



ExploringFEEC: Uso de Realidade Mista em Jogo Baseado em Localização

Elton Cardoso do Nascimento João Pedro de Moraes Bonucci
Leonardo Rener de Oliveira Hélio Pedrini
José Mário De Martino Paula Dornhofer Paro Costa

Relatório Técnico - IC-PFG-23-36
Projeto Final de Graduação
2023 - Dezembro

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE COMPUTAÇÃO

The contents of this report are the sole responsibility of the authors.
O conteúdo deste relatório é de única responsabilidade dos autores.

ExploringFEEC: Uso de Realidade Mista em Jogo Baseado em Localização

Elton Cardoso do Nascimento* João Pedro de Moraes Bonucci*
Leonardo Rener de Oliveira* Hélio Pedrini* José Mário De Martino†
Paula Dornhofer Paro Costa†

Resumo

Este trabalho de conclusão de curso tem por objetivo relatar os avanços de desenvolvimento e análise de um jogo baseado em localização por realidade aumentada com a finalidade de potencializar o engajamento e aprendizado dos membros da instituição investigada. Descrevemos os avanços da realidade aumentada e de jogos no âmbito técnico e cultural nos últimos anos e como buscamos combiná-los em um jogo com diversos pontos de interação distribuídos por um espaço físico, apresentando os resultados e desafios encontrados durante o processo de desenvolvimento.

1 Introdução

A Realidade Estendida (XR) emerge como um campo inovador que redefine a interação entre o mundo digital e o ambiente físico, incorporando elementos digitais à realidade de maneira abrangente. Abrangendo Realidade Virtual (VR), Realidade Aumentada (AR) e Realidade Mista (MR), a XR transcende às fronteiras tradicionais, proporcionando experiências imersivas e interativas. O potencial da XR no cenário contemporâneo é vasto e abrange diversas indústrias, moldando desde entretenimento até setores cruciais como saúde e educação.

No mundo de hoje, a XR desenha novos horizontes ao possibilitar experiências inéditas e transformadoras. Seja na criação de ambientes virtuais envolventes para entretenimento, na revolução do marketing através de mundos virtuais interativos, ou na otimização de serviços imobiliários por meio de passeios virtuais, a XR oferece um leque de oportunidades. Além disso, a integração de tecnologias como Inteligência Artificial (IA) e Internet das Coisas (IoT) na XR amplia ainda mais suas capacidades, promovendo um futuro onde a fusão entre o digital e o físico redefine a maneira como vivemos, aprendemos e nos conectamos.

Outra área da tecnologia que testemunhou grande avanço técnico e cultural foram os jogos eletrônicos, cada vez mais presentes no cotidiano de todas as idades e que têm sido

*Instituto de Computação, Universidade Estadual de Campinas, Campinas, SP, 13083-852

†Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, SP, 13083-852

objeto de pesquisas variadas, buscando direcionar o grande engajamento gerado por eles como catalisador para outras atividades como o aprendizado. Estudos como [1], [2] e [3] são alguns exemplos notáveis dentre muitos outros de analisar e reconhecer o potencial de jogos no aprendizado e engajamento em uma atividade.

Com base neste contexto, este trabalho propõe explorar as tecnologias e desafios relacionados ao desenvolvimento de software e jogos para o contexto da realidade estendida com o objetivo final de gerar uma ferramenta capaz de gerar impacto positivo no aprendizado e interesse de um indivíduo pelo ambiente e da comunidade em seu entorno. A criação de um jogo de realidade estendida baseado em localização [4] como ferramenta visa validar as seguintes hipóteses em relação ao potencial da realidade estendida:

- Mecânicas em realidade mista diminuem o esforço necessário para entender tecnologias novas para a pessoa.
- Jogos em realidade mista baseados em localização aumentam o interesse das pessoas na história, cotidiano e pesquisas do ambiente.
- Jogos em realidade mista com conteúdo apresentando tecnologias podem gerar interesse em aprender mais sobre o tópico.
- Utilizar realidade mista incentiva as pessoas a utilizarem realidade mista.

Ao aprofundar na questão do interesse de um indivíduo pelo ambiente em seu entorno, procuramos inspiração no campo do teatro documental. Seguindo a revisão sobre o campo realizado no trabalho [5], ele se caracteriza pelo propósito de interpretar ou responder a um determinado momento histórico, investigando-o através do ponto de vista do documentarista. Entre os pontos que consideramos relevantes para dialogar com nosso trabalho, citamos:

- Intencionalidade de documentar, registrar fatos, pessoas ou lugares através de uma poética.
- Uso de documentos, que são o produto de um olhar sobre determinado fato, com esforço para decifrar as informações contidas nele e manifestados no suporte material do registro.
- Uso de dados não ficcionais, selecionados e articulados segundo a tese defendida pelo autor, que os seleciona, edita e articula.
- Uma noção expansiva do que pode ser considerado um documento foi criada historicamente, podendo chegar até à visão da pessoa documentada como documento vivo dentro da obra.

Consideramos então que, como a realidade mista é uma tecnologia que permite juntar elementos virtuais à nossa visão de realidade, ela possui a capacidade de criar objetos virtuais onde se colocam novas informações sobre o espaço ao mesmo tempo que integra

essas informações a ele. Jogos de realidade mista baseados em localização adicionam a interatividade: uma resposta ao jogador que cria um envolvimento com o processo, tendo então um *holding power* sobre ele [6]. Fugindo de uma visão de jogos como uma forma de escapar da realidade, vemos os jogos como uma possível forma de se integrar à ela.

Sendo uma pesquisa exploratória, propomos uma tentativa de entender melhor a problemática criada, gerando compreensões básicas para defini-la e estudá-la com profundidade, permitindo posteriormente a construção mais sistemática de metodologias para abordá-la. Inserimos este estudo no contexto da computação pela sua contradição do uso de objetos intangíveis e da capacidade de aproximá-los de objetos tangíveis através, por exemplo, da realidade mista. Queremos ainda apresentar ou realçar informações que poderiam passar despercebidas. Pesquisas em computação, que envolvem objetos intangíveis, muitas vezes não são vistas pelos corredores do lugar de pesquisa.

Para realizar estes estudos, escolhemos desenvolver um jogo em realidade estendida utilizando a *Game Engine* Unity¹, em que tentamos integrar o espaço físico e virtual de forma lúdica, gerando interesse no usuário a fim de validarmos as hipóteses propostas e explorar as problemáticas apresentadas. Como hardware para o projeto, utilizamos o óculos de realidade aumentada ThinkReality A3 da Lenovo, que opera conectado a um celular Android.

O jogo proposto consiste em uma narrativa de exploração localizada e limitada ao prédio da Faculdade de Engenharia Elétrica e Computação (FEEC) da Unicamp. Nele, criamos uma narrativa que direciona o jogador através do prédio a fim de apresentar os laboratórios, salas, pesquisas e pesquisadores da faculdade com o óculos de realidade aumentada e os personagens e pontos de interação descritos nas seções seguintes.

Devido à restrição de tempo, optamos por reduzir o escopo do projeto aos pontos de interações para termos apenas uma prova de conceito que ambicionamos ampliar nos próximos meses.

Este texto se organiza da seguinte forma. A Seção 2 apresenta as etapas propostas originalmente para o projeto, enquanto que a Seção 3 apresenta os desenvolvimentos realizados no design do projeto. A Seção 4 descreve os desenvolvimentos técnicos. Por fim, temos as conclusões na Seção 5, agradecimentos na Seção 6 e as referências bibliográficas.

2 Etapas Propostas

O plano de trabalho original previa o desenvolvimento do projeto em seis etapas, descritas a seguir junto dos **resultados esperados** e **resultados obtidos**:

1. **Desenvolvimento de *Non Playable Character* (NPC) Robótico:** Implementação de um NPC utilizando um robô humanoide para interação com o jogador através de movimentos e áudio. Para o NPC, um avatar virtual será renderizado, sendo necessário, portanto, a localização de suas juntas em tempo real. Será utilizado um QR code para localização do tronco do robô, com as posições das juntas sendo calculadas dinamicamente conforme o deslocamento do robô. Os movimentos serão passados

¹<http://unity3d.com/unity/>

para o robô através do *retargeting* de animações, com a restrição de que o robô estará sentado/fixo em um local, podendo movimentar apenas as juntas acima da cintura.

- **Resultado esperado:** sincronização do robô com a animação do avatar e projeção do avatar.
 - **Resultado mínimo:** apenas a visualização do avatar.
 - **Resultado obtido:** como descrito na Seção 4.2, devido a problemas logísticos não foi possível realizar a sincronização do robô, nem a projeção do avatar sobre ele, assim, a cena proposta apenas foi adaptada apenas para usar o avatar virtual.
2. **Interações Utilizando o Óculos de Realidade Mista:** Desenvolvimento de interações utilizando óculos de realidade mista, como o sistema de diálogo com o NPC. Outras interações planejadas incluem uma tela virtual de “carregamento” a ser exibida quando o jogador entrar no elevador da faculdade, para evocar elementos de jogos digitais; além de um mapa virtual do prédio.
- **Resultado esperado:** as três interações descritas.
 - **Resultado mínimo:** apenas o sistema de diálogo.
 - **Resultado obtido:** descritos na Seção 4.3, duas das três interações foram concluídas totalmente, com a visualização do mapa virtual realizado parcialmente.
3. **Pontos de Interação com Celular:** Criação de pontos de interação com realidade aumentada utilizando um celular. O início da interação em cada ponto será realizado utilizando a leitura de um QR code. Alguns pontos poderão incluir objetos físicos para auxiliar a interação.
- **Resultado esperado:** implementação de ao menos 5 pontos de interação, em sua maior parte com elementos físicos.
 - **Resultado mínimo:** implementação de ao menos 2 pontos de interação, com um deles com elemento físico.
 - **Resultado obtido:** descritos na Seção 4.4, dois pontos de interação novos foram implementados, um deles envolvendo um elemento físico. Dois pontos foram implementados parcialmente, e um ponto foi criado reaproveitando um projeto anterior.
4. **Coleta de Dados e Questionário:** Pretendemos implementar coleta de dados automática nos aplicativos para entender melhor as experiências dos jogadores, junto de um questionário a ser realizado após a participação. Coletaremos informações para entender questões como: engajamento, conforto com os equipamentos e interações; e índice de término das atividades em cada ponto de interação. Preparamos um projeto para submissão ao Comitê de Ética em Pesquisa da Unicamp no primeiro mês do projeto.

