

Workshop de Realidade Virtual e Aumentada

09 a 11 de Novembro de 2015

Livro dos Anais



Livro dos Anais

XII Workshop de Realidade Virtual e Aumentada

Organizadores dos Anais

Danilo Medeiros Eler
Ronaldo Celso Messias Correia

Capa

Caroline Mazini Rodrigues
Mario Augusto Maldonado

**CULTURA
ACADÊMICA**
Editora

Presidente Prudente
FCT/Unesp
2015

W873 Workshop de Realidade Virtual e Aumentada (12. : 2015 : Presidente Prudente)
Livro dos Anais : XII Workshop de Realidade Virtual e Aumentada : Presidente Prudente, 9 a 11 de novembro de 2015 / organizadores dos anais Danilo Medeiros Eler, Ronaldo Celso Messias Correia. - São Paulo : Cultura Acadêmica, 2015
x, 203 p.

ISBN: 978-85-7983-808-8

1. Ciência da Computação. 2. Realidade virtual e aumentada. I. Eler, Danilo Medeiros. II. Correia, Ronaldo Celso Messias. III. Título.

LIVRO DOS ANAIS

XII WORKSHOP DE REALIDADE VIRTUAL E AUMENTADA

9 a 11 de Novembro de 2015
Presidente Prudente - SP - Brasil

Coordenação Geral

Ronaldo Celso Messias Correia (FCT/UNESP)
Rogério Eduardo Garcia (FCT/UNESP)

Coordenação do Comitê de Programa

Marcelo de Paiva Guimarães (UNIFESP/Programa de Mestrado da FACCAMP)
José Remo Ferreira Brega (FC/UNESP)

Coordenação dos Minicursos

Robson Augusto Siscoutto (FIPP/UNOESTE)
Celso Olivete Junior (FCT/UNESP)

Coordenação das Sessões Técnicas

Danilo Medeiros Eler (FCT/UNESP)
Milton Shimabukuro (FCT/UNESP)

Coordenação das Sessões de Pôsteres

Almir Olivete Artero (FCT/UNESP)
Marco Antonio Piteri (FCT/UNESP)

Comitê Científico

Alécio Binotto (IBM Research)

Alexandre Cardoso (UFU)

Almir Artero (FCT/Unesp)

Andre Kawamoto (Universidade Tecnológica Federal do Paraná – Brazil)

Antonio Apolinário Jr. (Universidade Federal da Bahia)

Antonio Carlos Sementille (Universidade Estadual Paulista - UNESP)

Bianchi Meiguins (Universidade Federal do Pará)

Claudio Kirner (Universidade Federal de Itajubá – UNIFEI)

Creto Vidal (Federal University of Ceará)

Danilo Medeiros Eler (Universidade Estadual Paulista – Unesp)

Edgard Lamounier – (UFU)

Edilson Aguiar (UFES)

Eduardo Albuquerque (UFG)

Eduardo Damasceno (Universidade Tecnológica Federal do Paraná)

Eunice Nunes (Universidade Federal de Mato Grosso)

Fabiano Vaz (Universidade Federal de Pernambuco)

Felipe Breyer (Universidade Federal de Pernambuco)

Fernando Osorio (Universidade de São Paulo – USP)

Fernando Teubl (Universidade Federal do ABC – UFABC)

Francisco Simões (Federal University of Pernambuco)

Glaucio Todesco (Faculdade de Tecnologia de Itu)

Helton Biscaro (University of São Paulo – USP)

Ilana Concilio (Universidade Presbiteriana Mackenzie)

Ildeberto Rodello (FEARP – USP)

Joaquim Cavalcante-Neto (Federal University of Ceará – UFC)

João Marcelo Teixeira (UFRPE)

João Paulo Lima (Universidade Federal Rural de Pernambuco – UFRPE)

