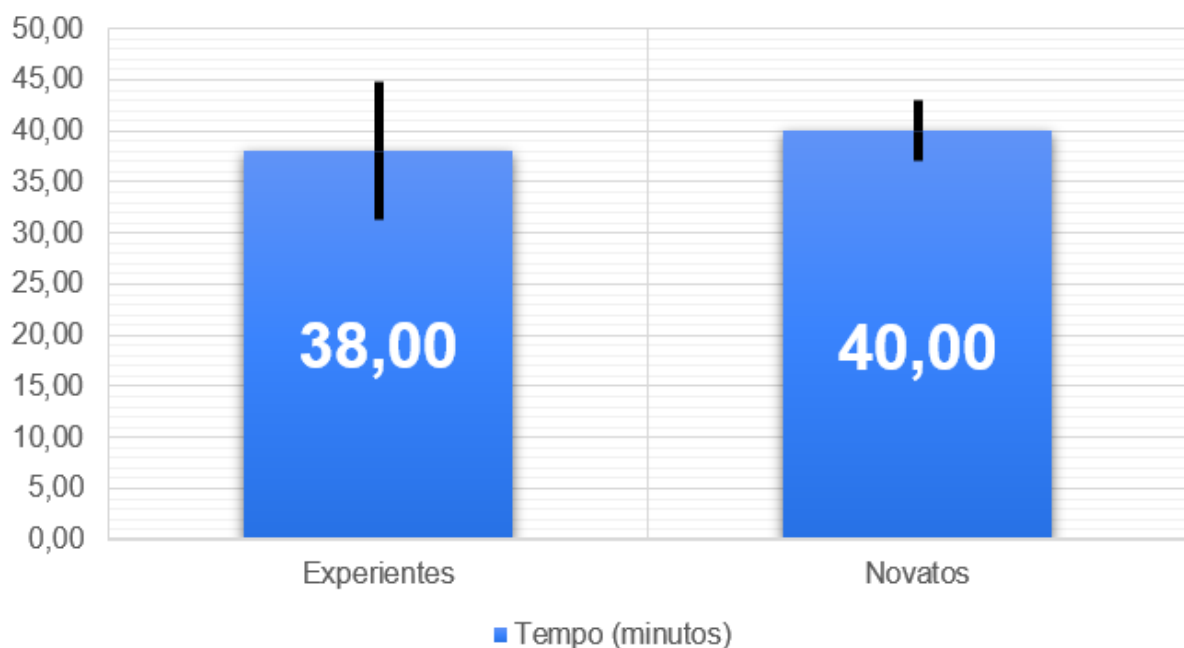


Figura 6.8: Tempo médio para Q2.

Como pode ser observado na Figura 6.9, ambos os grupos obtiveram médias de acerto (incluindo o desvio padrão) acima de 70%, o que indica que ambos os grupos tiveram uma boa assimilação da linguagem MultiSEL na prática. Na Figura 6.10, é possível visualizar a

média de acertos por *tag* XML da linguagem MultiSEL com os respectivos desvios padrão para os experientes e Na Figura 6.11 para os novatos.

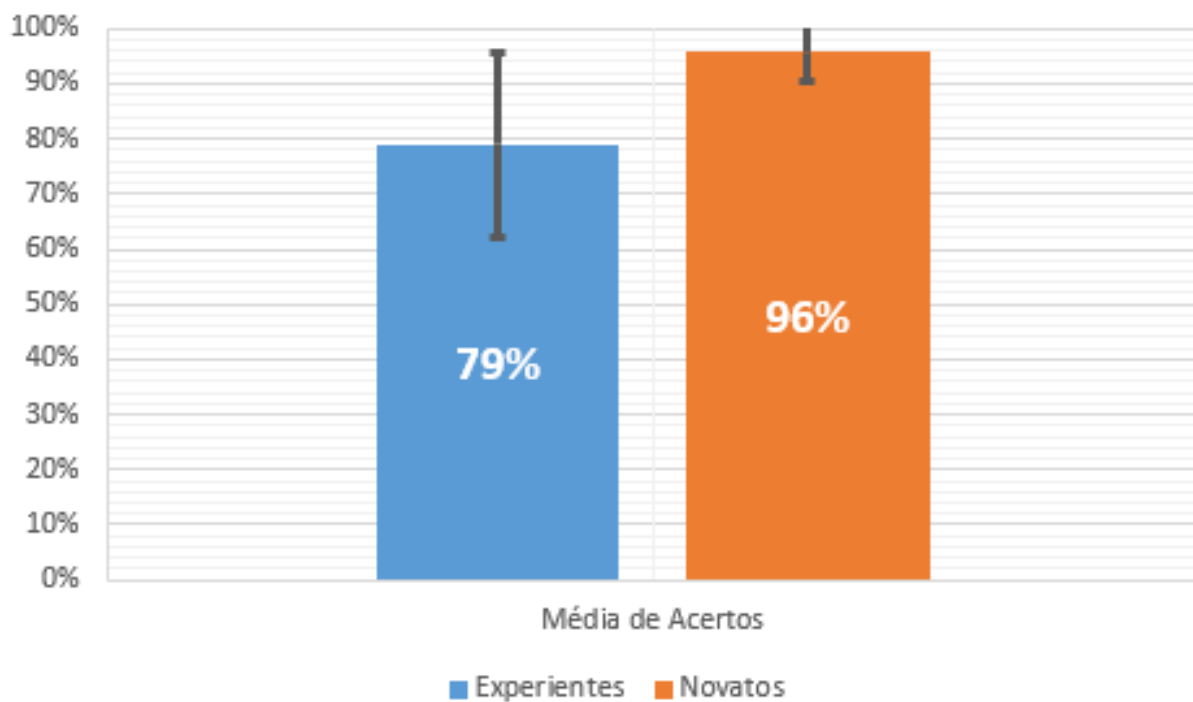


Figura 6.9: Média de sucesso em acertos para Q4.

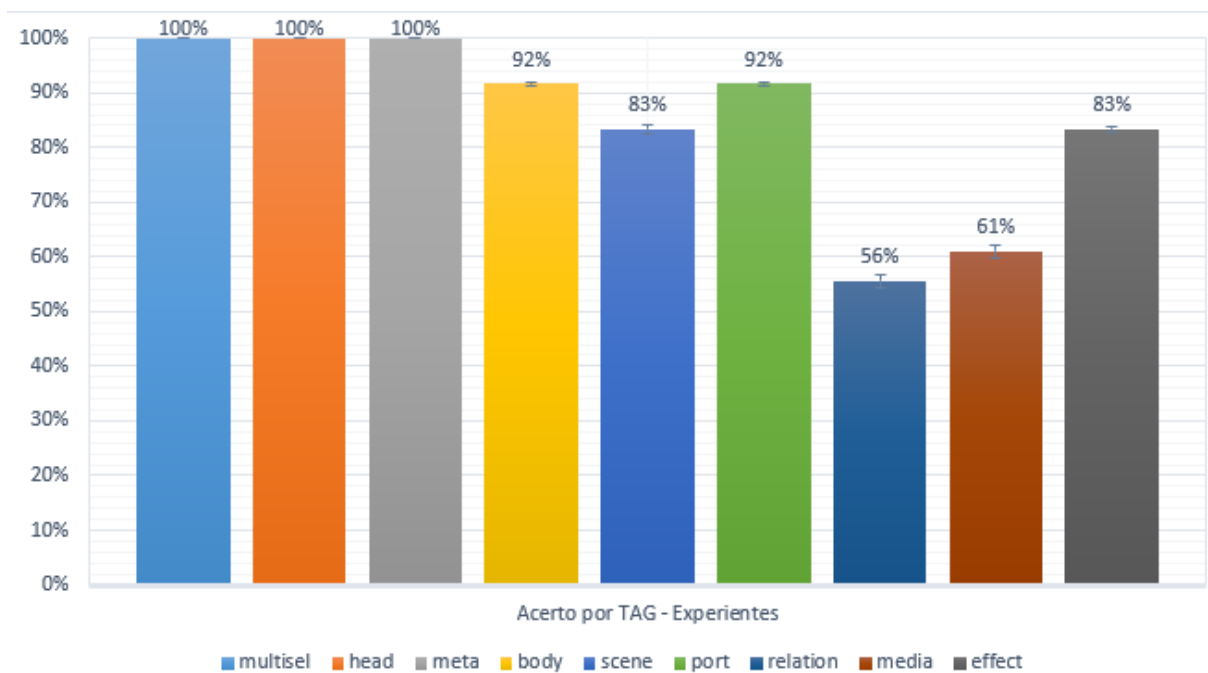


Figura 6.10: Média de sucesso em acertos para Q4 por elemento XML (experientes).

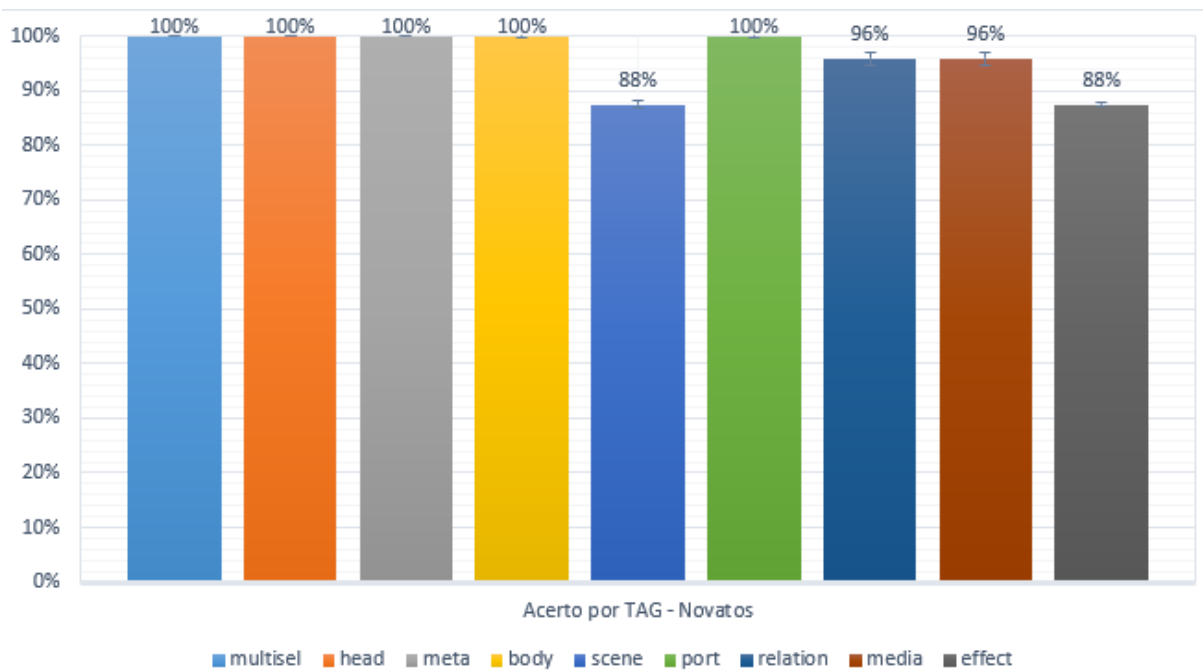


Figura 6.11: Média de sucesso em acertos para Q4 por elemento XML (Novatos).

Analisando os códigos entregues, é possível observar que participantes do grupo de experientes erraram mais na implementação das *tags relation, media e effect*, provavelmente por conta do vício de uso de outras linguagens baseadas em XML (como NCL) com conceitos similares. Com isso, é possível deduzir que os participantes do grupo novato estudaram mais o tutorial, enquanto parte dos experientes confiaram em sua experiência em outras linguagens e codificaram diretamente. Como possível reflexo disso, podemos ver nas Figuras 6.10 e 6.11 a melhor eficiência do grupo dos novatos.

6.2.3.2 G2 - Análise dos Resultados

Para avaliar o objetivo G2, ou seja, "analisar a linguagem MultiSEL com o objetivo de avaliar sua usabilidade do ponto de vista dos usuários", os dados referentes às questões do framework CDN foram coletados e analisados, considerando os dois grupos de usuários, novatos e experientes.

Para comparar os dois grupos, o método estatístico Mann-Whitney [43] foi utilizado. Os resultados obtidos para todas as questões CDN podem ser vistos na Tabela 6.15, bem como o resumo dos dados de Mediana, Desvio Padrão (*Standard Deviation* - SD), Soma de Ranks (*Rank Sum* - RS), com foco em "U" (estatística do teste) e valor "p-exato" para cada questão CDN na Tabela 6.15. Como o p-exato foi $> 0,05$ para todas as questões CDN, não há diferença significativa entre as opiniões de ambos os grupos de participantes.

Tabela 6.15: Respostas das questões do objetivo G2.

Questão	experientes em XML			Novatos em XML			Mann-Whitney	
	Mediana	SD	RS	Mediana	SD	RS	U	p-exact
CDN-Q1	4	1,142	123,0	4	0,866	87,0	45,0	0,425
CDN-Q2	4	0,862	123,5	4	0,927	86,5	44,5	0,425
CDN-Q3	3,5	0,942	111,0	4	0,927	99,0	33,0	0,135
CDN-Q4	2,5	1,440	125,0	2,5	0,866	85,0	47,0	0,468
CDN-Q5	1,5	0,986	104,0	2,5	0,500	106	26,0	0,051
CDN-Q6	3	1,320	123,5	3	0,866	86,5	45,5	0,425
CDN-Q7	2	0,759	141,5	1	0,707	68,5	32,0	0,118
CDN-Q8	3,5	1,343	125,0	3,5	1,089	85,0	47,0	0,484
CDN-Q9	1	0,759	111,0	2	0,707	99,0	33,0	0,135

Com base nos dados na Tabela 6.15 e nos gráficos nas Figuras 6.12 e 6.13, será apresentada uma breve discussão sobre os resultados das respostas às questões para os dois grupos. Primeiro, serão analisadas as respostas relacionadas às dimensões cognitivas ligadas a fatores de linguagem negativos, como Difusão/Verbosidade (CDN-Q4), Comprometimento Prematuro (CDN-Q5), Dependências Ocultas (CDN-Q6), Propensão a Erros (CDN-Q7) e Viscosidade (CDN-Q9). Lembrando que para essas questões, a resposta 1 (totalmente discordo) é considerada muito boa, 2 (discordo) como boa, 3 (neutra), 4 (concordo) como ruim e 5 (totalmente concordo) como muito ruim para avaliação do aspecto cognitivo, devido ao fator negativo dessas questões. Essa inversão do fator de resposta serve como um método de encorajar os participantes a pensarem melhor nas respostas e evitar respostas automáticas sequencialmente em valores altos.

Para Difusão/Verbosidade (CDN-Q4), uma linguagem verbosa é aquela que precisa de mais palavras ou palavras mais longas do que o necessário para expressar adequadamente a intenção do código. Pode-se analisar que tanto os experientes quanto os novatos tiveram uma mediana de 2,5 pontos, ou seja, eles apontaram esta questão como entre neutra e uma discordância. Por ser uma questão com um fator negativo, esta característica pode ser considerada aceitável. Pelas Figuras 6.12 e 6.13, percebe-se que os experientes (4 deles)

foram mais enfáticos ao discordar fortemente sobre a linguagem ser difusa/verbosa.

Para Comprometimento Prematuro (CDN-Q5), ele está relacionado a restrições sobre a ordem de fazer algo usando a notação. Podemos analisar que os experientes tiveram 1,5 ponto de mediana classificado como fortemente, ou seja, eles apontam este item como excelente, enquanto a mediana para os iniciantes foi 2,5 classificada como discordar parcialmente, ou seja, eles apontam este item como bom. Portanto, ambos os grupos avaliaram bem este aspecto da linguagem.