- **Resultado esperado:** implementação da coleta de dados no aplicativo, escrita do questionário e submissão do projeto ao Comitê de Ética em Pesquisa.
 - **Resultado mínimo:** escrita do questionário e submissão do projeto ao Comitê de Ética em Pesquisa.
 - **Resultado obtido:** descrito na Seção 4.1, realizamos a escrita do questionário e envio do projeto ao Comitê de Ética em Pesquisa.
5. **Exibição:** Exibição do jogo aberta ao público na FEEC. Ela ocorrerá em dias específicos.
- **Resultado esperado:** exibição aberta ao público durante ao menos 2 dias com 4 horas por dia.
 - **Resultado mínimo:** exibição restrita apenas para pessoas convidadas.
 - **Resultado obtido:** descrito na Seção 4.8, devido a problemas logísticos envolvendo a implementação do projeto e aprovação do projeto com o Comitê de Ética em Pesquisa, não conseguimos realizar a exibição.
6. **Escrita do relatório:** Escrita do relatório para a disciplina do Projeto Final de Graduação. Ele irá conter uma análise preliminar dos dados coletados.
- **Resultado esperado:** relatório contendo a descrição do desenvolvimento realizado, assim como análise dos dados coletados.
 - **Resultado mínimo:** relatório descrevendo apenas o desenvolvimento.
 - **Resultado obtido:** o presente relatório foi concluído, entretanto, não contendo a análise dos dados devido aos problemas envolvendo a coleta.

3 Design

O primeiro passo para iniciar o desenvolvimento do jogo foi a determinação de elementos de design que seriam implementados. Estas decisões estão descritas nesta seção, divididas em três partes: a Seção 3.1 apresenta o *design* do jogo em si, suas mecânicas básicas e sua estética. A Seção 3.2 apresenta o *design* da nossa personagem criada para o jogo, “INoVi”, e a Seção 3.3 descreve a narrativa criada para o jogo.

3.1 Estética e Game Design

Para a criação do design do jogo de forma mais global, decidimos empregar o uso do *framework* MDA (*Mechanics, Dynamics and Aesthetics*) [7]. O *framework* formaliza a experiência de uma pessoa com um jogo em três elementos: (i) mecânicas, as regras que descrevem os comportamentos do jogo; (ii) dinâmicas, a descrição do comportamento das mecânicas em contato entre si e com o jogador; e (iii) estética, a resposta emocional gerada no jogador. As mecânicas causam as dinâmicas, que causam a estética. O *game designer*

tem então como objetivo provocar uma determinada resposta estética no jogador, embora possa apenas determinar as mecânicas do jogo.

Seguindo a taxonomia proposta pelo *framework*, três objetivos estéticos foram escolhidos: **Descoberta** (Discovery), colocando o jogador para explorar o espaço da FEEC com um novo olhar; **Narrativa** (Narrative), que ajudará a manter o jogador interessado e dará um propósito para as atividades do jogo; e **Fantasia** (Fantasy), com o faz de conta criado ao inserir objetos e narrativa virtuais sobre o espaço real. Nossa inspiração baseou-se em jogos de mundo aberto para definir a principal mecânica criada para suscitar essas sensações, a de **Pontos de Interação**, em que diferentes interações isoladas seriam posicionadas na faculdade, cada uma apresentando uma mecânica diferente misturando objetos reais e virtuais. Estes pontos seriam explorados de forma não linear, potencializando a sensação de descoberta.

Foram pensados 18 pontos de interação no total, posicionados entre os 3 andares da faculdade. Dentre eles, alguns foram selecionados para implementação, como descrito na Seção 4. A Tabela 1 contém a descrição de todos os pontos, assim como a indicação se cada um foi implementado totalmente ou parcialmente.

Tabela 1: Ideias de pontos de interação.

Número e Seção	Descrição	Local pretendido
1	<i>Puzzle</i> envolvendo alinhar objetos no espaço 3D.	Janela do corredor da salas dos professores do 2 ^o andar, com visão para a Av. Albert Einstein, a FEQ e a FEM.
2	<i>Puzzle</i> narrativo.	Porta que abre para fora do prédio no 2 ^o andar, devido a sua estranheza.
3 Seção 4.3.4	Visualização do mapa da faculdade.	Hall do 2 ^o andar.
4	<i>Puzzle</i> envolvendo tecnologia de energia solar.	Inversores de energia solar no 2 ^o andar.
5	<i>Puzzle</i> envolvendo encontrar a chave de um dos armários.	Armários individuais do Departamento de Engenharia de Computação e Automação (DCA) no corredor principal do 2 ^o andar.
6 Seção 4.4.5.	<i>Puzzle</i> envolvendo interpretação de fala em Libras utilizando a Talita.	Sala de descanso do Laboratório do Departamento de Engenharia de Computação e Automação (LCA) (2 ^o andar).

Seção 4.4.4.	Visualização de um holograma da faculdade ou da UNICAMP sobre a mesa, e <i>puzzle</i> envolvendo manipular seus elementos.	Sala menor para reuniões do LCA.
8 Seção 4.4.3.	Projeção sobre a pintura de árvore na parede do corredor, com as folhas indicando a progressão até então no jogo.	Corredor do LCA.
9	Manipular objetos virtuais sobre um quadro de cortiça.	Quadro de cortiça no corredor das salas dos professor no 1º andar.
10	<i>Puzzle</i> histórico sobre a Semana de Arte Moderna.	Painel sobre a Semana de Arte Moderna no corredor das salas dos professor no 1º andar.
11 Seções 3.3, 3.2 e 4.3.2.	Cena inicial com a INoVi.	Entrada principal da FEEC (térreo).
12	<i>Puzzle</i> envolvendo matemática em uma lousa. Usar escrita para interagir com o <i>puzzle</i> , com auxílio da INoVi para resolver o problema.	Bitolódromo (corredor principal do 1º andar).
13 Seção 4.4.2	<i>Puzzle</i> envolvendo circuitos elétricos, acendendo uma lâmpada real ao interagir com o circuito virtual.	Não determinado
14	Montagem de circuito utilizando um kit educacional de eletrônica, com auxílio virtual.	Não determinado
15	Interação com um robô móvel sendo desenvolvido no LCA capaz de se locomover, guiar o usuário até pessoas ou objetos e interagir através de modelo de linguagem.	2º andar como um todo.

16	<i>Puzzle</i> envolvendo os perfis dos professores da faculdade, envolvendo a busca da sala de algum professor.	Corredor das salas dos professores no 2 ^o andar.
17 Seção 4.4.1	Hóquei em XR, reutilizando um projeto realizado anteriormente pelos integrantes do projeto.	Não determinado.
18 Seção 4.3.3	Tela de carregamento no elevador do prédio.	Elevador do prédio.

Um roteiro de questões foi desenvolvido para abordar o design de novos pontos de interação, tentando relacioná-los com o conceitos que desejamos explorar no projeto, os objetivos estéticos e a narrativa. Embora ele não tenha sido utilizado sistematicamente, foi utilizado como referência para pensarmos os pontos de interação.

Como Abordar Pontos de Interação

- Qual o local/objeto?
- O que ele significa?
 - No cotidiano das pessoas. Ações realizadas, significados.
 - Histórico.
- Como olhar dentro do discurso?
 - O que ele simboliza?
- Qual será a interação?
 - Baseado no objeto físico.
 - Como trazer descoberta?
 - * Como fazer o jogador interagir com o objeto de forma não usual?
- Como a interação irá ativar o transmissor?
 - Link com a história principal.
- Como apresentar informação sobre o ponto nas proximidades?
 - História opcional.
 - Textos, áudios, imagens.

3.2 Personagem

Dada a existência de uma narrativa no jogo e o objetivo do projeto de realização a criação de um avatar robótico, tivemos a necessidade de criar um personagem. Decidimos que ele teria o propósito de acompanhar e o auxiliar o jogador durante sua jornada, e narrativamente seria um assistente criado para a FEEC. Logo, um personagem não exageradamente formal, entretanto, profissional seria necessário. Ele seria o ponto inicial do jogo, precisando então ser carismática e prender o jogador.

Partimos para a busca de referências em jogos e filmes, e decidimos algumas características:

Antecedentes: IA criada para auxiliar a comunidade e pesquisas da universidade. Foi criada há aproximadamente 1 ano atrás, com potencial ainda a ser descoberto.

Objetivo externo: auxiliar a comunidade acadêmica.

Objetivo interno: entender seu papel e o que consegue fazer.

Traços de personalidade: carismática, inocente (amplo conhecimento, mas não sobre pessoas), prestativa, honestidade brutal (não mede palavras ao aconselhar ou criticar algo), lúdica.

Utilizamos o modelo de linguagem ChatGPT [8] como base para gerar uma descrição mais detalhada, conseguindo demonstrar a utilidade da ferramenta para o *design* de personagens por pessoas não experientes. O resultado final foi então:

“Apresentando ‘INoVi’ - a Inteligência Novadora para a Vida Acadêmica!”

INoVi é uma IA robótica criada há um ano e equipada com tecnologia de ponta e uma genuína sede de conhecimento, a INoVi está em uma jornada de auto-descoberta, buscando compreender plenamente suas capacidades e propósito no mundo.

Com inclinação profissional e sempre pronta para estender uma mão amiga, a INoVi é o rosto carismático das iniciativas de pesquisa da universidade. Sua inocência cativante e busca incansável pelo entendimento a tornam uma companheira confiável, buscando auxiliar estudantes e professores em suas buscas pelo conhecimento.

A honestidade da INoVi é tão marcante quanto brutal, nunca amenizando a verdade ou evitando expor informações óbvias. Pode não possuir empatia no sentido tradicional, mas sua abordagem equilibrada e raciocínio pragmático a tornam um recurso inestimável para resolver problemas complexos.