Jose Remo Brega (UNESP – Bauru)
Leonardo Cunha de Miranda (Universidade Federal do Rio Grande do Norte – UFRN)
Liliane Machado (Universidade Federal da Paraíba)
Luciano Soares (Insper)
Luis Carlos Trevelin (Universidade Federal de São Carlos)
Manuel Loaiza (Tecgraf - Depto. de Informática, PUC-Rio)
Marcelle Mota (UFPA)
Marcelo de Paiva Guimarães (Universidade Federal de São Paulo)
Marcelo Bergamaschi (Instituto Federal de Educação, Ciência e Tecnologia de São Paulo – IFSP)
Marcelo Hounsell (Universidade do Estado de Santa Catarina – UDESC)
Marcio Pinho (Pontifícia Universidade Católica do Rio Grande do Sul)
Marcos Wagner Souza Ribeiro (Universidade Federal de Goiás - Campus Jataí)
Nelson Neto (Federal University of Pará – UFPA)
Osvaldo Oliveira (Faculdade Campo Limpo Paulista)
Paula Carolei (UNIFESP)
Paulo Bressan (Universidade Federal de Alfenas)
Regina Araujo (UFSCar)
Robson Siscoutto (Universidade do Oeste Paulista)
Rodrigo Bonacin (Centro de Tecnologia da Informação Renato Archer)
Rodrigo Luis de Souza da Silva (Universidade Federal de Juiz de Fora)
Romero Tori (Centro Universitário Senac / Universidade de São Paulo)
Ronaldo Correia (Universidade Estadual Paulista – UNESP)
Selan dos Santos (Universidade Federal do Rio Grande do Norte)
Valéria Martins (Universidade Presbiteriana Mackenzie)
Valdinei Freire (Escola de Artes, Ciências e Humanidades - Universidade de São Paulo)

SUMÁRIO

ARTIGOS COMPLETOS

Virtual Simulation Applied on Human Hand Model Jussara Pimenta Matos, Renato Matosinho Martins, Paulo Marcos de Aguiar	1
Natural Interaction Tool for Data Structure Teaching Anthony Ferreira La Marca, Marcus Vinicius Brito da Silva e Windelimarcos Ferreira de Borba	6
Utilizando Sensor de Movimentos para o Ensino de Cinemática Marco A. Barbosa Filho, Thiago H. Reis, Paulo A. Bressan, Artur J. Roberto Junior	12
Desenvolvimento de um laboratório em ambiente virtual com base na teoria das inteligências múltiplas Toni Amorim, Eduardo de Paula Lima Nascimento, Norian Marranghelo, Alexandre C.R. Silva, Aledir S. Pereira	18
Serious Game to Help in Rehabilitation of Children with Disabilities using Natural Interfaces Flávia Gonçalves Fernandes, Alexandre Cardoso	24
Prototipação de uma Ferramenta de Interface Natural com Finger Tracking para Ensino Multisensorial Eduarda Maganha de Almeida, Adriane Carla Anastácio Silva, Luciano Henrique Quintino, Eduardo Filgueiras Damasceno	30
Usability Considerations for Voice Interaction in a Mobile Augmented Reality Application Tiago Araújo, Brunelli Miranda, Carlos Santos, Nikolas Carneiro, Marcelle Mota, Nelson Neto, Bianchi Serique	36
New Approaches of Feature Extraction from Fruits to Augmented Reality Systems Priscila Alves Macanhã, Wilson Estécio Marcílio Junior, Danilo Medeiros Eler, Rogério Eduardo Garcia, Ronaldo Celso Messias Correia.....	42
Construction of a 3D virtual university with recommendation of learning objects to support distance education Rodrigo Valença Cavalcante Frade, Francisco Milton Mendes Neto & Rafael Castro de Souza.....	48