Para Dependências Ocultas (CDN-Q6), esta dimensão cognitiva tem uma definição muito abstrata e é frequentemente difícil de entender. Esta dimensão está relacionada ao fato de que as relações importantes entre entidades não são visíveis. Para esta dimensão cognitiva, os experientes tiveram uma mediana de 3 pontos classificados como neutro e a mediana para iniciantes também foi classificada como 3 neutro. Portanto, ambos os grupos trataram esta questão de linguagem como neutra como pode ser vista nas Figuras 6.12 e 6.13.

Para Propensão a Erros (CDN-Q7), que indica quando se acredita que uma linguagem pode levar os usuários a erros, os experientes tiveram uma mediana de 2 pontos classificados como um resultado bom, enquanto que a mediana para iniciantes foi classificada como 1, um ótimo resultado. Portanto, ambos os grupos avaliaram bem este aspecto da linguagem, não levando a erros.

A Viscosidade (CDN-Q9) está relacionada a quanto as modificações de um elemento de código requerem modificações de outros elementos. Para esta dimensão cognitiva, os experientes tiveram uma mediana de 1 ponto avaliada como um ótimo resultado, enquanto que a mediana para iniciantes foi avaliada como 2, um bom resultado. Portanto, ambos os grupos avaliaram bem este aspecto da linguagem como observado nas Figuras 6.12 e 6.13.

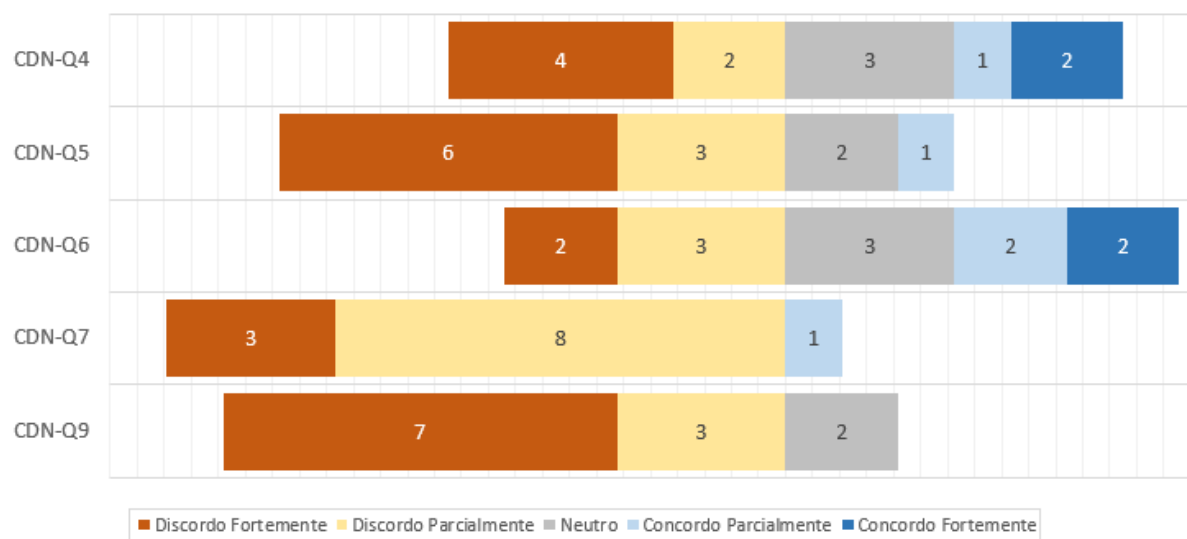


Figura 6.12: CDN - Grupo de experientes em XML, fatores negativos.

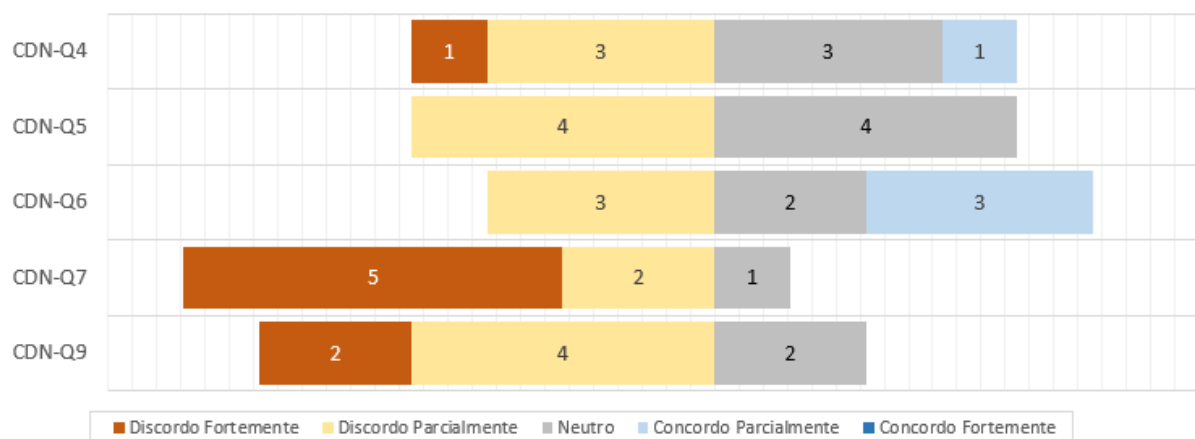


Figura 6.13: CDN - Grupo de novatos em XML, fatores negativos.

Com base nos dados da Tabela 6.15 e nas Figuras 6.14 e 6.15, será apresentada uma breve discussão sobre os resultados das respostas às questões positivas para ambos os grupos. As respostas relacionadas às dimensões cognitivas estão agora ligadas a fatores positivos da linguagem que são mais fáceis de compreender, seguidas pelas respostas à Visibilidade (CDN-Q1), Expressividade de papel (CDN-Q2), Proximidade de mapeamento (CDN-Q3) e Consistência (CDN-Q8).

Visibilidade (CDN-Q1) é a capacidade de visualizar facilmente os componentes da linguagem. Para essa dimensão cognitiva, tanto os experientes quanto os novatos têm mediana de 4 (concordam), classificando-o como um bom resultado. Como pode ser visto nas Figuras 6.14 e 6.15, com exceção de dois participantes dos grupos especialistas discordarem, todos os demais 18 participantes consideram como neutro ou positivo este

quesito. Portanto, ambos os grupos avaliaram bem esse aspecto da linguagem, pois foram respostas com um fator positivo.

Expressividade de papel (CDN-Q2) indica quando a utilidade de qualquer entidade é facilmente percebida. Para essa dimensão cognitiva, novamente tanto os experientes quanto os novatos tiveram 4 pontos para a mediana, classificando-o como um bom resultado. Portanto, ambos os grupos avaliaram bem esse aspecto da linguagem.

A Proximidade de mapeamento (CDN-Q3) indica o grau de relação entre a notação e o resultado que está sendo descrito. Para esta dimensão cognitiva, os experientes tiveram uma média de 3,5 pontos, classificando-a como um resultado positivo, enquanto a média para os iniciantes foi classificada como 4, também um resultado positivo. Portanto, ambos os grupos avaliaram bem este aspecto da linguagem como pode ser visto nas Figuras 6.14 e 6.15.

Consistência (CDN-Q8) está relacionada ao fato de que semânticas similares são expressas de maneiras sintaticamente similares. Para esta dimensão cognitiva, tanto os experientes quanto os novatos tiveram uma média de 3,5 pontos, classificando-a como um resultado positivo, porém mais próximo do neutro. Pelas Figuras 6.14 e 6.15, percebe-se que 4 experientes e 3 novatos não avaliaram bem esse aspecto da linguagem, enquanto 6 experientes e 4 novatos fizeram boas avaliações, portanto, não é possível concluir claramente o resultado sobre esse aspecto da linguagem.

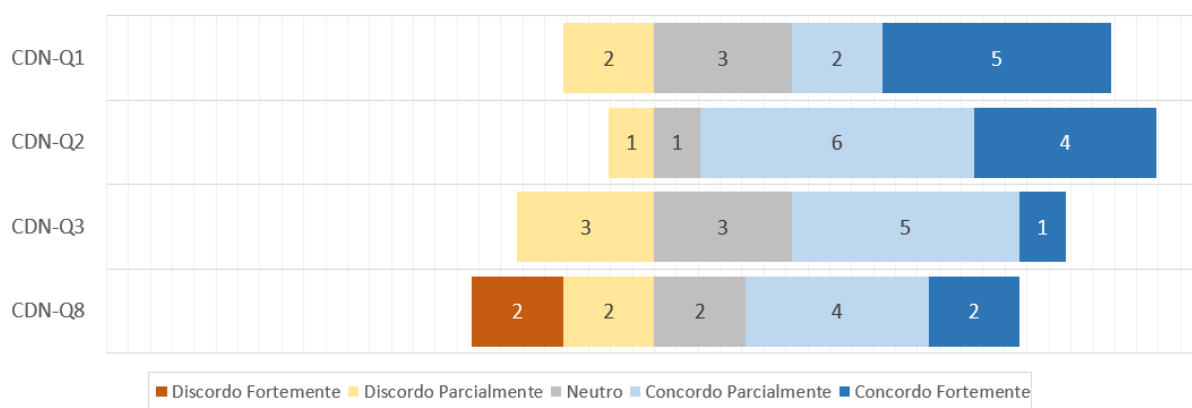


Figura 6.14: CDN - Grupo de experientes em XML, fatores positivos.

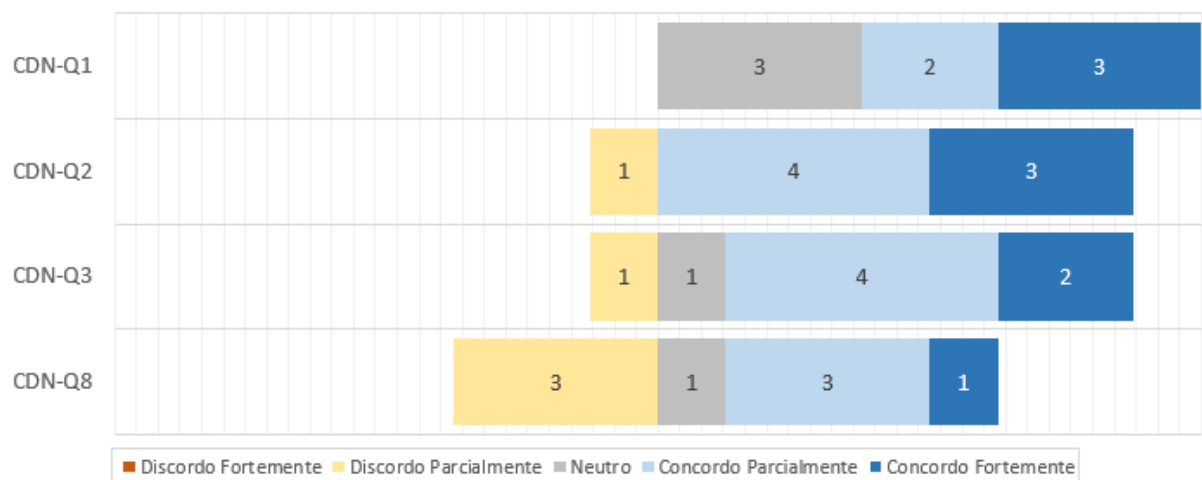


Figura 6.15: CDN - Grupo de novatos em XML, fatores positivos.

Considerando os resultados obtidos, as questões Q1, Q2 e Q3 relacionadas aos fatores positivos tiveram valor mediano de 4 para ambos os grupos, ou seja, uma boa classificação. Com relação às respostas relacionadas aos fatores negativos, para ambos os grupos, houve uma classificação boa que variou de neutra a excelente, indicando que a linguagem MultiSEL possui uma boa usabilidade pelos participantes considerando os aspectos da avaliação das dimensões cognitivas, atingindo o objetivo G2.

6.2.4 Análise Subjetiva

Neste segundo experimento apesar de ser um experimento remoto, foi observado através dos projetos enviados que alguns códigos estavam muito distintos da maioria. Estes códigos não seguiam a ordem dos demais, não respeitavam algumas regras especificadas no esquema XML disponibilizado. Em análise, é possível deduzir que alguns participantes ignoraram não só o esquema mas também exemplos semelhantes disponibilizados na página de treinamento do tutorial disponibilizado. Porém os mesmos apesar de possivelmente não terem seguido o tutorial, conseguiram finalizar a tarefa de projeto, entretanto com mais erros que os demais.

Ao analisar os dados em grupos, foi verificado que os que erraram mais nos projetos, eram justamente os que se declararam experientes, demonstrando que eles foram mais autoconfiantes e com isso decidiram codificar usando até mesmo editores próprios sem usar o esquema XML da linguagem MultiSEL.

6.2.5 Discussão

O número reduzido de usuários no experimento de autoria com MultiSEL, 20 no total, pode ser uma limitação da pesquisa realizada. Outra limitação do experimento foi o fato de os participantes não poderem testar o código MultiSEL gerado, pois realizaram o teste remotamente e não tinham disponível um HMD Oculus com AMUSEVR instalado para executar a aplicação criada.

No entanto, o trabalho indicou resultados muito positivos, permitindo chegar às seguintes conclusões:

- Considerando o propósito do G1, conclui-se que a linguagem MultiSEL pode ser utilizada tanto por iniciantes quanto por usuários experientes em XML, pois ambos os grupos conseguiram criar o projeto em tempo aproximado após passarem por treinamento semelhante. Tanto experientes quanto novatos tiveram um bom entendimento da linguagem e todos conseguiram completar as tarefas propostas em tempo hábil, em torno de 39 minutos; e
- Considerando a finalidade do G2, o teste de Mann-Whitney mostrou que não há diferença significativa nos resultados dos grupos experientes e novatos (considerando U e p-exato $> 0,05$). Além disso, a linguagem MultiSEL foi no geral bem avaliada com relação à sua usabilidade para a criação de projetos multissensoriais, considerando as dimensões cognitivas do framework CDN.

6.3 Experimento com AMUSEVR Oculus com Renderização de Efeitos Sensoriais

O terceiro experimento para o AMUSEVR Oculus realiza a avaliação da proposta desta tese sobre autoria imersiva por meio de experimentos onde os usuários utilizaram a ferramenta AMUSEVR como um protótipo de estudo de caso. Também foram utilizadas as técnicas e métodos de avaliação do *Goal Question Metric* (GQM) [6], o Questionário de Experiência do Usuário (UEQ)⁷ [40, 55] e o questionário Escala de Usabilidade do Sistema - SUS⁸ [10].

Com o objetivo de descrever o processo, a metodologia usada para estruturar e elaborar esta avaliação é discutida na Seção 6.3.1. O experimento com usuários é discutido

⁷<https://www.ueq-online.org/>

⁸<https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html>

na Seção 6.3.2 e uma avaliação quantitativa é descrita na Seção 6.3.3. Na Seção 6.3.4, são apresentados os resultados obtidos e finalmente, na Seção 6.3.6, é feita uma análise e discussão das conclusões finais e das limitações atuais.

6.3.1 Metodologia

Como metodologia, foi usada a abordagem *Goal Question Metric* (GQM) [6] para estruturar esta avaliação, como foi realizado nas Seções 6.1.1 e 6.2.1 relativas aos primeiro e segundo experimento.

A Tabela 6.16 mostra os objetivos definidos para a avaliação do AMUSEVR. Para cada objetivo G1-G4, a coluna "Descrição" apresenta o objeto sendo analisado.

Tabela 6.16: Objetivos para os experimentos com o AMUSEVR.

Meta	Descrição
G1	Analisar a ferramenta AMUSEVR com o objetivo de avaliar sua usabilidade sob a perspectiva do usuário.
G2	Analisar a ferramenta AMUSEVR com o objetivo de avaliar sua qualidade de experiência sob a perspectiva do usuário.
G3	Analisar a ferramenta AMUSEVR com o objetivo de avaliar o esforço de autoria exigido tanto para usuários experientes quanto iniciantes.
G4	Analisar a ferramenta AMUSEVR com o objetivo de avaliar o esforço de autoria em comparação com um método alternativo.
G5	Avaliar a abordagem de autoria imersiva do ponto de vista do usuário.