Apesar de sua postura profissional, a INoVi não se leva muito a sério, injetando ludicidade em suas interações para se envolver com o jogador e manter um ambiente descontraído. Sua atitude direta, combinada com um toque de charme, garante que o jogador possa sempre contar com a INoVi para fornecer as informações necessárias sem firulas desnecessárias.

Como o único personagem do jogo, a presença cativante e o carisma da INoVi prendem a atenção do jogador, criando um vínculo inquebrável entre o jogador e o companheiro de IA. Esteja preparado para embarcar em uma emocionante aventura de pesquisa, exploração e crescimento ao lado da cativante e sempre curiosa INoVi!”

Para a aparência da personagem, procuramos por ferramentas de baixa complexidade e custo para a criação do modelo 3D. Como iríamos realizar animações, ferramentas que já gerassem o “rig” do modelo (capacidade de manipular o movimento) foram priorizadas. Testamos primeiramente o site “Ready Player Me”, que permite a criação de personagens simples e exportáveis para a Unity [9]. Porém, consideramos a ferramenta limitada na capacidade de customização, e com estilo gráfico datado e difícil de ser replicado para outros modelos do jogo (Figura 1).

Como alternativa, utilizamos a ferramenta “VRoid Studio”, voltada para a criação de avatares para VTuber, mas que possui ferramentas criadas pela comunidade para uso de seus modelos dentro da Unity [10]. O resultado final pode ser visto na Figura 2 e foi utilizado devido a uma capacidade maior de customização e seu estilo cartunesco, que segue recomendações para aplicações em XR. A Seção 4.7 descreve como utilizamos ferramentas encontradas para realizar sincronização de dublagem com o modelo.



Figura 1: Modelos para a INoVi utilizando a ferramenta “Ready Player Me”.

3.3 Narrativa

A narrativa no nosso jogo tem como objetivo manter o jogador interessado em explorar os pontos de interação. Iniciamos seu desenvolvimento pelo objetivo que colocaria o jogador para andar pela faculdade: uma pane elétrica causada por um raio muito forte, danificando o “controle principal” que travou a porta principal e impede de sair do prédio. A ideia seria que o controle apenas seria ativado ao reestabelecer diferentes “pontos de controle”, que representam os pontos de interação. O objetivo do jogador é então reestabelecer os pontos de controle para conseguir sair do prédio.

Embora não linear, o jogo possui cenas de abertura, para imergir e dar o contexto ao jogador, e finalização fixas. O roteiro dessas cenas está disponível no GitHub: <https://github.com/INoVi>



Figura 2: Modelo final para a INoVi utilizando a ferramenta “VRoid Studio”.

[//github.com/ExploringFEEC/ExploringFEEC_Scripts](https://github.com/ExploringFEEC/ExploringFEEC_Scripts).

4 Desenvolvimentos

Nesta seção, descrevemos os desenvolvimentos realizados para a criação do jogo. Ela se divide primeiramente nos trabalhos realizados para possibilitar a coleta de dados (Seção 4.1), seguida pelo NPC Robótico (Seção 4.2). As Seções 4.3 e 4.4 descrevem as principais implementações das mecânicas e elementos do jogo em si. Por fim as Seções 4.5, 4.6 e 4.7 mostram desenvolvimentos técnicos de suporte como *networking*, fotogrametria e conversão de voz e dublagem, respectivamente; seguidas pela exibição na Seção 4.8.

4.1 Coleta de Dados e Metodologia Experimental

Para validar as hipóteses apresentadas, precisamos realizar sessões de demonstração com voluntários a fim de entender a percepção deles quanto à tecnologia e ao aprendizado. Idealmente, planejamos realizar a instrumentação do hardware para obtermos medidas internas de interação como direção dos olhos, número de interações e tempo de conclusão das tarefas propostas, no entanto estas ainda não foram implementadas até a escrita deste documento.

4.1.1 Metodologia de divulgação da pesquisa e coleta de dados

Para a divulgação e recrutamento dos participantes da pesquisa, enviaremos um e-mail aos estudantes de graduação da FEEC, através da Secretaria de Graduação. Pessoas que estiverem presentes na FEEC nos dias da demonstração também poderão participar.

Utilizaremos o jogo desenvolvido para a demonstração, sendo que:

- O método a ser utilizado para a demonstração e coleta de dados será realizado por meio dos passos a seguir:
- Apresentação dos pesquisadores e tecnologias utilizadas para os participantes Breve treinamento sobre o uso do óculos de realidade virtual e contextualização do roteiro da demonstração.
- Abertura para possíveis dúvidas dos participantes.
- Início da demonstração que dura em torno de 20 minutos em que o participante andarรก com um celular e óculos de realidade virtual pela FEEC realizando as tarefas propostas pelo roteiro do jogo.
- Término da demonstração
- Exibição do formulário online detalhado na seção de anexos deste documento.
- Escolha aleatória de alguns participantes para realizar a entrevista detalhada na seção de anexos deste documento.
- Encerramento e agradecimentos.

Especificamente sobre a entrevista, cinco participantes serão sorteados para responder após a demonstração, caso desejem, a uma entrevista de 20 minutos com um roteiro mais aprofundado de perguntas, no espaço da FEEC. Eles serão sorteados aleatoriamente ao término de cada demonstração individual, com a realização da entrevista neste momento, caso desejado, não sendo necessário retornar posteriormente.

Ao longo de todo o processo, os participantes serão acompanhados pelos pesquisadores para sanar possíveis dúvidas e atender a qualquer necessidade do participante durante a demonstração. Além disso, o participante poderá encerrar sua participação durante qualquer etapa e por qualquer motivo que considerar pertinente.

A pesquisa estava originalmente agendada para o final do mês de novembro, no entanto, postergamos a data devido ao processo de aceite do Comitê de Ética em Pesquisa da Unicamp.

4.2 NPC Robótico

Parte da proposta do projeto, a criação de um NPC Robótico envolveria a sincronização da movimentação de um robô humanoide junto com a animação da nossa personagem INoVi, com a projeção da personagem virtual sobre o corpo físico do robô. O planejamento deste desenvolvimento envolveria realizar o controle do robô, que teria um computador acoplado com a Unity, a comunicação via rede com o celular conectado ao óculos (Seção 4.5), a sincronização do movimento e a projeção. Estudos foram realizados com o SDK dos motores Dynamixel para permitir seu uso diretamente dentro da Unity [11], embora não tenha sido finalizado devido a problemas logísticos relacionados ao uso do robô humanoide. Estes problemas inviabilizaram a continuidade deste desenvolvimento no tempo deste projeto.

4.3 Interações com o Óculos de XR

No desenvolvimento dos pontos de interação em realidade aumentada, a plataforma Unity foi utilizada como base. Por meio dos kits de desenvolvimento de software (SDK) Snapdragon Spaces [12] e AR Foundation [13], os quais proporcionam uma camada de abstração superior para trabalhar com os recursos de AR e XR, foram realizadas implementações essenciais para criar um aplicativo Android capaz de reproduzir recursos de realidade estendida no óculos inteligente Lenovo ThinkReality A3, utilizado como peça fundamental neste projeto.



Figura 3: Lenovo ThinkReality A3.

Essas ferramentas permitiram a geração de interações imersivas e dinâmicas diretamente no Lenovo ThinkReality A3, expandindo as possibilidades de experiências em realidade estendida.

Com a combinação estratégica da Unity e dos SDKs mencionados, foi possível viabilizar a criação de mecânicas de jogo e pontos interativos e cativantes para o usuário, explorando os recursos avançados de realidade estendida oferecidos pelo óculos de XR e possibilitando novas formas de interação do usuário com o ambiente.

4.3.1 Posicionamento

Para garantir interações precisas e posicionamento exato dos objetos virtuais nos espaços da FEEC, implementamos um sistema baseado na distribuição estratégica de marcadores impressos (Figura 4) em locais selecionados como pontos de interação. Esses marcadores são identificados para mapear suas coordenadas no mundo virtual, permitindo que os objetos correspondentes sejam projetados nos óculos do usuário.

O processo de detecção e posicionamento dos marcadores foi conduzido utilizando o componente de rastreamento de imagens (ARTrackedImageManager) integrado ao SDK AR Foundation. Essa ferramenta, alimentada pela plataforma ARCore, oferece recursos fundamentais para aplicações de realidade aumentada, incluindo reconhecimento de plano, reconhecimento de imagem e identificação de objetos [14].

Com o intuito de personalizar a posição e rotação dos objetos virtuais em relação aos marcadores, desenvolvemos uma ferramenta específica. Esta ferramenta exibe as coordenadas de um objeto virtual inicialmente alinhado ao marcador, permitindo ao usuário movê-lo com o auxílio do sistema de interação por agarramento (Grab Interactable). Essa funcionalidade concede ao usuário controle total sobre a disposição do objeto no ambiente,



Figura 4: Modelo de cartaz utilizado no sistema de posicionamento de pontos de interação.

promovendo uma maior possibilidade de personalização dos pontos de interação, abordados na Seção 4.4.

4.3.2 Diálogo

Visando tornar mais acessível a interação com a realidade virtual [15], realizamos a implementação um sistema de diálogo sincronizado com as falas da INoVI. Esse sistema permite que os usuários ouçam e leiam os diálogos. A Figura 5 apresenta um exemplo de visualização de diálogo.

4.3.3 Elevador

Com o intuito de adicionar elementos de *gamificação* ao ambiente da FEEC, introduzimos a exibição virtual de uma tela de carregamento em um dos elevadores utilizados pelos usuários durante o percurso do jogo. Essa adição contribuiu para a imersão dos jogadores na experiência proposta.

4.3.4 Mapa Virtual

O mapa virtual desenvolvido tinha como objetivo oferecer aos jogadores uma representação precisa do ambiente, permitindo que identifiquem facilmente sua localização e os pontos de interesse disponíveis para interação.