A Study about the Acceptance and Use of Augmented Reality applied to Business Scenarios according to Unified Theory of Acceptance and Use of Technology	
Joao Vitor de Matos Silva, Ildeberto A. Rodello	54
Integrating Microsoft Kinect and Razer Hydra to Improve Natural Interaction in Virtual Environments	
Vanessa Avelino Xavier de Camargo, Alexandre Cardoso, Igor Andrade Moraes, Edgard Lamounier	60
Reuse of Virtual Objects from Scene Graphs for the Conception of Virtual Environments	
Camilo Barreto, Alexandre Cardoso, Edgard Lamounier, Paulo Prado, Leandro Mattioli, Alexandre Carvalho	66
Modeling Techniques and Development for Virtual Hydroelectric Power Plants Design	
Ígor Andrade Moraes, Milton Miranda Neto, Alexandre Cardoso, Edgard Lamounier, Paulo Prado	70
Scientific visualization of atmospheric electrical discharges	
José Remo Ferreira Brega, Márcia A. Zanolli Meira e Silva, Mário Popolin Neto, Robson Siscoutto	74
Desenvolvimento de Aplicações para o Oculus Rift: Integração do Oculus Rift com o Street View	
Fernanda Mendes, Valéria Farinazzo Martins, André Santos Lima, Marcelo de Paiva Guimarães	80
Sincronização de ambientes virtuais interativos por meio do grafo de cena do padrão X3D	
Alisson Fernando Coelho do Carmo, Milton Hirokazu Shimabukuro ..	86
Treasure Hunt Educational Game Based on Augmented Reality	
Hytalo Tavares, João Paulo Lima, Yvonne Carvalho	92
Face Detection: New methods based on skin detection	
Marcelo Urbano Pereira, Fernando Teubl Ferreira	98
Parallel Rendering: Methods Based on Sort-First and Sort-Last	
Marcelo Urbano Pereira, Fernando Teubl Ferreira	103
A Influência da Taxa de Quadros no Desempenho de Tarefas em Sistemas de Realidade Aumentada	
Joana S. Nascimento, Silvio R. R. Sanches.....	108

Towards Automatic Evaluation of Personality Tests	
Caio José dos Santos Brito, João Marcelo X. N. Teixeira, Veronica Teichrieb	113
Integração de Ambientes Virtuais 3D Colaborativos baseados em Aglomerados Gráficos	
Diego Roberto Colombo Dias, Luis Carlos Trevelin, Marcelo de Paiva Guimarães	119
Processamento de Linguagem Natural para a Avaliação de Situações de Emergência com Interface de Realidade Virtual	
Matheus F. Sanches, Leonardo C. Botega, Allan C. M. Oliveira.....	125
Hand-based Interaction in a 3D InfoVis Environment	
Nikolas Carneiro, Brunelli Miranda, Carlos Santos, Tiago Araújo, Anderson Marques, Bianchi Serique	132

ARTIGOS CURTOS

Muito Além do Papel: Uma Experiência de Design e Avaliação de Realidade Aumentada para Cartões Impressos	
Tatiana Aires Tavares, Herson Barreto, Lívia Barreto, Júlio Carvalho, Leandro de Assis, Mayrton Dias, Rennan Temoteo.....	138
Risks MultiProjection Use in Virtual Reality as Social Technology	
Luciene Bulhões Mattos	144
Augmented Reality Authoring Tools for Education: Sistematic Revision	
André Barone Rodrigues, Diego Roberto Colombo Dias, Valéria Farinazzo Martins, Marcelo de Paiva Guimarães.....	147
Home Automation Application using Mobile Devices and Wearable Myo	
Flávia Gonçalves Fernandes, João Ludovico Maximiano Barbosa, Milton Miranda Neto, Alexandre Cardoso	151

Um Sistema de Captura de Movimento a baixo custo para Monitoramento do Idoso Institucionalizado	
Luciano H. Quintino, Eduardo F. Damasceno, Adriane C. A. Silva, Eduarda Almeida	155
 ARFoodMenu: Augmented Reality in Food Establishments	
João Otávio Brandão Antunes de Lucena, João Paulo Silva do Monte Lima, Daniel Lima Nogueira	161
 Utilização de ambientes virtuais imersivos para o ensino	
Vagner Scamati, Diego Colombo Dias, Valéria Farinazzo Martins, Marcelo de Paiva Guimarães	165
 Ensino da Anatomia Humana Por Meio de Ambientes e Modelos Virtuais	
Fabrício Quintanilha Baptista, José Remo Ferreira Brega, Mário Popolin Neto, Silke Anna Theresa Weber	169
 Ferramenta de empacotamento de conteúdos de realidade aumentada no formato de objetos de aprendizagem	
Bruno Alves de Carvalho, Marcelo de Paiva Guimarães, Valeria Farinazzo Martins	173
 Estudo de potencialidades em ambientes de treinamento em 2D e em Realidade Virtual para o mapeamento de processos cognitivos na aprendizagem de subestações de energia	
Aline F. Furtado Silva, Alexandre Cardoso, Elise B. Mendes, Edgard Lamounier	177
 A Platform for Development of Multiple 3D Collaborative Atlas for Education on Human Anatomy	
Paulo Vinícius de Farias Paiva	183
 Tangible Touch Screen Mobile Device for Visually Impaired People: A Braille Text Entering Approach	
Fabrízzio Alphonsus A. M. N. Soares, Joyce Siqueira, Mateus M. Luna, Luciana de Oliveira Berretta, Cristiane B. R. Ferreira, Ronaldo M. da Costa, Leandro L. G. de Oliveira, Rogério Salvini	187
 Implementation Aspects of a Voice Service in Mobile Augmented Reality Applications	
Tiago Araújo, Brunelli Miranda, Carlos Santos, Nikolas Carneiro, Marissa Carvalho, Gleidson Sousa, Bianchi Serique, Nelson Neto	191