6.3.1.1 G1 - Questões e Métricas

Para atingir o objetivo G1, ou seja, analisar a ferramenta com o objetivo de avaliar sua usabilidade do ponto de vista do usuário, foi utilizado o questionário *System Usability Scale* (SUS) [10], semelhante ao visto na Seção 6.1.1.1 do primeiro experimento. O SUS especifica um questionário de dez itens mostrados na Tabela 6.3 presente também no primeiro experimento, onde cada um aceita respostas na escala Likert, variando de "Discordo fortemente" classificado como 1 até "Concordo fortemente" classificado como 5 [42]. Usamos a pontuação do SUS como métrica para as perguntas de G1.

6.3.1.2 G2 - Questões e Métricas

Para atingir o objetivo G2, isto é, analisar a ferramenta com o objetivo de avaliar sua qualidade de experiência do ponto de vista do usuário, foi utilizado o Questionário de Experiência do Usuário (UEQ-S) [40, 55], semelhante ao visto na Seção 6.1.1.2 do primeiro experimento. Como pode ser visto na Tabela 6.4 apresentado no primeiro experimento, o questionário contém oito itens (UEQ-Q1 a UEQ-Q8) com feedback negativo ou positivo em resposta à frase "A experiência com AMUSEVR foi". As respostas dos participantes foram numa escala de (-3) (feedback negativo) a (+3) (feedback positivo). Foram usados os scores pragmáticos e hedonistas do UEQ como métricas para as perguntas de G2.

6.3.1.3 G3 - Questões e Métricas

Para atender ao objetivo G3, ou seja, analisar a ferramenta com o objetivo de avaliar o esforço de autoria necessário tanto para usuários experientes quanto iniciantes, foi utilizado um conjunto de perguntas específicas, como pode ser visto na Tabela 6.17, semelhante ao visto na Seção 6.1.1.3 do primeiro experimento. A Q1 verificou se o participante conseguiu completar uma tarefa e, caso positivo, o tempo necessário para a conclusão da tarefa foi anotado na Q4.

Tabela 6.17: Perguntas personalizadas para G3.

Q#	Perguntas
Q1	Você conseguiu finalizar a tarefa?
Q2	Indique seu nível de experiência na utilização de realidade virtual ou ambientes imersivos.
Q3	Indique seu nível de experiência na utilização de ferramentas de autoria multimídia.
Q4	Quantos minutos você levou para completar a tarefa?

As questões Q2 e Q3 foram projetadas para separar os participantes em dois grupos, aqueles que já se consideravam com alguma experiência em ambientes imersivos e os que se consideravam com alguma experiência com ferramentas de autoria. Cada questão aceita respostas na escala Likert, variando de "Discordo fortemente" classificado como 1 até "Concordo fortemente" classificado como 5 [42]. A hipótese é que os participantes com experiência em ambientes imersivos e com ferramentas de autoria multimídia teriam um desempenho melhor do que aqueles sem experiência.

O tempo médio de edição para concluir a tarefa solicitado na questão Q4 foi usado como métrica para comparar o esforço de autoria entre usuários experientes e novatos/iniciantes.

6.3.1.4 G4 - Questões e Métricas

Para atingir o objetivo G4, ou seja, analisar a ferramenta AMUSEVR com o objetivo de avaliar o esforço de autoria em comparação a um método alternativo, foi utilizada a linguagem MultiSEL como o método alternativo de autoria. MultiSEL foi utilizada, pois não foi encontrado um método alternativo que permitisse criar aplicações multissensoriais com conteúdo 360 graus similar ao que AMUSEVR fornece.

As perguntas relativas ao objetivo G4 estão apresentadas na Tabela 6.18. A questão Q5 investiga se os participantes já utilizaram a Linguagem MultiSEL anteriormente. Em caso positivo, o participante deve responder as perguntas Q6 e Q7 sobre o formato de autoria de preferência e o formato mais produtivo, respectivamente, conforme sua própria visão. Se o participante nunca utilizou MultiSEL, as perguntas Q6 e Q7 devem ser ignoradas. A métrica para as questões Q6 e Q7 é a porcentagem de participantes que preferem um dos formatos indicados.

Tabela 6.18: Perguntas personalizadas para G4.

Q#	Pergunta
Q5	Já utilizou a Linguagem MultiSEL?
Q6	Na sua opinião, qual formato de autoria de sua preferência: AMUSEVR, MultiSEL, Ambos, Nenhum?
Q7	Na sua opinião, qual formato de autoria acredita ser mais produtivo para criar projetos multissensoriais em 360°: AMUSEVR, MultiSEL, Ambos, Nenhum?
Q8	Qual é o número necessário de cliques na ferramenta AMUSEVR em comparação ao número de linhas de um arquivo correspondente em MultiSEL com uma especificação semelhante?

Além das questões Q5-Q7, uma outra questão Q8 é definida, que é "*Qual é o número necessário de cliques na ferramenta AMUSEVR em comparação ao número de linhas de um arquivo correspondente em MultiSEL com uma especificação semelhante?*", conforme apresentado na Tabela 6.18. Para responder a essa pergunta, um teste comparativo específico para o AMUSEVR foi feito contra o método de autoria declarativo baseado

na linguagem MultiSEL, considerando diferentes tarefas, como será visto na Seção 6.3.3. Será utilizada uma métrica quantitativa baseada no número de cliques entre o AMUSEVR e o número de linhas de código MultiSEL para cada tarefa.

Reconhecemos que essa comparação entre o AMUSEVR e o MultiSEL pode não ser tão justa, já que as abordagens de autoria declarativa (textual) e gráfica são bem diferentes, mas como não há outro método ou ferramenta alternativa disponível para comparar o esforço de autoria do AMUSEVR, MultiSEL acabou sendo usada para avaliar o objetivo G4.

6.3.1.5 G5 - Questões e Métricas

Para atingir o objetivo G5, ou seja, avaliar a abordagem de autoria imersiva do ponto de vista do usuário, vamos usar as questões Q9 a Q16 de um formulário próprio presente na Tabela 6.19, onde de Q9 até Q12 são questões referentes a avaliar a abordagem de autoria imersiva, enquanto de Q13 até Q16 são questões referentes a avaliar a sensação de uso da ferramenta. Cada questão aceita uma resposta na escala Likert [42], variando de "Discordo fortemente" classificado como 1 até "Concordo fortemente" classificado como 5. O método da análise estatística será o teste não paramétrico Mann-Whitney [43] para comparar grupos com perfis distintos de usuários.

Tabela 6.19: Perguntas personalizadas para G5.

Q#	Perguntas
Q9	Eu me senti imerso na ferramenta de autoria.
Q10	A organização do projeto multissensorial em cenas imersivas facilita a sua autoria.
Q11	Foi fácil posicionar espacialmente as mídias e efeitos sensoriais no ambiente imersivo do AMUSEVR.
Q12	Eu acredito que NÃO necessitaria da realidade virtual para realizar as tarefas autorais.
Q13	Eu me senti incomodado visualmente ao usar o sistema.
Q14	Eu senti enjoo ao usar o sistema.
Q15	Eu senti tontura ao usar o sistema.
Q16	Eu senti desconforto físico ao usar o sistema.

6.3.1.6 Grafo GQM

A Figura 6.16 resume e ilustra a estrutura do modelo GQM construído. As experimentações realizadas para a coleta de dados serão apresentadas nas seções subsequentes. As experimentações foram divididas em duas partes, uma experimentação com participantes para atender aos objetivos G1, G2, G3 e G5, e a segunda realizada em outro momento para analisar dados quantitativos para atender ao objetivo G4.

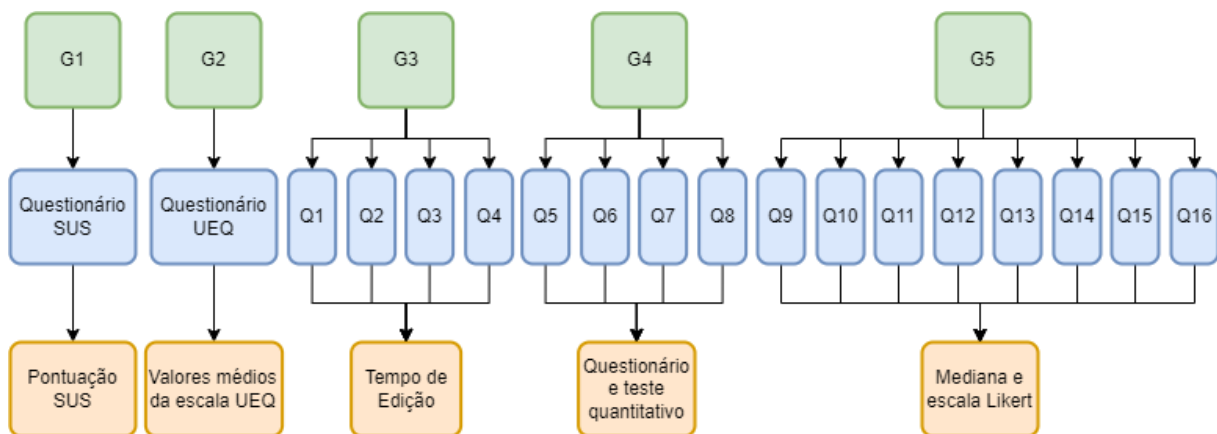


Figura 6.16: Diagrama direcionado com Objetivos, Questões e Métricas (GQM) do experimento com AMUSEVR Oculus.

6.3.2 Procedimentos com Participantes

Os procedimentos experimentais com usuários foram realizados considerando os objetivos definidos na Seção 6.3.1 para investigar a pergunta principal de pesquisa desta tese: *"Qual é a experiência do autor na criação de aplicações multissensoriais interativas em 360 graus usando um ambiente de autoria imersivo?"*. Os testes foram realizados presencialmente na UFF nas cidades de Niterói-RJ e no IFAC em Rio Branco-AC durante o ano de 2022, com quarenta e dois (42) participantes. Eles eram 32 homens e 10 mulheres com idades entre 18 e 48 anos, com níveis de educação variando do ensino médio à pós-graduação. Dois participantes eram canhotos.

Durante o experimento, os participantes tiveram a opção de escolher entre sentar ou ficar em pé durante a experiência de autoria imersiva, assim como foi realizado no primeiro experimento apresentado na Seção 6.1. Um vídeo sobre o terceiro experimento está disponível⁹. A Figura 6.17 exibe como foi realizado o terceiro experimento. Nela é possível observar um ventilador responsável pelo efeito sensorial de vento, uma lâmpada infravermelha voltada a participante responsável pelo efeito de calor, os controles

⁹<https://youtu.be/vRYaKAtB0F8>

responsáveis pelos efeitos hápticos de vibração, a cadeira giratória no qual o participante trabalhava, um auxiliar que ajudava com experimento e um nootebook que era utilizado para acompanhar a visão do participante.



Figura 6.17: Terceiro experimento com efeitos sensoriais.

Assim como no primeiro experimento, todos os participantes tiveram tempo para se familiarizar com o ambiente e as ferramentas usadas (cinco minutos em média), após o qual, cada um começou as tarefas e teve seus resultados registrados.

As tarefas foram definidas de acordo com alguns requisitos:

- o projeto deve ser curto, com no máximo duas cenas, as cenas 1 e 2, já que os participantes podem não ter experiência em ferramentas de autoria ou em realidade virtual. Além disso, é desejado realizar possíveis comparações com outras ferramentas ou mesmo um projeto MultiSEL;
- para cobrir alguns destaques da ferramenta, deve-se incluir uma mídia de fundo, algumas mídias 2D clássicas e um efeito sensorial, além de um exemplo de relação de sincronização e interação entre as cenas;
- os participantes devem ativar o modo de visualização para avaliar se estão satisfeitos com seu projeto e testar a interatividade programada e a renderização dos efeitos sensoriais.

As tarefas também visaram criar uma apresentação interativa de 360 graus de um showcase turístico do Rio de Janeiro, semelhantes às apresentadas no primeiro e no segundo experimentos, descritos nas Seções 6.1 e 6.2. Para guiar os usuários que participaram do experimento, assim como no primeiro experimento, foi fornecida uma lista de tarefas no menu virtual do controle descrevendo as tarefas a serem realizadas.

Após completar as tarefas, os participantes foram solicitados a responder um questionário com todas as perguntas planejadas no terceiro experimento. Neste experimento, ao ambiente de execução (*Preview mode*) contemplou a renderização de efeitos sensoriais de vento, calor e vibração com atuadores físicos, diferente do primeiro experimento realizado.

6.3.3 Experimento Comparativo para Esforço de Autoria - AMUSEVR x MultiSEL

Para comparar quantitativamente o esforço de autoria entre o AMUSEVR e o MultiSEL, foi utilizado o número de cliques necessários para completar tarefas no AMUSEVR em comparação com o número de linhas em um documento MultiSEL com a mesma especificação. Para realizar esta comparação, foram criados diferentes aplicativos com o AMUSEVR. Depois de editar cada projeto no AMUSEVR, exportamos as aplicações no formato MultiSEL para estabelecer um padrão comparativo para cada nível de especificação diferente.

Para a primeira tarefa comparativa, foi configurado o vídeo de fundo em 360 graus demo.mp4. Completar esta tarefa no AMUSEVR envolveu: um clique com o ponteiro do controle da mão direita no botão de menu à esquerda "*New Media*", conforme mostrado na Figura 5.1, seguido de um clique no botão "*Video 360*" no menu recém-aberto. Foi feito outro clique no menu da lista estendida (*Dropdown*) de mídia contida dentro da janela de menu flutuante de edição de mídia (veja no item 2 da Figura 5.6), que lista todos os arquivos de mídia disponíveis para serem usados como um vídeo de fundo 360. Para exportar o projeto, um clique no botão de menu à esquerda "*Back*", pelo qual o usuário é levado de volta ao menu principal. Outro clique em "*Exit*" no qual fecha o projeto, e para completar a tarefa, um clique final em "*Yes*" para confirmar a exportação. Em resumo, foram necessários seis cliques para completar a tarefa. O arquivo exportado resultante da tarefa AMUSEVR produz a Listagem 6.1, um documento MultiSEL com 22 linhas. Neste exemplo, foi usado um vídeo de 360 graus no fundo, mas para cada tipo de mídia, há uma avaliação do tipo de propriedades, resultando em uma variação no número de linhas.

```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <multisel id="Project-1" title="MultiSEL for AMUSEVR" xmlns="DefaultProfile">
3   <head>
4     <meta name="author" value="anonymous" />
5     <meta name="year" value="2022" />
6   </head>
7   <body primaryComponent="scene 1">
8     <scene id="scene 1" primaryComponent="video360_s1">
9       <media id="video360_s1" src="demo.mp4" type="video">
10        <property name="duration" value="259" />
11        <property name="background" value="true" />
12        <property name="loop" value="false" />
13        <property name="volume" value="0.72" />
14        <property name="mute" value="false" />
15      </media>
16      <relation id="Relation-1" type="finishes">
17        <primary component="scene 1" />
18        <secondary component="video360_s1" />
19      </relation>
20    </scene>
21  </body>
22 </multisel>

```

Listing 6.1: Exemplo MultiSEL equivalente a tarefa T1 feita no AMUSEVR.

Foram estabelecidas seis tarefas idênticas, mas progressivamente mais difíceis, para comparar o esforço de autoria do AMUSEVR e do MultiSEL, como mostrado na Tabela 6.20 (T1 a T6).

Tabela 6.20: Tarefas comparativas - AMUSEVR x MultiSEL.