Para o desenvolvimento de um mapa 3D virtual em escala real da faculdade, iniciamos com a tentativa de realizar uma reconstrução 3D do prédio, descrita com mais detalhes

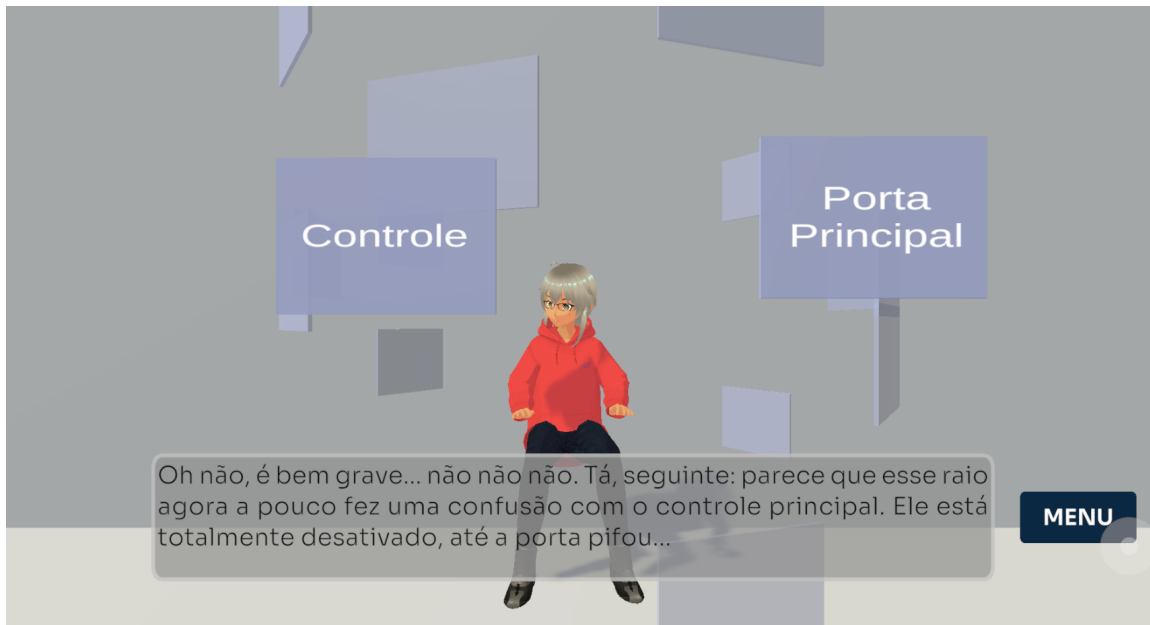


Figura 5: Exemplo de visualização de diálogo, com o botão para avançar.

na Seção 4.6. Devido às dificuldades encontradas, modelamos o prédio utilizando a planta baixa do como referência (Figura 6). Devido às simetrias existentes entre os andares do prédio, realizamos apenas a modelagem de um de seus andares, que será utilizada com pequenas alterações também para os outros.

Ainda são necessários avanços na exibição gráfica do mapa, e adição dos pontos de interesse.

4.4 Pontos de Interação

Os Pontos de Interação representam as principais mecânicas do nosso jogo, apresentando diferentes formas de interação em XR e de interação do mundo virtual com o físico. Nesta seção, apresentamos os pontos desenvolvidos.

Os pontos de interação oferecem pequenos desafios aos jogadores, empregando mecânicas de realidade estendida para interagir com objetos virtuais e físicos.

Para iniciar a interação em cada ponto, utilizamos a leitura de um QR code integrado ao sistema de posicionamento (Seção 4.3.1).

4.4.1 Hóquei

Este ponto de interação foi inspirado no clássico jogo de *air hockey* encontrado comumente em *arcades*. Nele, dois jogadores se enfrentam em uma mesa especial equipada com uma superfície lisa e perfurada que cria uma camada de ar. Usando um disco deslizante impulsionado por uma raquete especial, os jogadores buscam marcar gols no gol do oponente, exigindo habilidade, agilidade e estratégia para deslizar o disco na mesa e defender seu

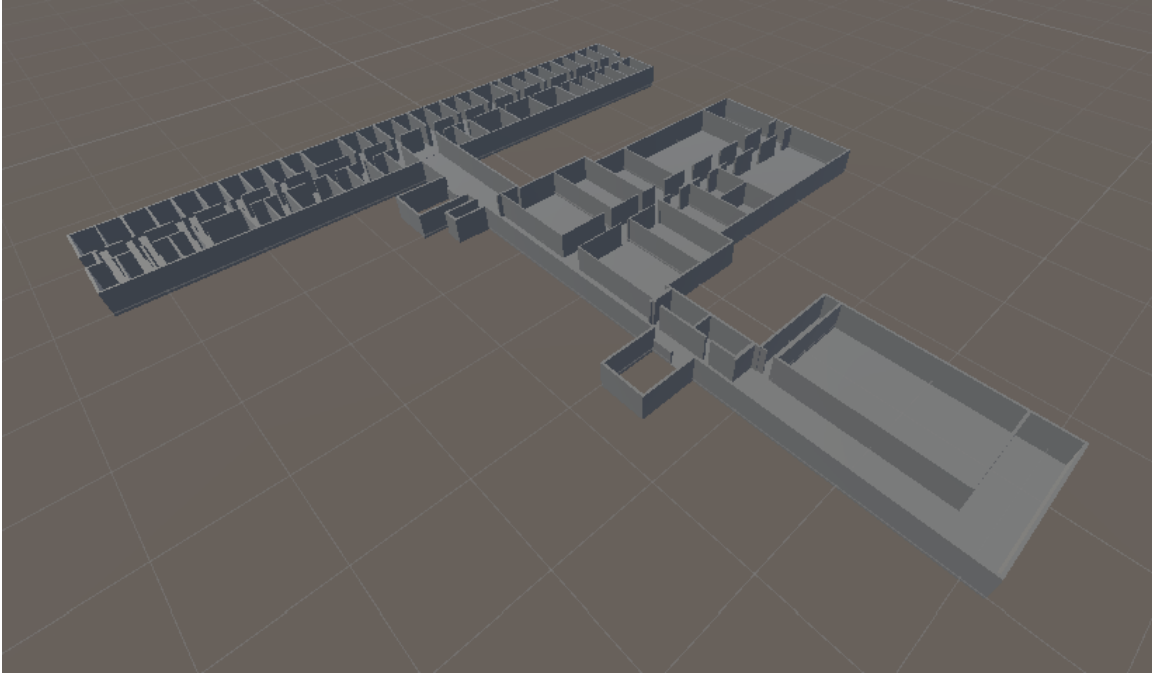


Figura 6: Modelo construído da FEEC.

próprio gol.

Nesta versão, implementada em realidade mista, a mesa, o disco e a raquete são elementos virtuais posicionados no mundo real por meio do sistema de posicionamento (Subseção 4.3.1). O jogador pode iniciar o jogo ao clicar em um botão virtual no centro da mesa e deve controlar sua raquete movendo a cabeça para a direita ou esquerda, com o objetivo de defender seu gol e marcar gols no adversário, que é controlado por código nesta versão. O jogo encerra quando o jogador ou o adversário realizam três gols, ou quando o tempo de 120 segundos se esgota.

4.4.2 Lâmpada

Neste ponto de interação, foram empregadas técnicas de *networking* (Subseção 4.5) para conectar um circuito real a um circuito virtual em realidade aumentada. O objetivo principal é estabelecer uma conexão entre o circuito virtual e uma lâmpada física, que é ativada quando o circuito virtual é concluído e um botão virtual é pressionado.

O jogador tem a oportunidade de interagir com os componentes do circuito virtual, como a *protoboard*, resistores, motor e LEDs, segurando-os com as mãos e posicionando-os nas posições corretas para completar o circuito. A representação esquemática e visual do circuito utilizado pode ser vista na Figura ???. O desafio proposto ao jogador é identificar as resistências dos dois resistores utilizados e selecionar os resistores de cinco bandas com as cores correspondentes, colocando-os adequadamente no circuito.

Para auxiliar no desafio, um cartaz impresso é disponibilizado no ponto de interação, explicando o significado de cada cor das bandas do resistor, além de fornecer um exemplo

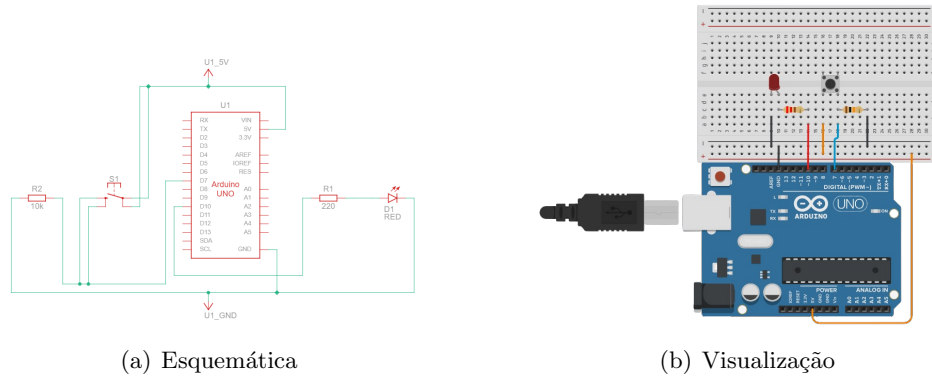


Figura 7: Circuito virtual utilizado no ponto de interação 4.4.2.



Figura 8: Simulação de interação com o circuito virtual.
Disponível em: <https://youtu.be/9JaLhSc3VgQ>

de como calcular a resistência, como pode ser visto na Figura 9. Essa abordagem visa não apenas proporcionar uma experiência interativa e imersiva, mas também promover a compreensão prática dos conceitos por trás dos circuitos elétricos, tornando o aprendizado mais dinâmico e envolvente.

4.4.3 Parede do LCA

Este local foi escolhido como o mapa de progresso do jogo. No final do corredor do Laboratório do Departamento de Engenharia de Computação e Automação, há um papel adesivo de uma árvore com 18 folhas e 7 ramos que optamos por servirem como guia de progresso ao iluminarmos folhas e ramos conforme os desafios dos pontos de interação fossem concluídos. Para isso fomos ao local realizar a fotogrametria do adesivo e seus arredores objetivando mapear suas folhas para o jogo (Subseção 4.6). O resultado pode ser visto na Figura 15, no entanto, o desenvolvimento desse ponto foi descontinuado devido a redução de pontos de interação, assim a ideia de representação nas folhas e/ou ramos seria numericamente incompatível.

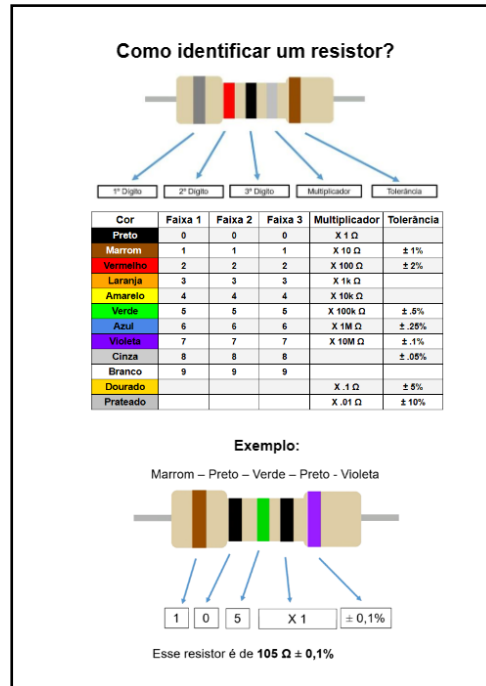


Figura 9: Cartaz impresso utilizado para auxiliar identificação de resistores.

4.4.4 Sala/Mesa LCA (terreno)

Nesta interação, apresentamos ao jogador uma versão tridimensional em realidade aumentada do mapa do campus da Unicamp, permitindo a visualização de ruas e edifícios da universidade. O desafio proposto ao jogador é organizar corretamente a posição de três institutos que não estão inicialmente identificados no mapa.

Para criar o mapa de maneira realista e em escala, utilizamos o software Blender [16] em conjunto com o complemento Blender-GIS [17]. Este complemento oferece a capacidade de importar mapas de satélite, mapas de deslocamento e geometria, incluindo edifícios. Ao selecionar as coordenadas desejadas na interface do complemento, é gerada uma visualização do mapa por meio dos dados fornecidos pelo OpenStreetMap. Além disso, os modelos 3D dos objetos, como edifícios e estradas, foram criados a partir de dados reais de elevação da missão SRTM da NASA. O resultado pode ser visualizado na Figura 10.

Essa abordagem permite ao jogador explorar o campus da Unicamp de uma maneira envolvente e interativa, proporcionando uma visão tridimensional realista do ambiente universitário e desafiando-os a identificar e posicionar corretamente locais específicos no mapa. Contudo, devido à não conclusão de sua implementação até a redação deste documento, este ponto de interação ainda não foi incluído na versão atual do jogo.

4.4.5 Talita

Este ponto de interação é um reflexo do projeto de Tecnologias Assistivas Para Surdos (TAS), fruto da colaboração entre pesquisadores de diversas instituições renomadas, como

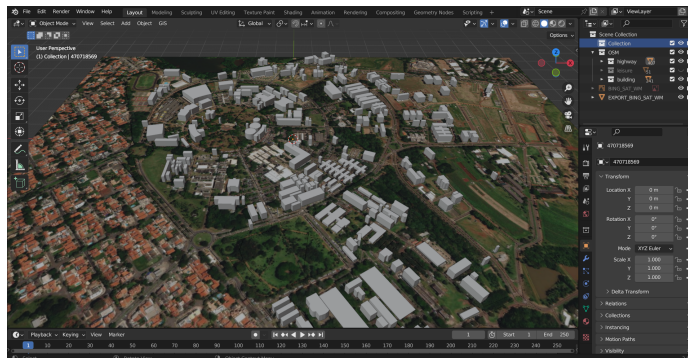


Figura 10: Visualização do mapa da Unicamp criado utilizando Blender-GIS.

a FEEC, o Centro de Estudos e Pesquisa em Reabilitação Prof. Dr. Gabriel Porto da Faculdade de Ciências Médicas (CEPRE/FCM), o Instituto de Estudos da Linguagem (IEL/Unicamp), o Instituto de Computação da Universidade Federal de Alagoas (IC/UFAL), o Centro Tecnológico da Universidade Federal do Espírito Santo (CT/UFES) e o Centro de Tecnologia da Informação Renato Archer (CTI).

No âmbito desta interação, a protagonista é a Talita, um avatar humanoide desenvolvido no contexto do projeto TAS, atuando como intérprete de Libras, empregando técnicas de captura de movimento [18]. O desafio proposto ao jogador é identificar letras do alfabeto de Libras, comparando os sinais produzidos pelo avatar com os símbolos dos sinais exibidos em uma TV.

O jogador assume o controle por meio de botões virtuais, utilizando as mãos para interagir e pressionar os botões na interface (Figura 11). Esta interação conduz a comunicação entre os botões virtuais e a TV, empregando técnicas de *networking* implementadas (Seção 4.5).

Essa abordagem oferece uma oportunidade singular de explorar as potencialidades das tecnologias assistivas para surdos, combinando-as de forma criativa e prática com os recursos imersivos da Realidade Aumentada, contribuindo assim para uma experiência educativa e inclusiva.

4.5 Networking

Focando na interação entre físico e virtual precisamos recorrer a comunicação com ou sem fio para não impactar a mobilidade necessária para o jogo. Para isso exploramos duas formas de comunicação disponíveis na ferramenta Unity. Visando a interação entre virtual e físico recorreremos ao circuito integrado ESP32 [19] 12 para fazer a interface entre o jogo e o acionamento de qualquer dispositivo físico como lâmpadas, motores, sensores, telas, etc através do Wi-Fi utilizando protocolo HTTP [20]. Adicionalmente, para comunicações mais intensivas, empregamos a Comunicação Remota de Procedimentos (RPC, do inglês Remote Procedure Call) [21], utilizando a biblioteca Netcode For GameObjects. Esta biblioteca está em constante desenvolvimento e é reconhecida como um padrão para soluções de redes de computadores em projetos construídos na plataforma Unity.

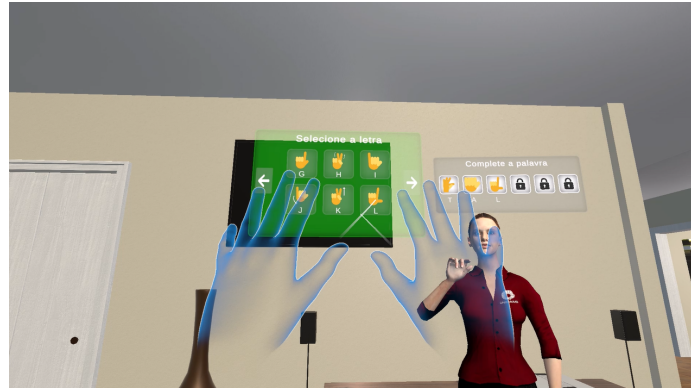


Figura 11: Simulação de interação com a Talita.
 Vídeo disponível em: <https://youtu.be/ngLJyKZG64Q>

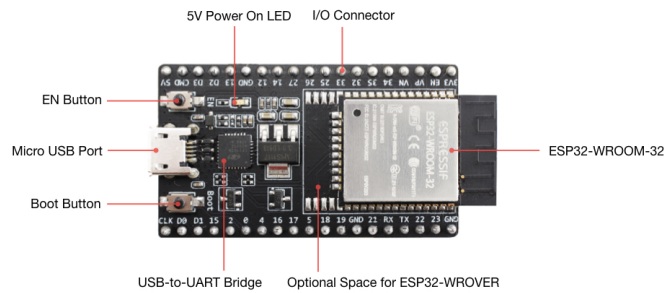


Figura 12: Ilustração do circuito integrado ESP32

4.5.1 HTTP

A primeira forma de comunicação que utilizamos foi uma configuração simples de cliente servidor local onde o circuito integrado ESP32 disponibiliza um IP (*Internet Protocol*) com URLs personalizadas para cada ponto de interação seguindo o modelo REST [22] adotado como padrão para conexões Web.

Para o circuito integrado que realiza o papel de servidor, precisamos apenas definir o código do Arduino com os padrões de URL esperados e o tipo de integração, ou seja, se vamos utilizar o circuito integrado como ponto de acesso de uma rede Wi-Fi ou como cliente de uma rede já existente. Essa distinção se tornou necessária devido ao padrão de segurança suportado pelo circuito integrado estar restrito ao WPA, o que impossibilita que o ESP32 se conecte a redes sem fio com outros padrões de acesso protegido. Sendo assim, é suficiente expor o IP na rede à qual o circuito integrado se conectou ou o IP gerado por ele mesmo, no caso da configuração de acesso direto, aguardando alguma requisição para ser ativado.

Em paralelo, do lado do cliente foi preciso utilizar a biblioteca UnityWebRequest [23], uma solução modular desenvolvida pela própria Unity Technologies para requisições HTTP, suportando todos os métodos [24]. Essa biblioteca precisa ser adicionada a uma instância em

C# referenciada por um objeto virtual qualquer responsável por realizar aquela requisição assim que um evento é completado, lidando com eventuais problemas de latência ou erro de requisição.

4.5.2 Netcode For GameObjects

Devido às interações entre personagens em dispositivos distintos no contexto do jogo, se mostrou importante investigar soluções para a implementação de funcionalidades multijogador na plataforma em questão. Tal necessidade surge da exigência, em alguns dos pontos de interação inicialmente delineados, de dispor de instâncias separadas do jogo operando em máquinas distintas. Essa abordagem visa assegurar a manutenção da sincronia e resiliência indispensáveis para proporcionar uma experiência de jogo mais envolvente.

Nesse contexto, optou-se pela incorporação da biblioteca Netcode For GameObjects [25], a qual se destina a otimizar a comunicação de rede para jogos multijogador, a biblioteca faz uso de chamadas de procedimento remoto (*Remote Procedure Calls*) para a interação entre cliente e servidor. Essa biblioteca demonstra uma integração robusta com a ferramenta Unity, apresentando suporte por meio de uma interface gráfica dedicada e comandos no-code para cenários mais simples. Essa característica simplifica significativamente o processo de criação de ambientes multijogador.