T#	Tarefas
T1	Criar cena 1 contendo um vídeo em 360 graus como pano de fundo (como na Listagem 6.1).
T2	Criar cena 1 contendo um áudio 3D.
T3	Criar cena 1 contendo um vídeo em 360 graus como pano de fundo e um áudio 3D.
T4	Criar cena 1 contendo um vídeo em 360 graus como pano de fundo e um botão interativo para a cena 2. Cena 2 com uma imagem em 360 graus.
T5	Criar cena 1 contendo um vídeo em 360 graus como pano de fundo e um botão interativo para a cena 2. Cena 2 com uma imagem em 360 graus e um vídeo 2D.
T6	Criar cena 1 contendo um vídeo em 360 graus como pano de fundo e um botão interativo para a Cena 2. Cena 2 com uma imagem em 360 graus e um vídeo 2D interativo para a Cena 3. Cena 3 com um efeito sensorial de calor.

6.3.4 Resultados

Nesta subseção, são analisados e discutidos os dados obtidos nos experimentos e nos questionários, que respondem aos objetivos definidos anteriormente.

6.3.4.1 G1 - Análise dos Resultados

Considerando o objetivo G1, para analisar a usabilidade do AMUSEVR, o questionário SUS obteve uma média de 78,51 pontos. Essa pontuação foi obtida a partir dos dados gerados pelas questões G1 mostradas na Figura 6.18. O resultado da pontuação SUS é promissor para a usabilidade do sistema, já que é maior que 68 pontos, que representam a média de base aceitável como citado em [10].

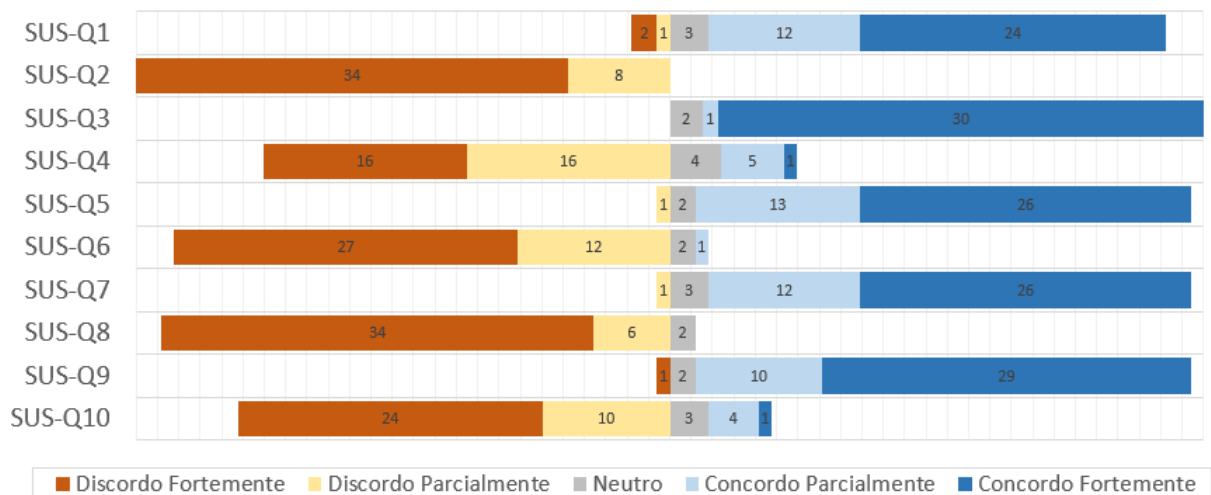


Figura 6.18: Número de respostas às afirmações do questionário do SUS (1 - discordo totalmente; 5 - concordo totalmente).

Além da escala da relativa à pontuação SUS como aceitável, é possível classificar a usabilidade do AMUSEVR Oculus, utilizando-se uma escala de adjetivos [3], como pode ser visto na Figura 6.19. Essa escala é dividida em seis categorias associadas à pontuação SUS. Baseado na escala adjetiva, podemos classificar o AMUSEVR Oculus com o adjetivo "Excelente".

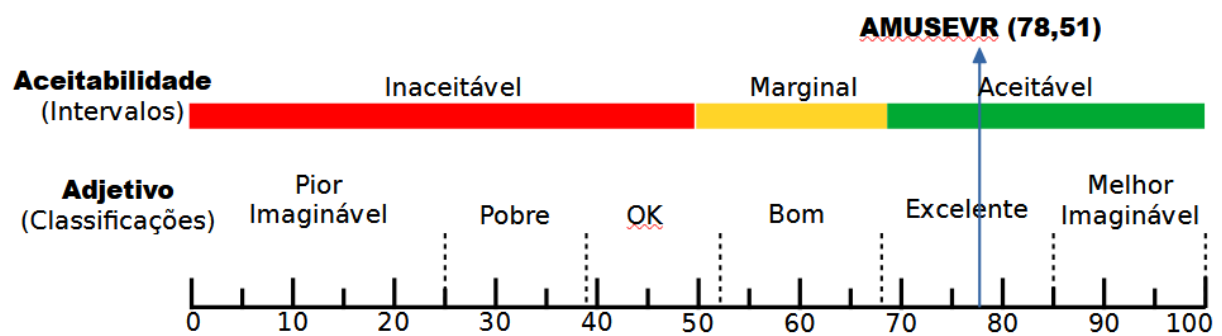


Figura 6.19: Escalas de Aceitabilidade e Adjetivos associadas à pontuação SUS.

Comparando o resultado SUS do primeiro experimento com AMUSEVR Vive com 10 participantes e o resultado do terceiro experimento com AMUSEVR Oculus com 42 participantes como pode ser visto na Figura 6.20, temos uma média de 80,38 pontos considerando um desvio padrão de 2,64 pontos. Se considerarmos essa média como uma pontuação SUS, ainda sim seria classificado como “Aceitável” na escala de aceitabilidade e “Excelente” na escala de Adjetivos, como poderia ser classificado no gráfico da última Figura 6.19.

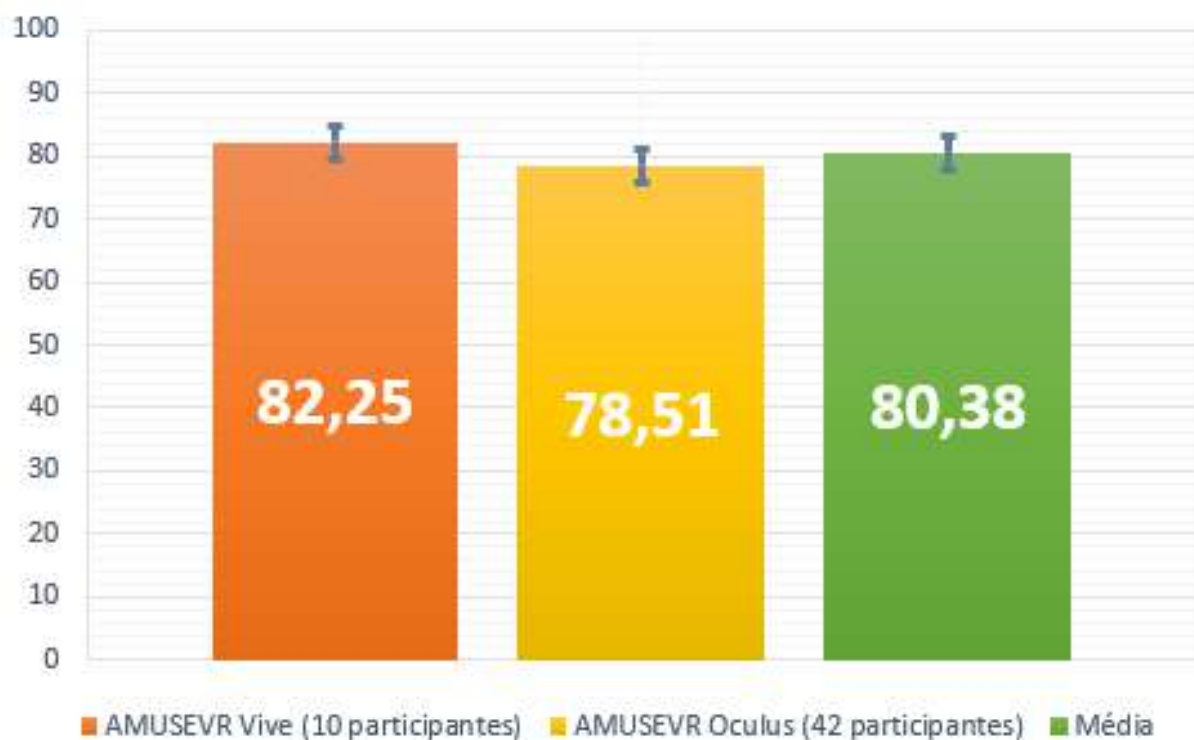


Figura 6.20: Comparativo de pontuação SUS dos experimentos.

6.3.4.2 G2 - Análise dos Resultados

Para alcançar o objetivo G2, o UEQ-S foi realizado de maneira complementar para avaliar a experiência do usuário. Foram utilizadas as médias das respostas às perguntas e a metodologia já estabelecida para o UEQ, com sua ferramenta de análise de dados¹⁰. De acordo com a interpretação dos resultados do UEQ [55], AMUSEVR Oculus obteve uma média aproximada de 2,0 com "Qualidade Pragmática" para alguns pontos avaliados e média aproximada de 2,5 com "Qualidade Hedônica" para os demais pontos, conforme pode ser visto na Tabela 6.21. A Figura 6.21 mostra o *benchmark* alcançado, o qual classifica a experiência do usuário como excelente de acordo com os requisitos de avaliação dos criadores do UEQ [40, 55].

Tabela 6.21: Resultado UEQ-S obtido.

Qualidade UEQ	Negativo	Positivo	Pont.(-3 à +3)	Variância
Pragmática	Obstrutivo	Autoexplicativo	+1.9	1,0
	Complicado	Fácil	+2,0	1,0
	Ineficiente	Eficiente	+2,2	1,2
	Confuso	Claro	+1,9	1,1
Hedônica	Entediante	Excitante	+2,5	0,9
	Desinteressante	Interessante	+2,8	0,5
	Convencional	Original	+2,5	0,8
	Comum	Vanguardista	+2,0	1,1

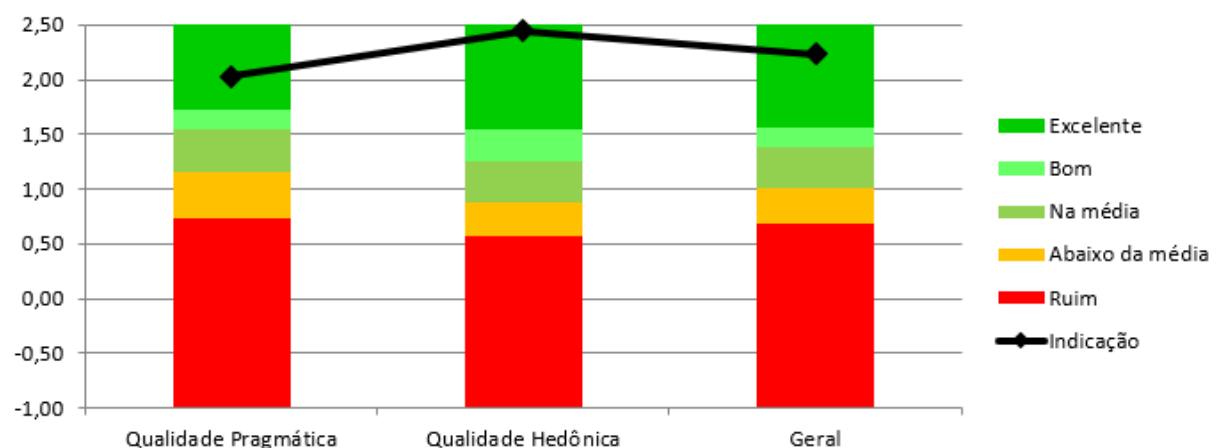


Figura 6.21: *Benchmark* UEQ para AMUSEVR Oculus.

Comparando o resultado UEQ-S em *Benchmark* do primeiro experimento com AMU-

¹⁰https://www.ueq-online.org/Material/Short_UEQ_Data_Analysis_Tool.xlsx

SEVR Vive com 10 participantes e o resultado do terceiro experimento com AMUSEVR Oculus com 42 participantes como pode ser visto na Figura 6.22, quando comparados os resultados dos experimentos, podemos observar que foram indicados o desempenho da versão Vive, Oculus e uma média, onde todos os resultados apresentam classificação “Excelente” para os quesitos “Qualidade Pragmática”, “Qualidade Hedônica” e no “Geral”.

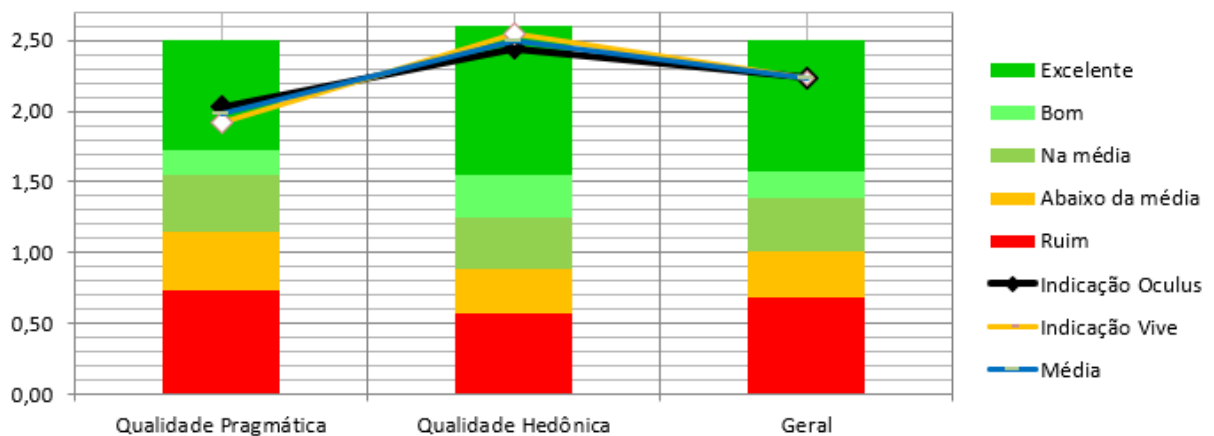


Figura 6.22: *Benchmark* UEQ comparativo entre AMUSEVR Oculus e Vive.

6.3.4.3 G3 - Análise dos resultados

As questões Q1, Q2, Q3 e Q4 foram utilizadas para analisar o objetivo G3. Para Q1, tivemos 100% de resposta Sim, ou seja, todos concluíram a atividade proposta.

A Tabela 6.22 mostra o nível de experiência do grupo de participantes, através das respostas de Q2 e Q3, em realidade virtual e em ferramentas de autoria multimídia. Para obter os grupos de novatos ou experiente, somamos as pontuações atribuídas para cada uma das respostas dadas por cada participante e dividimos por dois, obtendo uma média simples de experiência para cada participante. Com isto, consideramos os participantes com médias < 3 como novatos e o restante no grupo experiente. Exemplo, participante 22 considerou para Q2 Discordo Fortemente, pontuando 1, soma-se isso a Q3 Concordo Parcialmente, pontuando 4, como média, temos 2,5. Desta forma, classificamos o participante 22 no grupo de Novatos. Utilizando esse critério, obtivemos um grupo de 30 novatos e 12 experientes em realidade virtual ou ferramentas de autoria multimídia.

Tabela 6.22: Respostas sobre experiência dos usuários.

Q#	Pergunta	1	2	3	4	5
Q2	Indique seu nível de experiência na utilização de realidade virtual ou ambientes imersivos	23	5	7	6	1
Q3	Indique seu nível de experiência na utilização de ferramentas de autoria multimídia.	16	5	11	7	3

A Q4 é referente ao tempo médio levado para completar as tarefas com o AMUSEVR, que foi de 5,23 minutos, com desvio padrão de 1,47 para o grupo todo. O tempo levado para grupos distintos de usuários experientes e novatos é apresentado na Tabela 6.23 e Figura 6.23.

Tabela 6.23: Tempo de conclusão do experimento por usuários experientes e novatos.