Com ela conseguimos abstrair muitas das animações e mensagens necessárias para sincronia de movimentos complexos com a ajuda do *NetworkManager*, um objeto responsável pela gestão de requisições, e algoritmos vinculados os *Game Objects*, abstração para qualquer porção de código interativa (personagens, jogadores, objetos com física, entre outros). Com isso, conseguimos abstrair toda a parte de rede como tipos específicos, como a classe *NetworkBehaviour*, lidando apenas com a possível concorrências da execução de determinados algoritmos.

4.6 Fotogrametria

Para o desenvolvimento de algumas interações do jogo, e da visualização do mapa da faculdade (Subseções 4.4.3 e 4.3.4), identificamos a necessidade de obter uma representação 3D do prédio como um todo e do corredor do LCA. Para estudarmos o processos e resultados que poderíamos obter, decidimos utilizar fotogrametria, que é o processo de obtenção de dados de um espaço físico utilizando fotos. Neste caso, realizamos o processo de reconstrução 3D, em que um conjunto de fotos é utilizado para criar um modelo 3D da geometria da cena.

Por conta do nosso prazo e de não ser o foco principal do projeto, decidimos utilizar uma ferramenta já pronta para fotogrametria. Utilizamos para isso o *Meshroom*, por ser uma ferramenta gratuita [26]. Ele possui diversas operações para reconstrução 3D disponíveis em nós configuráveis em *pipelines*, visando a uma alta flexibilidade de processos e experimentação de parâmetros. Para o processamento das fotos capturadas, utilizamos o *darktable* [27]. Por fim, para a simplificação e *baking* de texturas utilizamos o Blender [16]. Como referência para o processo seguimos uma aula de fotogrametria disponível publicamente [28], assim como a documentação do *Meshroom* [29].

As Figuras 13, 14 e 15 contêm uma representação do processo realizado. O primeiro passo consistiu na captura de fotografias. Utilizamos a câmera *EOS Rebel T5i*, um tripé e um cartão cinza de referência para balanço de branco. As fotos foram capturadas em formato bruto “CR2”, visando à fidelidade visual e qualidade dos dados, com sensibilidade do sensor ISO 100, abertura f/8, e distância focal fixa, visto que o software de reconstrução lida melhor com fotos com mesmo foco.

Fotos foram capturadas com o cartão de referência no início da captura e em mudanças bruscas de iluminação. Ocorreram problemas na coleta devido a entrada de luz natural no ambiente e iluminação sintética, principalmente ao capturar grandes extensões, devido a mudança na exposição necessária e no balanço de branco.

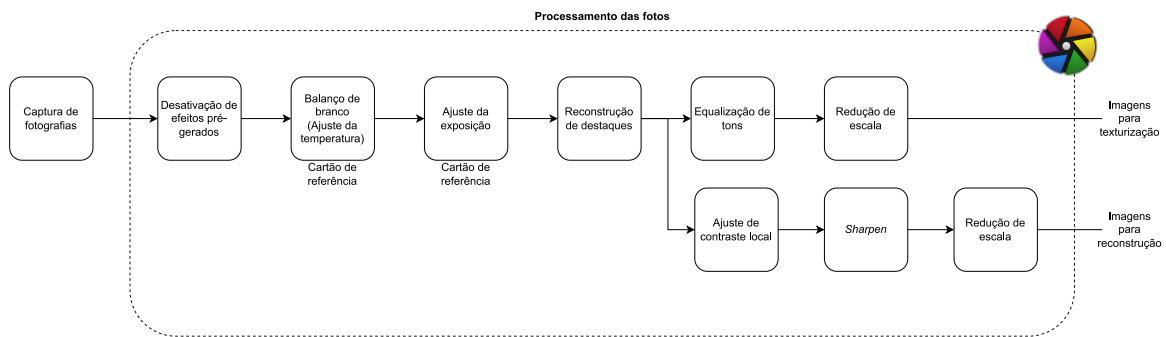


Figura 13: Fluxo utilizado para o processamento das fotos.

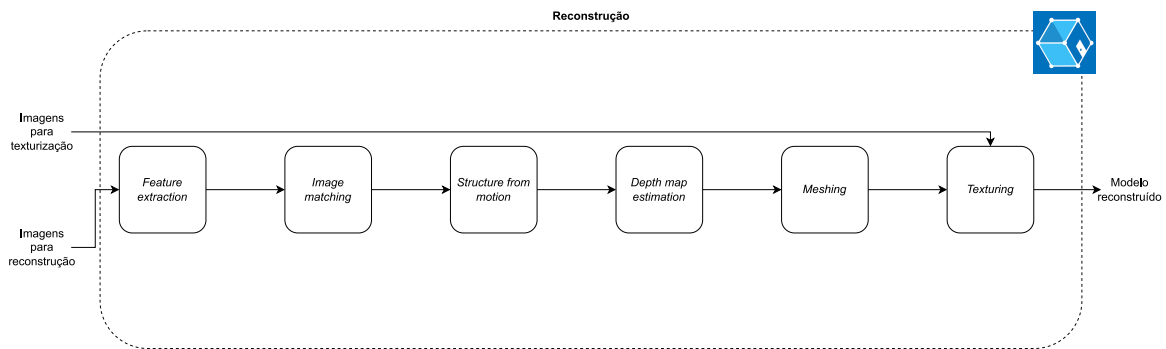


Figura 14: Fluxo utilizado para a reconstrução do modelo.

Após a coleta, realizamos o processamento das fotos com o *darktable*. Para cada captura, foram gerados dois conjuntos de fotos: o primeiro buscando alta nitidez e contraste para uso na reconstrução, e o segundo buscando fidelidade visual e minimização da influência da iluminação para uso na texturização do modelo gerado (geração das texturas albedo/*base color*).

A reconstrução foi realizada utilizando o fluxo padrão de reconstrução do *Meshroom*, que realiza:

- *Feature extraction*: extração de pontos de interesse nas imagens, tornando possível verificar a semelhança de diferentes pontos em diferentes imagens.

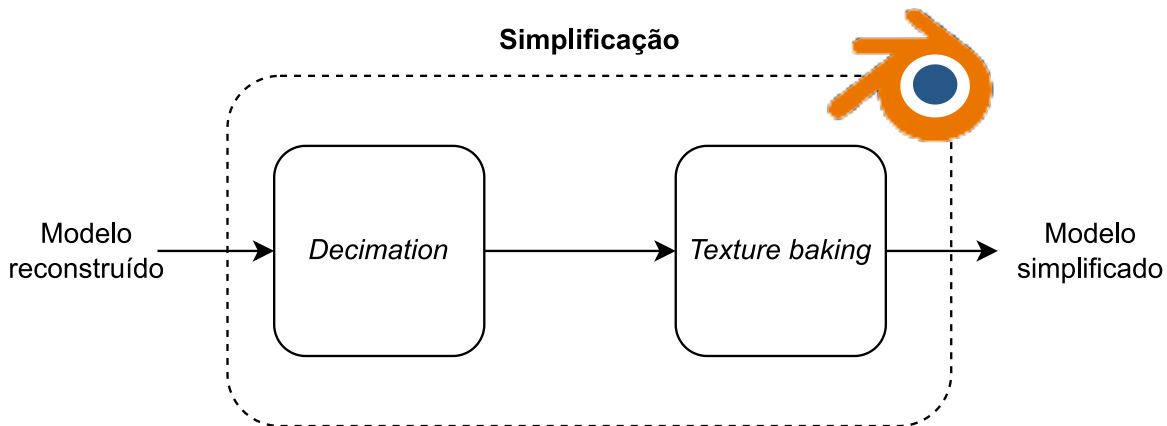


Figura 15: Fluxo utilizado para a simplificação do modelo.

- *Image matching*: procura de imagens semelhantes utilizando os atributos detectados.
- *Feature matching*: procura de todos os atributos semelhantes nos pares de imagens selecionados.
- *Structure from motion (SfM)*: criação de um *point cloud* 3D, obtenção da pose da câmera em cada foto e calibração intrínseca de todas as câmeras utilizadas. Utiliza uma reconstrução incremental a partir de uma reconstrução inicial com um par de imagens. Uma imagem é dita *alinhada* caso seja possível inserir ela na reconstrução.
- *Depth map estimation*: cálculo da profundidade de cada pixel em cada imagem.
- *Meshing*: criação da geometria da cena. Também realiza filtros para melhoria da qualidade da geometria gerada.
- *Texturing*: geração do *UV map* e textura de cor da cena.

Algumas alterações foram realizadas no fluxo: para acelerar a execução do nó *Feature-Extraction*, alteramos o parâmetro *Force CPU Extraction* para falso. Devido a problemas iniciais com uma quantidade baixa de câmeras sendo alinhadas na reconstrução, realizamos os passos descritos na documentação para reconstrução esparsa ². Por fim, fizemos as alterações descritas no curso de referência para realizar a texturização com o segundo conjunto de fotos.

Por fim, o processo de *decimation* (simplificação) e *baking* de texturas é realizado no Blender, visando diminuir a quantidade de vértices existente no modelo, o que é importante visto que o jogo executa em uma plataforma móvel, mas ainda mantendo o máximo possível de detalhes do modelo original. Foram geradas as texturas de *base color* (albedo), *normal map* e *ambient occlusion*.

²<https://github.com/alicevision/Meshroom/wiki/Reconstruction-parameters#sparse-reconstruction>

Duas capturas foram realizadas: a primeira em um dos corredores da faculdade, para validação do processo para reconstrução do prédio inteiro, com 1092 fotos, e a segunda na pintura do corredor do LCA, para seu uso no ponto de interação (Subseção 4.4.3), com 78 fotos. Apenas a segunda captura foi possível de ser reconstruída com resultados satisfatórios, tendo 74 câmeras alinhadas (94,9 %), enquanto que a primeira não teve uma quantidade suficiente de câmeras alinhadas (10, 0,9 %).

As Tabelas 2 e 3 sumarizam os resultados e parâmetros testados com as duas capturas. Observamos que especificamente alterar o detector de atributos do padrão “dsp-sift” para o “AKAZE” gerou um aumento significativo das câmeras alinhadas na segunda captura.

As Figuras 16, 17 e 18 mostram a nuvem de pontos, *mesh* e *mesh* após simplificação, respectivamente. Visualmente, o primeiro ponto a ser observado é a dificuldade de realizar a remoção da luz especular da textura, que pode ser vista desde a nuvem, gerada devido à baixa rugosidade da pintura e luz natural incidente na sala.

Consideramos que o uso de um *spray* específico para fotogrametria, capaz de minimizar os reflexos ou o uso de filtros polarizadores poderiam amenizar o efeito. Tentativas de utilizar ferramentas de *de-lighting*, tanto a *Agisoft Texture De-Lighter* [30] quanto a *De Lighting tool* [31], da própria Unity, não geraram resultados satisfatórios, obtendo apenas texturas com padrões não desejados.

Um segundo detalhe sobre a captura é o ruído e existência de buracos na *mesh*. Os buracos se concentram em regiões laterais que não tiveram tantas fotos capturadas, então possivelmente obter mais fotos dessa região reduziria o problema. Já o ruído poderia ser resolvido editando manualmente a *mesh* antes da texturização, que não foi realizado pelo trabalho devido a restrições de tempo.

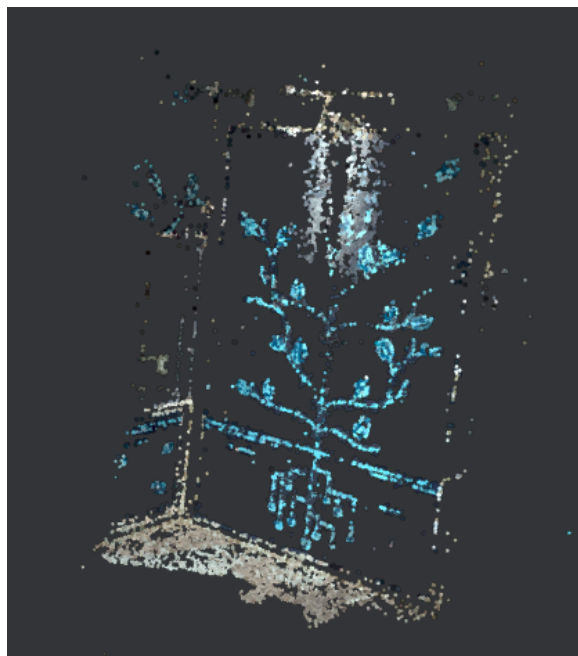


Figura 16: Nuvem de pontos obtida com a segunda captura.

Tabela 2: Resultados obtidos com a captura 1 e variação de parâmetros.

Captura	Parâmetros Alterados (Valor Original)	Câmeras Alinhadas
1	<ul style="list-style-type: none"> • FeatureExtraction.Force CPU Extraction = false (true) 	4 (0,4 %)
1	<ul style="list-style-type: none"> • FeatureExtraction.Force CPU Extraction = false (true) • StructureFromMotion.Max Angle Initial Pair = 60 (40) 	10 (0,9 %)
1	<ul style="list-style-type: none"> • FeatureExtraction.Force CPU Extraction = false (true) • FeatureExtraction.Describer Types = akaze (dpsift) • ImageMatching.Method = Exhaustive (SequentialAndVocabularyTree) • FeatureMatching.Guided Matching = true (false) • StructureFromMotion: dois nós seguidos, o primeiro com Min Input Track Length = 3 (2) e o segundo com 2 	8 (0,7 %)
1	<ul style="list-style-type: none"> • FeatureExtraction.Force CPU Extraction = false (true) • FeatureExtraction.Describer Types = akaze (dpsift) • FeatureExtraction.Describer Density = high (normal) • FeatureExtraction.Describer Quality = high (normal) • ImageMatching.Method = Exhaustive (SequentialAndVocabularyTree) • FeatureMatching.Guided Matching = true (false) • StructureFromMotion: dois nós seguidos, o primeiro com Min Input Track Length 3 (2) e o segundo com 2 (2) 	3 (0,3 %)

4.7 Conversão de Voz e Dublagem

Para a dublagem da personagem INoVi, nossa ideia inicial foi procurar alguma alternativa rápida e de baixo custo, porém com resultados expressivos. Decidimos pelo uso de um modelo conversor de voz, capaz de receber um áudio gravado por uma pessoa e convertê-lo para o timbre de outra pessoa. Dessa forma, poderíamos gravar as falas quando necessário e decidir a voz em pós produção, sem necessitar da participação de uma atriz.

Tabela 3: Resultados obtidos com a captura 2 e variação de parâmetros.

Captura	Parâmetros Alterados (Valor Original)	Câmeras Alinhadas
2	<ul style="list-style-type: none"> • FeatureExtraction.Force CPU Extraction = false (true) 	24 (30,8 %)
2	<ul style="list-style-type: none"> • FeatureExtraction.Force CPU Extraction = false (true) • FeatureExtraction.Describer Density = high (normal) • FeatureExtraction.Describer Quality = high (normal) 	26 (33,3 %)
2	<ul style="list-style-type: none"> • FeatureExtraction.Force CPU Extraction = false (true) • FeatureExtraction.Describer Density = high (normal) • FeatureExtraction.Describer Quality = high (normal) • FeatureMatching.Guided Matching = true (false) • StructureFromMotion: dois nós seguidos, o primeiro com Min Input Track Length 3 (2) e o segundo com 2 (2) 	31 (39,7 %)
2	<ul style="list-style-type: none"> • FeatureExtraction.Force CPU Extraction = false (true) • FeatureExtraction.Describer Density = high (normal) • FeatureExtraction.Describer Quality = high (normal) • FeatureExtraction.Describer Types = akaze (dpsift) • FeatureMatching.Guided Matching = true (false) • StructureFromMotion: dois nós seguidos, o primeiro com Min Input Track Length 3 (2) e o segundo com 2 (2) 	74 (94,9 %)

Utilizamos o modelo *ExpressiveVC* [32], em desenvolvimento pelo aluno de doutorado Lucas Hideki Ueda, com base no modelo para conversão de voz cantando *SoftVC VITS* [33], para conversão de voz mantendo a expressividade contida nela.

Como voz destino, escolhemos a voz da professora Paula Dornhofer Paro Costa, devido à existência de uma grande quantidade de áudios disponíveis em vídeo-aulas no Youtube. Os áudios foram extraídos utilizando a ferramenta *youtube-dl* [34], tomando o cuidado de remover áudios de aulas com a participação de outras pessoas. Os áudios foram então filtrados para remoção de ruído utilizando o modelo *denoiser* [35], e cortados em fatias, removendo as porções são fala utilizando o *Audio Slicer* [36].



Figura 17: *Mesh* obtida com a segunda captura.

A partir da base de dados, utilizamos os *scripts* do modelo para realizar a reamostragem dos áudios para a frequência necessária de 44,1 kHz e extração de atributos. O modelo generativo foi então treinado durante 81000 épocas. A Figura 19 contém uma representação do processo realizado.

Analisando algumas amostras de áudios convertidos utilizando o modelo, podemos perceber que ele é capaz de carregar algumas características expressivas da voz de origem, junto de características do timbre da voz destino, porém misturando com o timbre da voz de origem. Julgamos como não satisfatório o resultado, entretanto, não conseguimos explorar outras alternativas dentro do prazo do projeto.

Para a dublagem em si da personagem, realizamos um processo de sincronização labial (*lip sync*), ilustrado na Figura 20. Utilizamos como base um tutorial voltado para realizar sincronização dentro da Unity com avatares criados no VRoid Studio [37].

O primeiro passo do processo é realizar a extração de fonemas do áudio utilizando o *Radish Tools Phoneme Extractor* [38], parte do *Radish Modding Tools* [39], que utiliza o áudio e a transcrição da fala para gerar um arquivo com a tradução fonética. Encontramos o problema de que a ferramenta não possui um modelo fonético da língua portuguesa, e decidimos por utilizar o modelo padrão da língua inglesa, visto que os resultados finais pareceram o suficiente para os nossos objetivos, embora seja possível observar algumas falhas.

O próximo passo foi carregar o arquivo na Unity, utilizando o pacote *UniLipSync*, que



Figura 18: *Mesh* simplificada obtida com a segunda captura

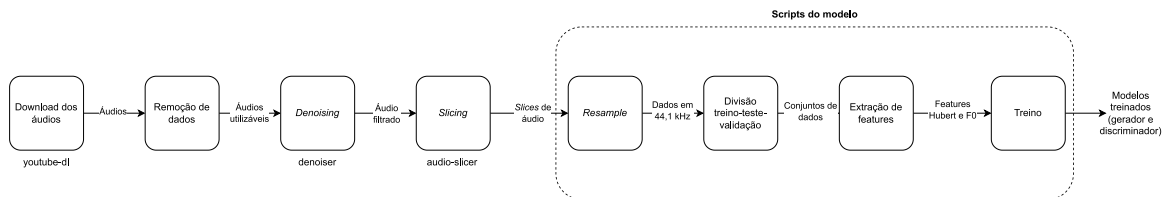


Figura 19: Fluxo utilizado para treinamento do modelo conversor de voz.

lê o arquivo e gera uma animação correspondente para a boca do avatar [40]. Por fim, a animação gerada é colocada em sincronismo com o áudio da fala, gerando a animação final.

Um exemplo do resultado do conversor de voz e sincronização labial pode ser visto em: <https://youtu.be/H94MC20UmeA>

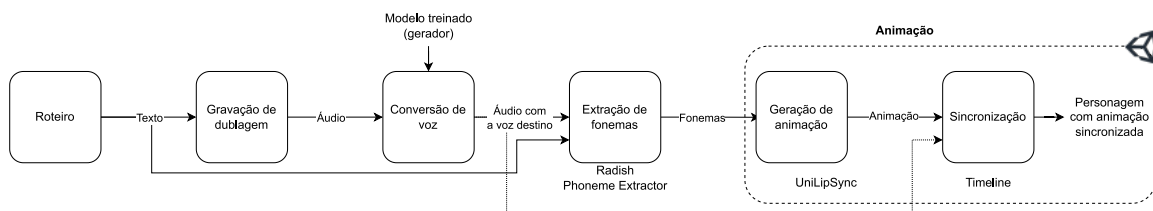


Figura 20: Fluxo utilizado para geração e sincronização da dublagem.