Participantes	Tempo (minutos)	
	Média	Desvio Padrão
experientes	4,75	1,075
Novatos	5,43	1,590

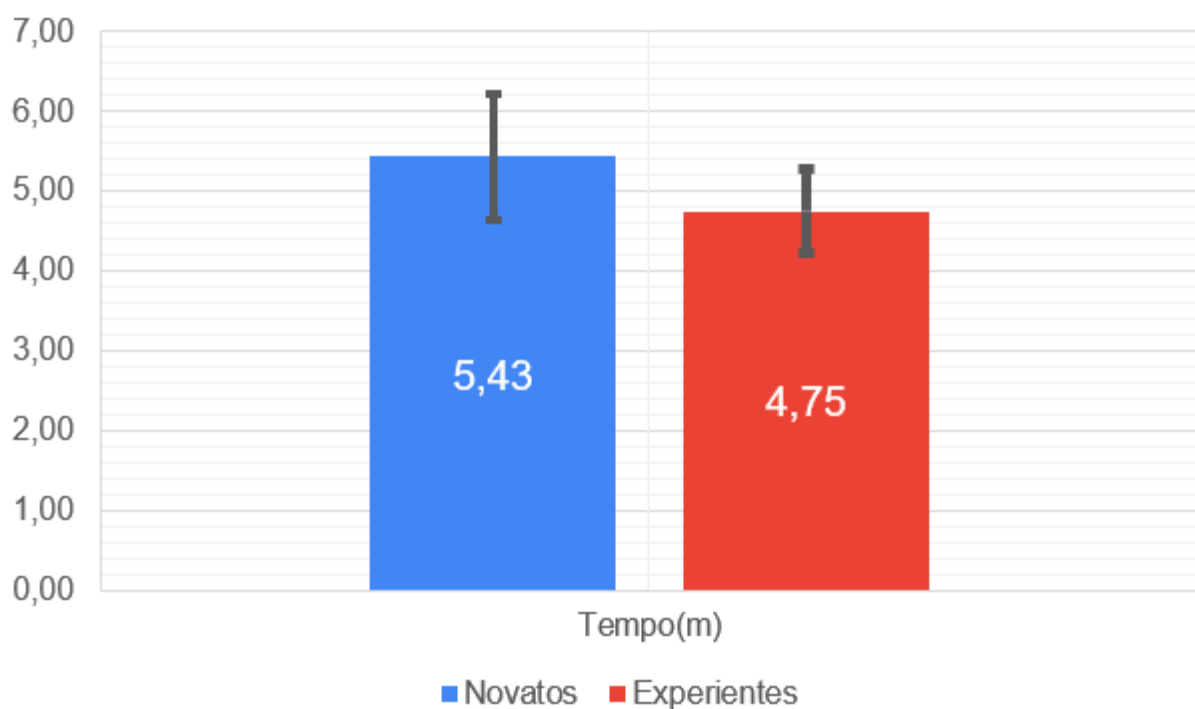


Figura 6.23: Tempo médio para os grupos em Q4.

6.3.4.4 G4 - Análise dos Resultados

Considerando o objetivo G4, a questão Q5 investigou se os participantes já tinham utilizado a Linguagem MultiSEL, obtendo-se o seguinte resultado: 29 participantes (69%) responderam *Não* e 13 participantes (31%) responderam *Sim*, ou seja, haviam participado do segundo experimento. Do grupo de participantes que responderam já ter utilizado a Linguagem MultiSEL, realizamos as duas perguntas descritas na Tabela 6.24.

Tabela 6.24: Comparativo da preferência dos usuários sobre métodos de autoria.

Pergunta	AMUSEVR	MultiSEL	Ambos	Nenhum
Q6 - Na sua opinião, qual formato de autoria de sua preferência?	65%	0%	20%	15%
Q7 - Na sua opinião, qual formato de autoria acredita ser mais produtivo para criar projetos multissensoriais em 360°?	71,5%	0%	19%	9,5%

A partir da Tabela 6.24, percebe-se que dos 13 participantes, que já haviam utilizado MultiSEL, 65% preferem AMUSEVR como ferramenta de autoria, enquanto que 71,5% indicam que AMUSEVR é uma ferramenta de autoria mais produtiva para criar projetos multissensoriais, quando comparada à MultiSEL.

Em relação à questão Q8 do objetivo G4, o processo de teste foi descrito na Seção 6.3.3. A Tabela 6.25 mostra a comparação entre o número de cliques e linhas necessários para diferentes tarefas T1-T6 realizadas. Com o aumento de objetos de mídia e interação relacionados a cada tarefa, o código MultiSEL aumentou consideravelmente, resultando em um desenvolvimento tedioso e cansativo através do MultiSEL. É importante, no entanto, observar que se as tarefas fossem realizadas por um único usuário, o custo de tempo seria muito maior, sem mencionar a necessidade de habilidade técnica e capacidade de prever o resultado de seu projeto quando falta o feedback visual que pode ser encontrado no AMUSEVR.

Tabela 6.25: Comparação do número de cliques no AMUSEVR com o número de linhas no arquivo MultiSEL.

T#	Tarefas	Cliques	Linhas
T1	Cena 1 contendo um vídeo de fundo de 360 graus (como na Listagem 6.1).	6	22
T2	Cena 1 contendo um áudio 3D.	6	24
T3	Cena 1 contendo um vídeo de fundo de 360 graus e um áudio 3D.	8	39
T4	Cena 1 contendo um vídeo de fundo de 360 graus e um botão interativo para a cena 2. Cena 2 com uma imagem de 360 graus.	12	52
T5	Cena 1 contendo um vídeo de fundo de 360 graus e um botão interativo para a cena 2. Cena 2 com uma imagem de 360 graus e um vídeo 2D.	14	69
T6	Cena 1 contendo um vídeo de fundo de 360 graus e um botão interativo para a Cena 2. Cena 2 com uma imagem de 360 graus e um vídeo interativo 2D para a Cena 3. Cena 3 com um efeito sensorial de calor.	19	88

A Figura 6.24 apresenta a comparação do número de cliques e linhas necessários para as diversas tarefas T1-T6. Podemos ver que, à medida que a complexidade da tarefa aumenta de T1 a T6, fica claro que existe uma distância progressiva quase linear em relação ao número de cliques do AMUSEVR (azul) e número de linhas do MultiSEL (laranja) na Figura 6.24.

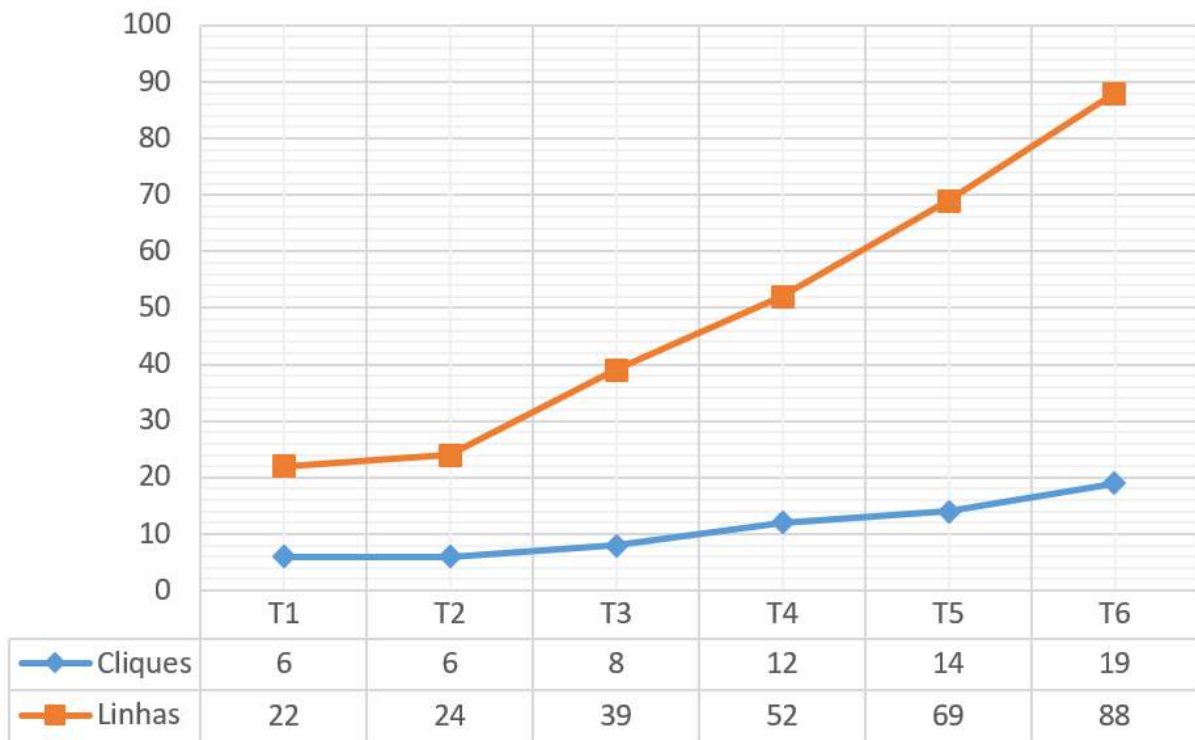


Figura 6.24: Cliques do AMUSEVR versus linhas do MultiSEL.

6.3.4.5 G5 - Análise dos resultados

Os dados completos do método estatístico Mann-Whitney [43] obtidos para todas as questões Q9-Q16 podem ser vistos na Tabela 6.26, bem como o resumo de todos os dados, Mediana (Med), Desvio Padrão (*Standard Deviation* - SD), Soma de Ranks (*Rank Sum* - RS), com foco em "U" (estatística do teste) e valor "p-exato" para cada questão na Tabela 6.26. O p-exato foi $> 0,05$ para as questões Q9, Q11, Q12, Q14, Q15 e Q16, portanto, não há diferença significativa entre as opiniões de ambos os grupos de participantes. Já em Q10 e Q13, o p-exato foi $< 0,05$, que significa divergência ou votação muito fora da curva para Q10 e de pequena significância estatística em Q13, que poderá ser melhor observado nas Figuras 6.25 e 6.26, que detalham as respostas dos grupos de experientes e novatos. Para Q10 "A organização do projeto multissensorial em cenas imersivas facilita a sua autoria", percebe-se que houve uma forte discordância no grupo de novatos. Para Q13 "Eu me senti incomodado visualmente ao usar o sistema", percebe-se novamente que houve uma concordância parcial no grupo de novatos, enquanto o grupo de experientes discordou fortemente.

Tabela 6.26: Resumo Q9-Q16.

Q#	experientes			Novatos			Mann-Whitney	
	Med	SD	RS	Med	SD	RS	U	p-exact
Q9	5,0	0,727	599,0	5	0,288	304,0	134,0	0,104
Q10	5,0	0,860	573,0	5	0,000	330,0	108,0	0,022
Q11	5,0	0,678	601,5	5	0,452	301,5	136,5	0,114
Q12	2,5	1,098	676,5	2	1,073	226,5	148,5	0,192
Q13	1,0	0,819	711,0	1,0	0,000	192,0	114,0	0,033
Q14	1,0	0,746	657,0	1,0	0,000	246,0	168,0	0,376
Q15	1,0	0,746	642,5	1,0	0,288	260,5	177,5	0,472
Q16	1,0	0,836	655,5	1,0	0,288	247,5	169,5	0,386

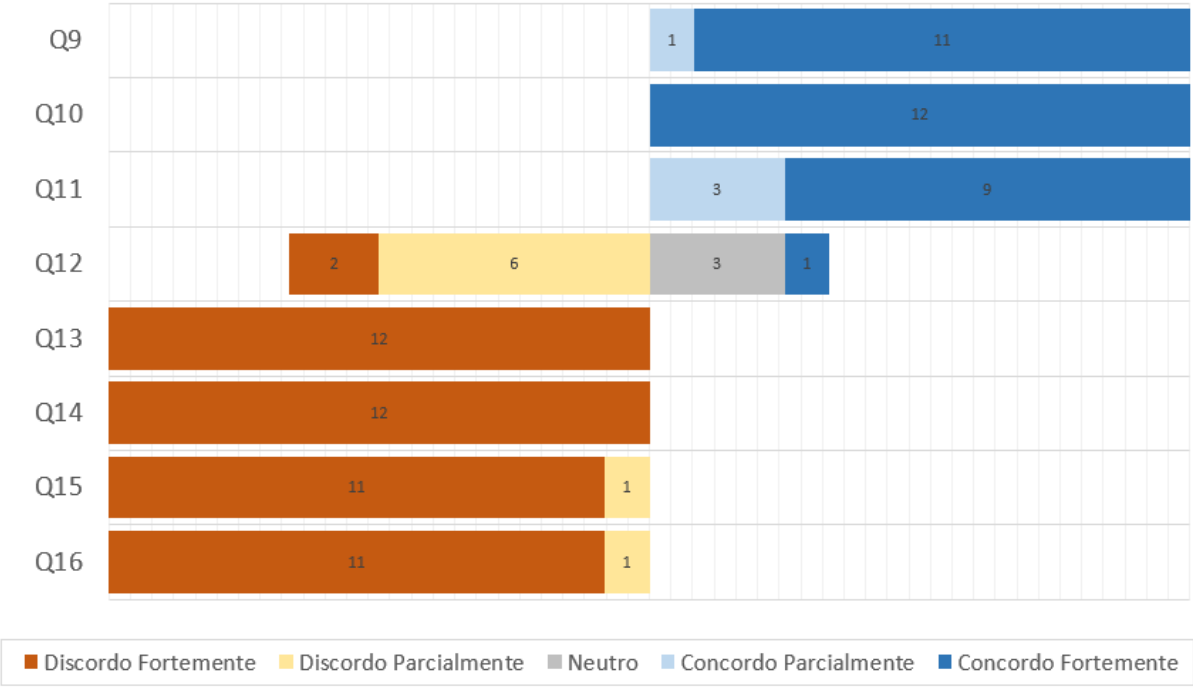


Figura 6.25: Gráfico comparativo quantitativo de Q9-Q16 para o grupo de experientes.

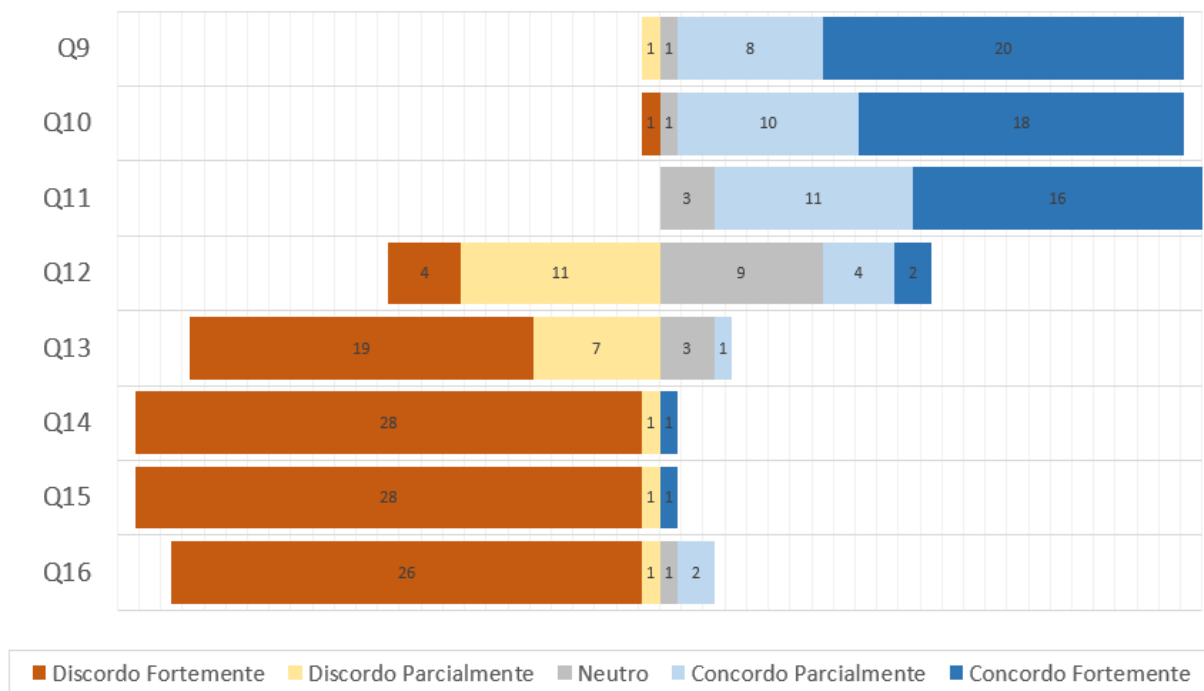


Figura 6.26: Gráfico comparativo quantitativo de Q9-Q16 para o grupo de novatos.

A seguir, utilizando as Figuras 6.25 e 6.26, serão descritos os resultados para as questões Q9 à Q16, considerando que as questões Q9 a Q11 são de fator positivo, enquanto as questões Q12 a 16 são de fator negativo:

Referente à questão Q9, demonstra que a maioria declarou ter se sentido imerso na ferramenta de autoria em ambos os grupos, com apenas uma discordância parcial e uma neutralidade no grupo de novatos.

Referente à questão Q10, demonstra que a maioria declarou que organização do projeto multissensorial em cenas imersivas facilita a sua autoria em ambos os grupos, com apenas uma neutralidade e uma discordância forte, ambos no grupo de novatos. Provavelmente esta forte discordância contrastando com a unanimidade do grupo de experientes que votou em “Concordo Fortemente”, foi responsável pela pontuação p-exato ter sido $< 0,05$.

Referente à questão Q11, demonstra que a maioria declarou que foi fácil posicionar espacialmente as mídias e efeitos sensoriais no ambiente imersivo do AMUSEVR, com apenas três neutralidades no grupo de novatos.

Referente à questão Q12, demonstra que a maioria discorda que NÃO necessitariam da realidade virtual para realizar as tarefas autorais nesta modalidade de projeto. Esta pergunta foi aplicada com um fator invertido (negativo), portanto é possível que isto tenha levado a mais diversidade nas respostas, porém para ambos os grupos, a maioria escolheu

respostas favoráveis à ferramenta AMUSEVR.

Referente às questões Q13, Q14, Q15 e Q16, que investigavam enjoos ou mal estar durante ao uso da ferramenta em realidade virtual (*Virtual reality sickness*), onde todas as questões apresentavam perguntas relacionadas a fatores negativos, observa-se a maioria em ambos os grupos discordando fortemente em todas as questões, com poucas exceções em algumas questões somente no grupo de novatos, apontando que o uso da ferramenta em realidade virtual é viável neste quesito.

6.3.5 Análise Subjetiva

Neste terceiro experimento foi observado que o tempo de ambientação livre antes da atividade, assim como no primeiro, foi muito importante, inclusive para apresentar os controles e as ações com os controles, além de diminuir possíveis receios com o uso da tecnologia.

Durante a atividade, assim como no primeiro experimento, foi observado em alguns participantes uma certa preocupação com a beleza da apresentação, em que outros exploravam mais opções do que os demais. Após isso foi detectado que os mais perfeccionistas e curiosos, normalmente eram mais experientes do que os que focavam diretamente em cumprir a atividade. Esse comportamento possivelmente foi o responsável pela menor média de tempo de quem executava diretamente as tarefas por não ter muita autoconfiança ou experiência com as tecnologias. Alguns usuários já haviam participado do primeiro experimento, classificando-se automaticamente no grupo de experientes e focando em conhecer melhor a ferramenta.

6.3.6 Discussão

Considerando os resultados dos experimentos realizados, pode-se concluir que:

- O objetivo G1 foi atingido, pois a pontuação do SUS de 78,51 pontos mostra que a usabilidade da ferramenta AMUSEVR foi aceitável e excelente do ponto de vista do usuário, conforme mostrado na Figura 6.19;
- O objetivo G2 foi atingido pelos resultados da UEQ, que indicaram que a experiência do usuário com a ferramenta AMUSEVR foi excelente considerando o benchmark da UEQ, conforme mostrado na Figura 6.21;

- Considerando o objetivo G3, concluímos que o AMUSEVR pode ser usado por usuários iniciantes, pois eles foram capazes de criar uma experiência interativa de 360 graus. Por outro lado, ao compararmos usuários experientes e iniciantes, nossos resultados mostram que apesar da média de tempo ser menor para o grupo experiente, a diferença não foi tão significativa, além disso podemos observar que a ferramenta foi bem assimilada por ambos os grupos.
- O objetivo G4 foi atingido considerando nossos resultados de teste, pois o AMUSEVR provou causar menos esforço de autoria do que o MultiSEL para criar diferentes experiências interativas de 360 graus; e
- O objetivo G5 também foi atingido considerando a avaliação da abordagem de autoria imersiva do ponto de vista do usuário. O ambiente de autoria teve uma boa avaliação dos participantes, que consideraram a imersão favorecida pelos diversos recursos como realidade virtual e efeitos sensoriais, também apontaram que a ferramenta facilita a autoria de projetos imersivos e corroboraram discordando da negativa na qual citava que a ferramenta não é necessária para esta atividade.

Finalmente, também pedimos aos usuários comentários e sugestões livres sobre o AMUSEVR. Recebemos sugestões sobre:

- Aspectos da interface do usuário (UI) - Alguns melhoramentos na UI foram sugeridos pelos usuários, como ampliar algumas fontes e pop-ups de mensagem que alertam o usuário para ações importantes, como funções de exclusão, por exemplo. Um exemplo de comentário do usuário sobre a UI foi "...Alguns itens de menu eram pequenos e o apontamento era difícil.";
- Feedback positivo - Recebemos elogios sobre imersão, fator de vanguarda e satisfação de uso. Um exemplo de elogio do usuário foi "O sistema é muito interessante de usar. Você pode se divertir enquanto cria conteúdo. Eu gostei muito, estou até pensando em estudar tópicos similares no futuro. Amei a inovação! Eu me senti como "O Cara da Tecnologia" em um filme de ficção científica.";
- Aspectos de portabilidade - sugestões para portar o AMUSEVR para outras plataformas de VR mais facilmente disponíveis, como aquelas semelhantes ao Google Cardboard¹¹. O objetivo é oferecer acesso mais fácil à ferramenta usando smartphones. Um exemplo de comentário do usuário sobre a portabilidade foi "...Eu gostaria de usar este sistema com meu smartphone..."

¹¹<https://arvr.google.com/cardboard/>

Sobre as sugestões dadas em formulário nesta nova versão, vimos que foi repetido a dificuldade em enxergar alguns textos, o que foi contornado nesta versão com a opção de aproximar ou afastar os menus, mas provavelmente devido à introdução de muita informação ao mesmo tempo no tutorial, alguns participantes podem não ter assimilado alguns comandos. Já sobre a possibilidade de portar para uma versão compatível com adaptações utilizando smartphones, é esperado grande perda de performance em alguns equipamentos devido à possibilidade de manipular inúmeras mídias em até alta resolução simultaneamente e aparelhos de baixa capacidade computacional.

6.4 Considerações Finais

Considerando a pergunta principal de pesquisa desta tese: *"Qual é a experiência do autor na criação de aplicações multissensoriais interativas em 360 graus usando um ambiente de autoria imersivo?"*. A pesquisa evoluiu e apresentou três fases de experimentação na qual obtiveram em todas, resultados promissores em relação aos experimentos. O primeiro experimento, foi usado como base para analisar a viabilidade do AMUSEVR, na qual teve a ferramenta avaliada através das metodologias da Escala de Usabilidade do Sistema (SUS), Questionário de Experiência do usuário (UEQ) e Questionário Personalizado. As etapas desse experimento foram documentadas com a metodologia *Goal Question Metric* (GQM). Como resultado, foi possível definir a viabilidade da ferramenta pois obtivemos resultados promissores, com isto e com as sugestões dos participantes, foi decidida a evolução e aprimoramento da pesquisa.

Durante o desenvolvimento, foi observada a necessidade da inclusão de um formato de portabilidade do projeto, no qual foi escolhido a linguagem MultiSEL como base. Para avaliação da mesma e com objetivo ainda de favorecer a resposta da pergunta principal da tese, elaboramos um segundo experimento totalmente online utilizando as metodologias *Goal Question Metric* (GQM), *Cognitive Dimensions of Notations* (CDN), de Questionário Personalizado e Método estatístico Mann-Whitney. Com elas foi possível avaliar e comparar o uso da linguagem para criação dos projetos, além de comparar com desenvolvimento pelo AMUSEVR. Como resultado, foi observado que ambas as abordagens são capazes de produzir conteúdos semelhantes, porém a proposta apresenta melhor receptividade dos participantes. Por fim, como experimento final e mais completo, foi utilizado o AMUSEVR em sua última versão (Oculus) com as metodologias *Goal Question Metric* (GQM), Escala de Usabilidade do Sistema (SUS), Questionário de Experiência do usuário (UEQ), teste Comparativo Quantitativo, um Questionário Personalizado e Método

estatístico Mann-Whitney, a fim de responder à pergunta principal da pesquisa.

Para tal, foram cruzados os resultados do primeiro e segundo teste a fim de encontrar algum desvio de resultado, que não foi encontrado, além de em ambos os experimentos apresentarem resultados aproximados. Além disso, em uso de ferramentas estatísticas diversas, buscamos analisar mais profundamente os resultados, que também foram bem avaliados neste terceiro experimento.

Como resultado geral dos experimentos, através do protótipo de ferramenta AMU-SEVR, foi revelado que o ambiente de autoria totalmente imersiva é viável e funcional. Os participantes dos experimentos concluíram com sucesso as tarefas, mostrando que é possível a criação de aplicações multissensoriais interativas em 360 graus usando um ambiente de autoria imersivo, além da experiência do usuário ser positiva.

Capítulo 7

Conclusão

Esta pesquisa propôs um método de autoria imersiva para criação e edição de aplicações mulsemídia interativas com conteúdo 360°. Atualmente, não há uma solução única para esta tarefa, apresentações deste gênero normalmente são desenvolvidas utilizando diversas ferramentas especializadas em outros conteúdos, mas desconhecemos alguma que ofereça uma solução totalmente imersiva para autoria. Esta proposta foi implementada em um ambiente de autoria imersiva denominado AMUSEVR, que vem demonstrando estar habilitado a cumprir o objetivo da pesquisa que é facilitar o desenvolvimento de projetos mulsemídia em ambientes imersivos.

O ambiente de autoria fornece uma imersão ao autor em realidade virtual, com liberdade criativa e a disponibilidade de usar uma ampla variedade de tipos de objetos de mídia e efeitos sensoriais. Os diversos experimentos realizados de usabilidade e experiência de usuário com dezenas de participantes apontam a viabilidade da mesma, pois todos os usuários concluíram suas tarefas de criação em todas as atividades propostas em todos os três experimentos, no qual precisavam produzir uma apresentação mulsemídia interativa 360° para um aplicativo turístico sem necessidade de auxílio adicional.

Dentre os resultados das metodologias utilizadas para avaliação, temos como a média de pontuação SUS 80,38 pontos e a média da avaliação UEQ-S foi considerada excelente pelos usuários em todos os quesitos. Na comparação de AMUSEVR com um método alternativo usando a linguagem MultiSEL, AMUSEVR se mostrou mais eficiente na avaliação quantitativa e, na avaliação subjetiva, se destacou com a preferência dos autores. Quando os resultados foram submetidos ao teste estatístico Mann-Whitney para comparar grupos de autores experientes e novatos, não encontramos diferença estatística significativa, indicando que AMUSEVR pode ser utilizada por usuários sem experiência em RV ou ferramentas de autoria multimídia.

Portanto, através dos experimentos conseguimos responder à pergunta principal de pesquisa desta tese: "*Qual é a experiência do autor na criação de aplicações multissensoriais interativas em 360 graus usando um ambiente de autoria imersivo?*", indicando grande potencial da abordagem imersiva para facilitar a criação de projetos multissensoriais interativos com conteúdo 360 graus na prática, já que a experiência relatada pelos usuários que realizaram os experimentos foi bastante positiva.

7.1 Publicações

Durante o tempo da concepção até a conclusão da tese, diversos trabalhos e publicações foram realizados como visto na Tabela 7.1.

Tabela 7.1: Tabela com trabalhos já publicados.

Ano	Publicação
2020	Memo-VR: Exercício Cognitivo para Idosos Utilizando Realidade Virtual e Interface com as Mãos [29].
2021	AMUSEVR: <i>A Virtual Reality Authoring Environment for Immersive Multimedia Applications</i> [27].
2022	<i>An Immersive Memory Game as a Cognitive Exercise for Elderly Users</i> [30].
2022	<i>Immersive Authoring of 360 Degree Interactive Applications</i> [28]
2022	AMUSEVR - <i>Authoring 360° Multimedia and Sensory Effects in VR</i> , software registrado no INPI - Instituto Nacional da Propriedade Industrial - registro número 512022003596-0, data de registro: 24/12/2022.

O primeiro trabalho publicado serviu de base para o aprendizado e adquirir experiência em desenvolvimento de aplicações em Realidade Virtual. Esta pesquisa resultou na publicação "Memo-VR: Exercício Cognitivo para Idosos Utilizando Realidade Virtual e Interface com as Mãos" [29] no Simpósio Brasileiro de Computação Aplicada à Saúde de 2020, apresentado no evento.

Em 2022, o mesmo trabalho foi atualizado e publicado como *An Immersive Memory Game as a Cognitive Exercise for Elderly Users* [30] nos Proceedings of the 2nd Life Improvement In Quality By Ubiquitous Experiences Workshop (LIQUE2022), evento internacional em conjunto com ACM IMX 2022, e apresentado no evento.

Em 2021, foram publicados os resultados preliminares já do AMUSEVR Vive no

workshop internacional SensoryX 2021, intitulado "AMUSEVR: *A Virtual Reality Authoring Environment for Immersive Multimedia Applications*" [27], em conjunto com ACM IMX 2021.

Em 2022, foi realizada uma publicação no journal IEEE Access intitulada "*Immersive Authoring of 360 Degree Interactive Applications*" [28], no qual apresentou uma atualização da pesquisa publicada na ACM IMX 2021, incluindo resultados de experimentos de mais usuários e a metodologia GQM para avaliação.

Por fim em 2022, foi realizado o registro de software da aplicação AMUSEVR Oculos com título "AMUSEVR - *Authoring 360° Multimedia and Sensory Effects in VR*" no INPI - Instituto Nacional da Propriedade Industrial.

Atualmente, temos dois artigos em processo de escrita, sendo um relacionado ao resultado dos experimentos do MultiSEL para AMUSEVR apresentado na Seção 6.2 e o outro mais completo sobre o AMUSEVR Oculus, apresentando dados sobre efeitos sensoriais e utilização do MultiSEL, dados esses obtidos e descritos apresentado na Seção 6.3.

7.2 Trabalhos futuros

Futuramente, como possíveis funcionalidades adicionais, será estudada a incorporação de objetos tridimensionais (3D) nos projetos do AMUSEVR. Outro objetivo futuro da pesquisa é avaliar a eliminação dos controles do HMD, fazendo uso somente do rastreamento das mãos como comandos no ambiente imersivo e implementar ferramentas que facilitem a visualização do cenário panoramicamente de forma a contribuir com a localização de objetos espalhados pelo cenário. Além disso, um possível gerenciador global da história de apresentação dos objetos em cena, oferecendo uma visão global de todas as cenas do projeto, o que pode ser muito útil para projetos maiores.