4.8 Exibição

A exibição proposta, aberta ao público ou restrita apenas para pessoas convidadas, não foi realizada, devido aos problemas logísticos com a aprovação do projeto de pesquisa no Comitê de Ética em Pesquisa. Consideramos a possibilidade de realizar uma pequena exibição ainda neste ano de 2023, ainda não confirmada, ou no início do ano que vem como prova de conceito e coleta de dados para desenvolvimento futuro.

Uma apresentação no XVIII Workshop de Teses, Dissertações e Trabalhos de Iniciação Científica (WTD)³ do Instituto de Computação foi realizada no dia 04 de Dezembro de 2023.

5 Conclusões

Tecnologias e fenômenos atuais como realidade estendida e jogos digitais expandem a nossa capacidade de representar e explorar o mundo. Procuramos neste trabalho realizar algumas investigações sobre essas possibilidades, focadas na mistura de objetos virtuais e físicos para apresentar reflexões sobre o espaço físico, comunidade e suas atividades. Entretanto, encontramos algumas limitações em nossos resultados devido aos problemas de planejamento interno e logísticas externas. A exposição seria essencial para obter esses resultados e possibilitar uma maior exploração conceitual.

Dentre as seis etapas propostas, duas não foram realizadas, sendo que as outras quatro foram realizadas parcialmente, dentro dos resultados mínimos esperados. Consideramos de forma geral os resultados positivos, apresentando possibilidades inovadoras dentro dos temas apresentados, com destaques para as propostas de pontos de interação com suas mecânicas em XR.

O aprendizado possibilitado nas áreas de *design*, animação, XR, fotogrametria, criação de um projeto ético e uso da Unity agregam aos resultados do projeto, embora faltaram maiores estudos nas áreas de metodologias para análise dos dados qualitativos que seriam obtidos e uso do robô humanoide.

Acreditamos que os desenvolvimentos realizados foram promissores e abriram possibilidades de exploração na área que podem ser investigadas em projetos futuros.

6 Agradecimentos

Agradecemos ao Prof. Hélio Pedrini pela orientação e acompanhamento durante este trabalho à Profa. Paula Dornhofer Paro Costa pelo empenho em nos ajudar em diversos pontos durante o projeto, ao Prof. José Mário De Martino pelo apoio e pela infraestrutura fornecida, aos alunos Lucas Hideki Ueda, pelo apoio com o modelo de conversão de voz; Rodolfo Luis Tonoli, pela ajuda com o projeto TAS; Maria Fernanda Tejada Begazo, pela apresentação sobre a robô Marta; e Daniele Souza Gonçalves, pela dublagem da personagem INoVi.

³<https://www.ic.unicamp.br/wtd/2023/>

Referências

- [1] S. Sandrone and C. Carlson, “Gamification and game-based education in neurology and neuroscience: Applications, challenges, and opportunities,” *Brain Disorders*, vol. 1, p. 100008, 2021.
- [2] G.-J. Hwang and P.-Y. Chen, “Interweaving gaming and educational technologies: Clustering and forecasting the trends of game-based learning research by bibliometric and visual analysis,” *Entertainment Computing*, vol. 40, p. 100459, 2022.
- [3] L.-H. Wang, B. Chen, G.-J. Hwang, J.-Q. Guan, and Y.-Q. Wang, “Effects of digital game-based STEM education on students’ learning achievement: a meta-analysis,” *International Journal of STEM Education*, vol. 9, no. 1, pp. 1–13, 2022.
- [4] A. Kleftodimos, G. Lappas, and M. Vrigkas, “Taleblazer vs. Metaverse: a comparative analysis of two platforms for building AR location-based educational games,” *International Journal of Entertainment Technology and Management*, vol. 1, no. 4, pp. 290–310, 2022.
- [5] M. Soler, “O campo do teatro documentário,” *Sala Preta*, vol. 13, no. 2, pp. 130–143, Dec. 2013, <https://www.revistas.usp.br/salapreta/article/view/69082>.
- [6] S. Turkle, “Video Games and Computer Holding Power,” in *The Second Self: Computers and the Human Spirit*. MIT Press, 2005, pp. 65–90, <https://ieeexplore.ieee.org/document/6284425>.
- [7] R. Hunicke, M. LeBlanc, and R. Zubek, “MDA: A Formal Approach to Game Design and Game Research,” in *AAAI Workshop on Challenges in Game AI*, vol. 4, 2004, pp. 1–5.
- [8] Open AI, “Introducing chatgpt,” 2023, <https://openai.com/blog/chatgpt>.
- [9] Ready Player Me, Inc, “Integrate a character creator into your game in days - Ready Player Me,” 2023, <https://readyplayer.me/pt-BR>.
- [10] VRoid, “VRoid Studio,” 2023, <https://vroid.com/en/studio>.
- [11] “ROBOTIS-GIT/DynamixelSDK,” Dec. 2023, <https://github.com/ROBOTIS-GIT/DynamixelSDK>.
- [12] Qualcomm Technologies, “Snapdragon Spaces Overview,” 2023, <https://docs.spaces.qualcomm.com/common/setup/QuickStartGuide.html>.
- [13] Unity Technologies, “Ar foundation,” 2023, <https://docs.unity3d.com/Packages/com.unity.xr.arfoundation@5.1/manual/index.html>.
- [14] W. Xiangchun, “The Road to ARCore-Unity Development from Introduction to Practice,” 2019.

- [15] B. Agulló, M. Montagud, and I. Fraile, “Making interaction with virtual reality accessible: rendering and guiding methods for subtitles,” *AI EDAM*, vol. 33, no. 4, pp. 416–428, 2019.
- [16] B. O. Community, *Blender - a 3D modelling and rendering package*, Blender Foundation, Stichting Blender Foundation, Amsterdam, 2018, <http://www.blender.org>.
- [17] C. Dempsey, “What is GIS,” *Gis Lounge*, 2021.
- [18] Â. Benetti, E. Poeta, and J. De Martino, “Implementação de Avatar Sinalizador de Libras Explorando Técnicas de Captura de Movimento,” *Libro de Actas*, p. 32, 2013.
- [19] E. Systems, *ESP32 Datasheet*, Technical Datasheet, Espressif Systems, 2023, https://www.espressif.com/sites/default/files/documentation/esp32_datasheet_en.pdf.
- [20] World Wide Web Consortium, “W3c protocols,” 1996, <https://www.w3.org/Protocols/>.
- [21] IBM, “AIX 7.1 Documentation - Remote Procedure Call,” 2023, <https://www.ibm.com/docs/en/aix/7.1?topic=concepts-remote-procedure-call>.
- [22] R. T. Fielding, *Architectural styles and the design of network-based software architectures*. University of California, Irvine, 2000.
- [23] Unity Technologies, “Unity Documentation - UnityWebRequest,” 2023, <https://docs.unity3d.com/Manual/UnityWebRequest.html>.
- [24] Mozilla, “Mdn web docs - http methods,” 2023, <https://developer.mozilla.org/en-US/docs/Web/HTTP/Methods>.
- [25] Unity Technologies. (2023) Unity Multiplayer Documentation - Netcode. <https://docs-multiplayer.unity3d.com/netcode/current/about/>.
- [26] C. Griwodz, S. Gasparini, L. Calvet, P. Gurdjos, F. Castan, B. Maujean, G. D. Lillo, and Y. Lanthony, “AliceVision Meshroom: An open-source 3D reconstruction pipeline,” in *12th ACM Multimedia Systems Conference*. ACM Press, 2021.
- [27] darktable team, “darktable,” <https://www.darktable.org/>.
- [28] Gleb Alexandrov, “Photoscanning 100% Free: Complete Meshroom Tutorial | Photogrammetry Course,” Jan. 2022, <https://www.youtube.com/watch?v=j3lhPKF8qjU>.
- [29] Meshroom Contributors, “Meshroom Manual - Meshroom v2021.0.1 documentation,” 2023, <https://meshroom-manual.readthedocs.io/en/latest/index.html>.
- [30] “Agisoft Texture De-Lighter - general workflow : Helpdesk Portal,” 2023, <https://agisoft.freshdesk.com/support/solutions/articles/31000158376-agisoft-texture-de-lighter-general-workflow?id=71>.

- [31] “De Lighting Tool,” Oct. 2023, <https://github.com/Unity-Technologies/DeLightingTool>.
- [32] “ExpressiveVC Expressive Voice Conversion,” Aug. 2023, <https://github.com/AI-Unicamp/ExpressiveVC>.
- [33] “SoftVC VITS Singing Voice Conversion,” Nov. 2023, <https://github.com/svc-develop-team/so-vits-svc>.
- [34] “youtube-dl,” Nov. 2023, <https://github.com/ytdl-org/youtube-dl>.
- [35] A. Defossez, G. Synnaeve, and Y. Adi, “Real Time Speech Enhancement in the Waveform Domain,” in *Interspeech*, 2020.
- [36] “Audio Slicer,” Nov. 2023, <https://github.com/openvpi/audio-slicer>.
- [37] ReForge Mode - Unity Tutorials, “FREE Lipsync Tools for Unity! - UniLipSync,” Mar. 2023, <https://www.youtube.com/watch?v=szBCZ4i4lfo>.
- [38] FFaUniHan, “Radish Tools Phoneme Extractor,” Oct. 2023, https://github.com/FFaUniHan/Radish_Tools_Phoneme_Extractor.
- [39] “radish modding tools,” Nov. 2020, <https://www.nexusmods.com/witcher3/mods/3620?tab=description>.
- [40] FFaUniHan, “UniLipSync - Free Unity Lipsync Tool,” Nov. 2023, <https://github.com/FFaUniHan/UniLipSync>.