Caso seja implementado um módulo de trabalho colaborativo, pretendemos realizar atividades de criação utilizando a Internet. Com isso, será possível, por exemplo, criar projetos com o auxílio de até dois tutores em modo autoria, além de permitir a participação de vários colaboradores na apresentação do projeto. Imagine um projeto relacionado a uma sala de aula na qual existem vários alunos com seus avatares e um professor interagindo com eles da mesma forma que as aplicações chamadas de metaverso estão sendo difundidas.

Vislumbrando a possibilidade de novos experimentos, poderão ser realizados testes de

estresse, escalabilidade e até de esforço por cliques, além de objetos em 3D.

Estamos particularmente interessados em aplicações para educação, onde os professores podem usar AMUSEVR para criar conteúdo interativo em 360 graus para seus alunos. Gostaríamos de investigar como as experiências multissensoriais podem ser usadas para melhorar a educação em diferentes níveis.

Além disso, aproveitando a experiência que tivemos com aplicações voltadas à saúde de idosos, como no trabalho [29], estamos inclinados a promover mais aplicações usando o AMUSEVR para criar ambientes que possam simular situações para o tratamento de lembranças pós-traumáticas e, com isso, elaborar mais pesquisas relacionadas.

Por se tratar de uma ferramenta de autoria, será possível trabalhar em conjunto com outros projetos, como uma extensão da TV digital, novos experimentos utilizando outros tipos de efeitos sensoriais e até mesmo aplicá-la a atividades turísticas e culturais.

Referências

- [1] ABREU, R.; DE MATTOS, D.; SANTOS, J. A. D.; GHINEA, G.; MUCHALUAT-SAADE, D. C. Towards content-driven intelligent authoring of mulsemmedia applications. *IEEE MultiMedia* 28 (2021).
- [2] AMORIM, G. F.; MUCHALUAT-SAADE, D. C. Leiautes dinâmicos para documentos multimídia baseados em templates. *Anais Estendidos do Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia)* (10 2019), 9–12.
- [3] BANGOR, A.; KORTUM, P. T.; MILLER, J. T. An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction* 24, 6 (2008), 574–594.
- [4] BARRETO, F.; DE ABREU, R. S.; JOSUÉ, M. I. P.; MONTEVECCHI, E. B. B.; VALENTIM, P. A.; MUCHALUAT-SAADE, D. C. Providing multimodal and multi-user interactions for digital tv applications. *Multimedia tools and applications* 82, 4 (2023), 4821–4846.
- [5] BARRETO, F.; DE ABREU, R. S.; MONTEVECCHI, E. B. B.; JOSUÉ, M. I. P.; VALENTIM, P. A.; MUCHALUAT-SAADE, D. C. Extending Ginga-NCL to Specify Multimodal Interactions With Multiple Users. In *Anais Estendidos do XXVI Simpósio Brasileiro de Sistemas Multimídia e Web* (Maranhão, MA, Brazil, 2020), SBC.
- [6] BASILI, V. R. Software modeling and measurement: the goal/question. Tech. rep., Metric paradigm, 1992.
- [7] BATTISTI, A. L. E.; MUCHALUAT-SAADE, D. C.; DELICATO, F. C. Enabling internet of media things with edge-based virtual multimedia sensors. *IEEE Access* 9 (2021).
- [8] BERNS, A.; REYES SÁNCHEZ, S.; RUIZ RUBE, I. Virtual reality authoring tools for teachers to create novel and immersive learning scenarios. In *Eighth International Conference on Technological Ecosystems for Enhancing Multiculturality* (New York, NY, USA, oct 2020), ACM, pp. 896–900.
- [9] BLACKWELL, A.; GREEN, T. Notational systems—the cognitive dimensions of notations framework. *HCI models, theories, and frameworks: toward an interdisciplinary science. Morgan Kaufmann* 234 (2003).
- [10] BROOKE, J. Sus-a quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.
- [11] COELHO, H.; MELO, M.; BARBOSA, L.; MARTINS, J.; SERGIO, M.; BESSA, M. Immersive edition of multisensory 360 videos. In *Advances in Intelligent Systems and Computing* (2018), vol. 746, Springer Verlag, pp. 309–318.

- [12] COELHO, H.; MELO, M.; BARBOSA, L.; MARTINS, J.; TEIXEIRA, M. S.; BESSA, M. Authoring tools for creating 360 multisensory videos-evaluation of different interfaces. *Expert Systems* (2019).
- [13] COELHO, H.; MELO, M.; MARTINS, J.; BESSA, M. Collaborative immersive authoring tool for real-time creation of multisensory VR experiences. *Multimedia Tools and Applications* 78, 14 (2019).
- [14] COMSA, I.-S.; SALEME, E. B.; COVACI, A.; ASSRES, G. M.; TRESTIAN, R.; SANTOS, C. A. S.; GHINEA, G. Do I Smell Coffee? The Tale of a 360° Mulsemedia Experience. *IEEE MultiMedia* 27, 1 (2019), 27–36.
- [15] CONSORTIUM, W. W. W. W. Synchronized Multimedia Integration Language. w3.org/TR/SMIL/, 2008.
- [16] CONSORTIUM, W. W. W. W. HTML5. w3.org/TR/html5/, 2014.
- [17] CORDEIL, M.; CUNNINGHAM, A.; BACH, B.; HURTER, C.; THOMAS, B. H.; MARRIOTT, K.; DWYER, T. IATK: An immersive analytics toolkit. In *26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 - Proceedings* (mar 2019), Institute of Electrical and Electronics Engineers Inc., pp. 200–209.
- [18] COVACI, A.; GHINEA, G.; LIN, C.-H.; HUANG, S.-H.; SHIH, J.-L. Multisensory games-based learning-lessons learnt from olfactory enhancement of a digital board game. *Multimedia Tools and Applications* 77, 16 (2018), 21245–21263.
- [19] COVACI, A.; TRESTIAN, R.; SALEME, E. B.; COMSA, I.-S.; ASSRES, G.; SANTOS, C. A. S.; GHINEA, G. 360° Mulsemedia: a way to improve subjective QoE in 360° videos. In *Proceedings of the 27th ACM International Conference on Multimedia* (2019), pp. 2378–2386.
- [20] COVACI, A.; ZOU, L.; TAL, I.; MUNTEAN, G. M.; GHINEA, G. Is multimedia multisensorial? - A review of mulsemedia systems, aug 2018.
- [21] CRAIG, A. D. Interoception: the sense of the physiological condition of the body. *Current Opinion in Neurobiology* 13, 4 (aug 2003), 500–505.
- [22] DA SILVA, V. A. F.; DO NASCIMENTO, M. L.; DA SILVA, J. V.; BOVÉRIO, M. A. REALIDADE VIRTUAL. *Revista Interface Tecnológica* 14, 2 (dec 2017), 12.
- [23] DAMASIO, A. R. Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition* 33, 1-2 (nov 1989), 25–62.
- [24] DE MATTOS, DOUGLAS PAULO, D. C. M. S.; GHINEA, G. *An Approach for Authoring Mulsemedia Applications Based on Events*. Tese de Doutorado, Universidade Federal Fluminense, Niterói - RJ, 2011.
- [25] DE MATTOS, D. P.; C., M.-S. D.; GHINEA GHEORGHITA. Beyond Multimedia Authoring. *ACM Computing Surveys (CSUR)* 54, 7 (jul 2021), 1–31.

- [26] DE MATTOS, D. P.; MUCHALUAT-SAADE, D. C.; GHINEA, G. An Approach for Authoring Mulsemmedia Documents Based on Events. *2020 International Conference on Computing, Networking and Communications, ICNC 2020* (feb 2020), 273–277.
- [27] FARIAS, F. M. D.; DE MATTOS, D. P.; MUCHALUAT-SAADE, D. C. AMUSEVR: A Virtual Reality Authoring Environment for Immersive Mulsemmedia Applications. In *Proceedings of the 1st Workshop on Multisensory Experiences (SensoryX 2021)* (jun 2021), Brazilian Computing Society.
- [28] FARIAS, F. M. D.; MATTOS, D. P. D.; GHINEA, G.; MUCHALUAT-SAADE, D. C. Immersive authoring of 360 degree interactive applications. *IEEE Access* 10 (2022), 115205–115221.
- [29] FARIAS, F. M. D.; MONTEVECCHI, E.; BOKEHI, J.; SANTANA, R.; MUCHALUAT-SAADE, D. Memo-VR: Exercício Cognitivo para Idosos Utilizando Realidade Virtual e Interface com as Mãos. In *Anais do Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS)* (sep 2020), Sociedade Brasileira de Computacao - SB, pp. 434–439.
- [30] FARIAS, F. M. D.; SANTANA, R. F.; MONTEVECCHI, E.; BOKEHI, J. R.; MUCHALUAT-SAADE, D. An immersive memory game as a cognitive exercise for elderly users. *Proceedings of the Life Improvement in Quality by Ubiquitous Experiences Workshop (LIQUE)* (6 2022), 17–21.
- [31] FASSOLD, H.; TAKACS, B. Towards Automatic Cinematography and Annotation for 360° Video. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video* (2019), pp. 157–166.
- [32] FRAGA, T.; MENEZES, C. D. Um ambiente para autoria e realização de aplicações educacionais com realidade aumentada. In *Anais do XXVIII Simpósio Brasileiro de Informática na Educação (SBIE 2017)* (nov 2017), vol. 1, Brazilian Computer Society (Sociedade Brasileira de Computação - SBC), p. 937.
- [33] GHEORGHITA, G.; CHRISTIAN, T.; LINWEISI, R., G. S. Mulsemmedia. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* 11 (oct 2014).
- [34] GUEDES, Á. L. V.; AZEVEDO, R. G. D. A.; BARBOSA, S. D. J. Extending multimedia languages to support multimodal user interactions. *Multimedia Tools and Applications* 76, 4 (feb 2017), 5691–5720.
- [35] HORST, R.; DIEZ, S.; DÖRNER, R. Highlighting Techniques for 360° Video Virtual Reality and Their Immersive Authoring. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (oct 2019), vol. 11844 LNCS, Springer, pp. 515–526.
- [36] HORST, R.; NARAGHI-TAGHI-OFF, R.; RAU, L.; DÖRNER, R. Bite-Sized Virtual Reality Learning Applications: A Pattern-Based Immersive Authoring Environment. *Journal of Universal Computer Science* 26, 8 (2020), 947–971.
- [37] HWANG, W.; SALVENDY, G. Number of people required for usability evaluation: the 10±2 rule. *Communications of the ACM* 53, 5 (2010), 130–133.

- [38] JOSUÉ, M. I. P. *Preparação de Objetos de Mídia e Efeitos Sensoriais para Formatação de Documentos Mulsemídia*. Tese de Doutorado, Universidade Federal Fluminense- UFF, 2021.
- [39] JOSUÉ, M.; ABREU, R.; BARRETO, F.; DE MATTOS, D.; AMORIM, G.; DOS SANTOS, J.; MUCHALUAT-SAADE, D. Modeling sensory effects as first-class entities in multimedia applications. In *Proceedings of the ACM Multimedia Systems Conference (MMSys)* (2018), ACM, pp. 225–236.
- [40] LAUGWITZ, B.; HELD, T.; SCHREPP, M. Construction and Evaluation of a User Experience Questionnaire. *HCI and Usability for Education and Work, 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society 5298* (2008), 63–76.
- [41] LÉCUYER, A. BCIs and Video Games: State of the Art with the OpenViBE2 Project. *Wiley Online Library* (2016), 85–99.
- [42] LIKERT, R. A method of constructing an attitude scale. *Scaling: A sourcebook for behavioral scientists* (1974), 233–243.
- [43] MCKNIGHT, P. E.; NAJAB, J. Mann-whitney u test. *The Corsini encyclopedia of psychology* (2010), 1–1.
- [44] MENDES, P. R. C.; GUEDES, Á. L. V.; MORAES, D. D. S.; AZEVEDO, R. G. A.; COLCHER, S. An Authoring Model for Interactive 360 Videos. *2020 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)* (2020), 2–7.
- [45] MILGRAM, P.; TAKEMURA, H.; UTSUMI, A.; KISHINO, F. Augmented reality: a class of displays on the reality-virtuality continuum. *Telemanipulator and Telepresence Technologies 2351* (12 1995), 282–292.
- [46] MONTEVECCHI, E. B. B.; JOSUÉ, M. I. P.; ABREU, R. S. D. Providing Eye Gaze Interaction for Ginga-NCL Applications. *Webmedia 2020* (2020).
- [47] PARK, N. S.; HAN, M. Sensory Effect Authoring Tool for Virtual Reality Simulation. *9th International Conference on Information and Communication Technology Convergence: ICT Convergence Powered by Smart Intelligence, ICTC 2018* (2018), 1497–1499.
- [48] REC, I. H. 761, nested context language (ncl) and ginga-ncl for iptv services, geneva, apr. 2009, 2012.
- [49] REEVES, L. M.; MARTIN, J.-C.; MCTEAR, M.; RAMAN, T.; STANNEY, K. M.; SU, H.; WANG, Q. Y.; LAI, J.; LARSON, J. A.; OVIATT, S.; BALAJI, T. S.; BUISINE, S.; COLLINGS, P.; COHEN, P.; KRAAL, B. Guidelines for multimodal user interface design. *Communications of the ACM* 47, 1 (jan 2004), 57.
- [50] ROGERS, Y.; SHARP, H.; PREECE, J. *Design de interacao : alem da interacao humano-computador*. Bookman, 2013.
- [51] SALEME, E. B.; COVACI, A.; MESFIN, G.; SANTOS, C. A.; GHINEA, G. Mulsemmedia DIY: A survey of devices and a tutorial for building your own mulsemmedia environment, jun 2019.

- [52] SALEME, E. B.; SANTOS, C. A.; FALBO, R. A.; GHINEA, G.; ANDRES, F. Towards a reference ontology on mulsemmedia systems. In *MEDES 2018 - 10th International Conference on Management of Digital EcoSystems* (New York, NY, USA, sep 2018), Association for Computing Machinery, Inc, pp. 23–30.
- [53] SALEME, E. B.; SANTOS, C. A.; GHINEA, G. A mulsemmedia framework for delivering sensory effects to heterogeneous systems. *Multimedia Systems* 25, 4 (2019), 421–447.
- [54] SANTOS, C. A.; NETO, A. N.; SALEME, E. B. An Event Driven Approach for Integrating Multi-sensory Effects to Interactive Environments. In *Proceedings - 2015 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2015* (jan 2016), Institute of Electrical and Electronics Engineers Inc., pp. 981–986.
- [55] SCHREPP, M.; HINDERKS, A.; THOMASCHEWSKI, J. Construction of a Benchmark for the User Experience Questionnaire (UEQ). *International Journal of Interactive Multimedia and Artificial Intelligence* 4, 4 (2017), 40.
- [56] SILVA, E. P.; VIEIRA, N.; AMORIM, G.; MOUSINHO, R.; GUEDES, G.; GHINEA, G.; SANTOS, J. A. D. Using multisensory content to impact the quality of experience of reading digital books. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* 17 (11 2021).
- [57] SLATER, M.; SANCHEZ-VIVES, M. V. Enhancing our lives with immersive virtual reality. *Frontiers Robotics AI* 3 (dec 2016).
- [58] SOUZA, G.; SILVA, D.; DELGADO, M.; RODRIGUES, R.; MENDES, P. R.; AMORIM, G. F.; GUEDES, A. L.; DOS SANTOS, J. Interactive 360-degree Videos in Ginga-NCL Using Head-Mounted-Displays as Second Screen Devices. *WebMedia '20: Proceedings of the Brazilian Symposium on Multimedia and the Web* (2020), 289–296.
- [59] TAKÁCS, B. Immersive interactive reality: Internet-based on-demand VR for cultural presentation. *Virtual Reality* 15, 4 (nov 2011), 267–278.
- [60] TAKACS, B.; VINCZE, Z. Deep authoring - an AI Tool set for creating immersive MultiMedia experiences. *Multimedia Tools and Applications* (jan 2021), 1–30.
- [61] TAKACS, B.; VINCZE, Z.; FASSOLD, H.; KARAKOTTAS, A.; ZIOULIS, N.; ZARPALAS, D.; DARAS, P. Hyper 360—Towards a Unified Tool Set Supporting Next Generation VR Film and TV Productions. *Journal of Software Engineering and Applications* 12, 05 (2019), 127–148.
- [62] UNITY. Unity - Manual: Vídeo panorâmico, 2021.
- [63] VELHO, L. Interactive 360 VR. Tech. rep., Instituto de matemática pura e aplicada - IMPA, Rio de Janeiro, 2019.
- [64] XU, M.; LI, C.; ZHANG, S.; LE CALLET, P. State-of-the-art in 360 video/image processing: Perception, assessment and compression. *IEEE Journal of Selected Topics in Signal Processing* 14, 1 (2020), 5–26.

-
- [65] ZHANG, L.; ONEY, S. FlowMatic: An immersive authoring tool for creating interactive scenes in virtual reality. In *UIST 2020 - Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, oct 2020), Association for Computing Machinery, Inc, pp. 342–353.

APÊNDICE A – Códigos utilizados no Arduino

A.1 Código Arduino para controle de Efeitos Sensoriais

```
1 int aquecedor = 10;
2 int ventilador = 11;
3 int unity = 0;
4
5 void setup() {
6   pinMode(aquecedor, OUTPUT);
7   pinMode(ventilador, OUTPUT);
8   Serial.begin(9600);
9   digitalWrite(aquecedor, LOW);
10  digitalWrite(ventilador, LOW);
11 }
12
13 void loop() {
14   if (Serial.available() > 0) {
15     unity = Serial.read();
16     if(unity == 'a'){
17       digitalWrite(aquecedor, HIGH);
18     }
19     if(unity == 'b'){
20       digitalWrite(aquecedor, LOW);
21     }
22     if(unity == 'c'){
23       digitalWrite(ventilador, HIGH);
24     }
25     if(unity == 'd'){
26       digitalWrite(ventilador, LOW);
27     }
28   }
29 }
```

Listing A.1: Código para efeitos sensoriais em Arduino.

APÊNDICE B – Códigos utilizados no experimento MultiSEL

B.1 Código Inicial entregue aos participantes

```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <multisel id="Project-1" title="MultiSEL for AMUSE VR" xmlns:xsi="http://www.w3.org/2001/
  XMLSchema-instance" xsi:noNamespaceSchemaLocation="MultiSEL_Schema_to_AMUSEVR.xsd">
3   <head>
4     <meta name="author" value="seu nome aqui" />
5     <meta name="year" value="2022" />
6   </head>
7   <body primaryComponent="scene 1">
8     <scene id="scene 1">
9
10    </scene>
11
12   </body>
13 </multisel>

```

Listing B.1: Código inicial entregue nos experimentos.

B.2 XML esquema MultiSEL para Projetos AMUSEVR

```

1 <!--
2   XML MultiSEL Language Schema for the AMUSE VR
3   Copyr ight : 2022 LABORATORIO MIDIA COM, All Rights Reserved.
4   See https://www.midiacom.uff.br
5   Author: Flávio Miranda de Farias
6   Revision: 05/04/2022
7 -->
8 <xs:schema attributeFormDefault="unqualified" elementFormDefault="qualified" xmlns:xs="
  http://www.w3.org/2001/XMLSchema">
9   <!-- ***** Data type definitions ***** -->
10  <xs:simpleType name="mediaType">
11    <xs:restriction base="xs:string">
12      <xs:enumeration value="video"/>
13      <xs:enumeration value="image"/>

```



```

14     <xs:enumeration value="audio"/>
15     <xs:enumeration value="text"/>
16     <xs:enumeration value="image/x-icon"/>
17   </xs:restriction>
18 </xs:simpleType>
19
20 <!--
21 Esta limitacao de criar apenas 5 cenas foi incluída no Schema somente para o experimento
    com AMUSEVR, evitando que o usuário criassem muitas cenas. A linguagem MultiSEL nao
    limita o número de cenas em uma aplicacao.
22 -->
23 <xs:simpleType name="sceneType">
24   <xs:restriction base="xs:string">
25     <xs:enumeration value="scene 1"/>
26     <xs:enumeration value="scene 2"/>
27     <xs:enumeration value="scene 3"/>
28     <xs:enumeration value="scene 4"/>
29     <xs:enumeration value="scene 5"/>
30   </xs:restriction>
31 </xs:simpleType>
32
33 <xs:simpleType name="propertymediatype">
34   <xs:restriction base="xs:string">
35     <xs:enumeration value="duration"/>
36     <xs:enumeration value="background"/>
37     <xs:enumeration value="pip"/>
38     <xs:enumeration value="loop"/>
39     <xs:enumeration value="volume"/>
40     <xs:enumeration value="mute"/>
41     <xs:enumeration value="x"/>
42     <xs:enumeration value="y"/>
43     <xs:enumeration value="z"/>
44     <xs:enumeration value="text"/>
45   </xs:restriction>
46 </xs:simpleType>
47
48 <xs:simpleType name="effecttype">
49   <xs:restriction base="xs:string">
50     <xs:enumeration value="WindType"/>
51     <xs:enumeration value="HeatType"/>
52     <xs:enumeration value="VibrationType"/>
53   </xs:restriction>
54 </xs:simpleType>
55
56 <xs:simpleType name="propertyeffecttype">
57   <xs:restriction base="xs:string">
58     <xs:enumeration value="duration"/>
59     <xs:enumeration value="x"/>
60     <xs:enumeration value="y"/>
61     <xs:enumeration value="z"/>
62     <xs:enumeration value="intensity"/>
63   </xs:restriction>
64 </xs:simpleType>
65

```

```

66 <xs:simpleType name="interrelationtype">
67   <xs:restriction base="xs:string">
68     <xs:enumeration value="starts"/>
69     <xs:enumeration value="finishes"/>
70     <xs:enumeration value="onSelectionStarts"/>
71     <xs:enumeration value="onSelectionEnds"/>
72   </xs:restriction>
73 </xs:simpleType>
74
75 <xs:simpleType name="externrelationtype">
76   <xs:restriction base="xs:string">
77     <xs:enumeration value="starts"/>
78     <xs:enumeration value="finishes"/>
79     <xs:enumeration value="onSelectionStarts"/>
80     <xs:enumeration value="onSelectionEnds"/>
81   </xs:restriction>
82 </xs:simpleType>
83
84 <xs:simpleType name="keycoderelationtype">
85   <xs:restriction base="xs:string">
86     <xs:enumeration value="trigger"/>
87   </xs:restriction>
88 </xs:simpleType>
89
90 <!-- ***** struture Definitions ***** -->
91
92 <xs:element name="multisel">
93   <xs:complexType>
94     <xs:sequence>
95       <xs:element name="head" minOccurs="1" maxOccurs="1">
96         <xs:complexType>
97           <xs:sequence>
98             <xs:element name="meta" maxOccurs="unbounded" minOccurs="0">
99               <xs:complexType>
100                 <xs:simpleContent>
101                   <xs:extension base="xs:string">
102                     <xs:attribute type="xs:string" name="name" use="optional"/>
103                     <xs:attribute type="xs:string" name="value" use="optional"/>
104                   </xs:extension>
105                 </xs:simpleContent>
106               </xs:complexType>
107             </xs:element>
108           </xs:sequence>
109         </xs:complexType>
110       </xs:element>
111       <xs:element name="body" minOccurs="1" maxOccurs="1">
112         <xs:complexType>
113           <xs:choice maxOccurs="unbounded" minOccurs="0">
114             <xs:element name="scene" maxOccurs="unbounded" minOccurs="0">
115               <xs:complexType>
116                 <xs:choice maxOccurs="unbounded" minOccurs="0">
117                   <xs:element name="port" minOccurs="0" maxOccurs="unbounded">
118                     <xs:annotation>
119                       <xs:documentation>Portas desta cena</xs:documentation>

```

```

120         </xs:annotation>
121     <xs:complexType>
122         <xs:simpleContent>
123             <xs:extension base="xs:string">
124                 <xs:attribute type="xs:string" name="id" use="required"/>
125                 <xs:attribute type="xs:string" name="component" use="required
126             "/>
127         </xs:extension>
128     </xs:simpleContent>
129 </xs:complexType>
130 </xs:element>
131 <xs:element name="media" maxOccurs="unbounded" minOccurs="0">
132     <xs:annotation>
133         <xs:documentation>Mídias e Efeitos Sensoriais desta cena</xs:
documentation>
134     </xs:annotation>
135     <xs:complexType>
136         <xs:sequence>
137             <xs:element name="property" maxOccurs="unbounded" minOccurs="3"
138             >
139                 <xs:complexType>
140                     <xs:simpleContent>
141                         <xs:extension base="xs:string">
142                             <xs:attribute type="propertymediatype" name="name" use=
143                             "required"/>
144                             <xs:attribute type="xs:string" name="value" use="
145                             required"/>
146                         </xs:extension>
147                     </xs:simpleContent>
148                 </xs:complexType>
149             </xs:element>
150             </xs:sequence>
151             <xs:attribute type="xs:string" name="id" use="required"/>
152             <xs:attribute type="xs:string" name="src" use="required"/>
153             <xs:attribute type="mediaType" name="type" use="required"/>
154         </xs:complexType>
155     </xs:element>
156     <xs:element name="effect" maxOccurs="unbounded" minOccurs="0">
157         <xs:complexType>
158             <xs:sequence>
159                 <xs:element name="property" maxOccurs="6" minOccurs="4">
160                     <xs:complexType>
161                         <xs:simpleContent>
162                             <xs:extension base="xs:string">
163                                 <xs:attribute type="propertyeffecttype" name="name" use
164                                 ="required"/>
165                                 <xs:attribute type="xs:string" name="value" use="
166                                 required"/>
167                             </xs:extension>
168                         </xs:simpleContent>
169                     </xs:complexType>
170                 </xs:element>
171             </xs:sequence>
172             <xs:attribute type="xs:string" name="id" use="required"/>

```

```

167         <xs:attribute type="effecttype" name="type" use="required"/>
168     </xs:complexType>
169 </xs:element>
170 <xs:element name="relation" maxOccurs="unbounded" minOccurs="0">
171     <xs:annotation>
172         <xs:documentation>Relacoes desta cena</xs:documentation>
173     </xs:annotation>
174     <xs:complexType>
175         <xs:sequence>
176             <xs:element name="primary" minOccurs="1" maxOccurs="1">
177                 <xs:complexType>
178                     <xs:simpleContent>
179                         <xs:extension base="xs:string">
180                             <xs:attribute type="xs:string" name="component" use="
required"/>
181                         </xs:extension>
182                     </xs:simpleContent>
183                 </xs:complexType>
184             </xs:element>
185             <xs:element name="secondary" minOccurs="1" maxOccurs="1">
186                 <xs:complexType>
187                     <xs:simpleContent>
188                         <xs:extension base="xs:string">
189                             <xs:attribute type="xs:string" name="component" use="
required"/>
190                         </xs:extension>
191                     </xs:simpleContent>
192                 </xs:complexType>
193             </xs:element>
194         </xs:sequence>
195         <xs:attribute type="xs:string" name="id" use="optional"/>
196         <xs:attribute type="interrelationtype" name="type" use="required"
/>
197         <xs:attribute type="keycoderelationtype" name="keyCode" use="
optional"/>
198         <xs:attribute type="xs:int" name="delay" use="optional"/>
199     </xs:complexType>
200 </xs:element>
201 </xs:choice>
202 <xs:attribute type="sceneType" name="id" use="required"/>
203 <xs:attribute type="xs:string" name="primaryComponent" use="optional"/>
204 <xs:attribute type="xs:int" name="delay" use="optional"/>
205 </xs:complexType>
206 </xs:element>
207 <xs:element name="relation" maxOccurs="unbounded" minOccurs="0">
208     <xs:annotation>
209         <xs:documentation>Relacionamentos entre cenas</xs:documentation>
210     </xs:annotation>
211     <xs:complexType>
212         <xs:sequence>
213             <xs:element name="primary" minOccurs="1" maxOccurs="1">
214                 <xs:complexType>
215                     <xs:simpleContent>
216                         <xs:extension base="xs:string">

```

```

217         <xs:attribute type="xs:string" name="component" use="required"
218     "/>
219         <xs:attribute type="xs:string" name="interface" use="required"
220     "/>
221         </xs:extension>
222         </xs:simpleContent>
223         </xs:complexType>
224         </xs:element>
225         <xs:element name="secondary" minOccurs="1" maxOccurs="1">
226             <xs:complexType>
227                 <xs:simpleContent>
228                     <xs:extension base="xs:string">
229                         <xs:attribute type="xs:string" name="component" use="required"
230                     "/>
231                     </xs:extension>
232                     </xs:simpleContent>
233                     </xs:complexType>
234                     </xs:element>
235                     </xs:sequence>
236                     <xs:attribute type="xs:string" name="id" use="optional"/>
237                     <xs:attribute type="externrelationtype" name="type" use="required"/>
238                     <xs:attribute type="keycoderelationtype" name="keyCode" use="required"
239                 />
240             </xs:complexType>
241             </xs:element>
242             </xs:choice>
243             <xs:attribute type="sceneType" name="primaryComponent" use="required"/>
244         </xs:complexType>
245         </xs:element>
246     </xs:sequence>
247 </xs:element>
248 </xs:schema>

```

Listing B.2: XML esquema MultiSEL para projetos AMUSEVR.

B.3 Gabarito para Correção da atividade prática do experimento

```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <multisel id="Project-1" title="MultiSEL for AMUSE VR" xmlns:xsi="http://www.w3.org/2001/
3   XMLSchema-instance" xsi:noNamespaceSchemaLocation="MultiSEL_Schema_to_AMUSEVR.xsd">
4   <head>
5       <meta name="author" value="Flávio Miranda de Farias" />
6       <meta name="year" value="2022" />
7   </head>
8   <body primaryComponent="scene 1">
9       <!-- Cena 1 principal -->
10      <scene id="scene 1" primaryComponent="video3d-1_s1">

```

```

10      <!-- Porta da cena -->
11      <port id="portChangeScene1" component="image2d_s1" />
12      <!-- Midia e ES -->
13      <media id="video3d-1_s1" src="paodeacucar360.mp4" type="video">
14          <property name="background" value="true" />
15          <property name="loop" value="true" />
16          <property name="volume" value="0.5" />
17      </media>
18      <!-- Imagem a 1m de distancia com atrase de 1s -->
19      <media id="image2d_s1" src="cristo2d.jpg" type="image">
20          <property name="x" value="0" />
21          <property name="y" value="0" />
22          <property name="z" value="1" />
23          <property name="loop" value="true" />
24      </media>
25      <!-- Relacao interna a cena -->
26      <!-- Atraso da imagem -->
27      <relation id="Relation-s1-1" type="starts" delay="1">
28          <primary component="scene 1" />
29          <secondary component="image2d_s1" />
30      </relation>
31  </scene>
32  <!-- Cena 2 -->
33  <scene id="scene 2">
34      <media id="video3d-1_s2" src="cristo360.mp4" type="video">
35          <property name="background" value="true" />
36          <property name="loop" value="true" />
37          <property name="volume" value="0.3" />
38      </media>
39      <!-- ES de vento com duracao de 3s iniciado junto com a cena-->
40      <effect id="sewind_s2" type="WindType">
41          <property name="duration" value="3" />
42          <property name="x" value="0" />
43          <property name="y" value="1.7" />
44          <property name="z" value="0.5" />
45          <property name="intensity" value="0.7" />
46      </effect>
47  </scene>
48  <!-- Relacionamentos entre cenas -->
49  <relation id="Relation-1" type="onSelectionStarts" keyCode="trigger">
50      <primary component="scene 1" interface="portChangeScene1" />
51      <secondary component="scene 2" />
52  </relation>
53  </body>
54  </multisel>

```

Listing B.3: Gabarito para correção da atividades prática do experimento.