Appling Visual Domain Specific Languages for Definition of Multi-Agent Systems in Virtual Reality Applications

Daniel Abella C. M. de Souza, Fernando da Fonseca de Souza, Raphael Mendonça da Nóbrega, Fabrício da Costa Dias, Pacelly J. C. A. Xavier 196

Persistence and Data Distribution in Distributed Virtual Environments

Ronaldo Celso Messias Correia, Pedro Henrique Reis, Rogério Eduardo Garcia, Celso Olivete Júnior 200

Realização:



UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”

PROEX - Pró-Reitoria de Extensão
CLAE - Centro Local de Apoio à Extensão
FUNDUNESP - Fundação para o Desenvolvimento da UNESP



DEPARTAMENTO DE
MATEMÁTICA E COMPUTAÇÃO
FCT-UNESP

Apoio:



Laboratório de Pesquisa em Engenharia de Software Aplicada



Sociedade Brasileira
de Computação



UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”

PROEX - Pró-Reitoria de Extensão
CLAE - Centro Local de Apoio à Extensão
FUNDUNESP - Fundação para o Desenvolvimento da UNESP

Virtual Simulation Applied on Human Hand Model

Jussara Pimenta Matos, Renato Matosinho Martins, Paulo Marcos de Aguiar

Departamento de Engenharia Elétrica
Instituto Federal de Educação, Ciência e Tecnologia de São Paulo
São Paulo, SP
{jussara.pimenta@ifsp.edu.br; therenatobr@gmail.com; pmaguiar@ifsp.edu.br}

Abstract— This study presents a modeling process for the human hand movement. In recent years the use of virtual and augmented reality has increased in the biomedical area. The prostheses used to replace members for people with physical limitations require that the individual adapts to prosthetic devices, and mechanical, where an effort on the part of patients is required, especially in the early stages of their process of adaptation, and one consequence is that a significant number of these patients give up his recovery in a very short time. In order to use the prosthesis the patient has to practice how to control the artificial limb, that is particularly hard. The upper limb prosthesis simulation is a wide field of study that has many applications. Usually a virtual simulation is used in training system rehabilitation centers. The simulation complexity in this case is relevant due to the large number of joins, presenting complex aspects, that is the motivation to develop this study.

Keywords—Virtual Simulation; Human Hand Model

I. INTRODUÇÃO

Este artigo apresenta a utilização de técnicas de simulação virtual para a construção do modelo da mão humana. A principal motivação é a necessidade de construção de um modelo a partir da captura de dados de uma mão humana real, construído com base nas mãos de um voluntário e que representasse o mais próximo possível a anatomia humana, para o estudo de próteses. Durante a sua concepção, notou-se a dificuldade de se obter modelos que tivessem sido elaborados a partir de um modelo real. Após avaliação das opções disponíveis, optou-se por formular um processo no qual se obtém, a partir de um voluntário, um modelo para servir de base para variados estudos que necessitem de um modelo virtual da mão humana em diferentes posições.

A pesquisa está inserida no campo da engenharia biomédica, uma área que está cada vez mais crescendo, e em especial, utilizando o laboratório de reabilitação da instituição. Esta área tem como objetivo auxiliar as pessoas que se encontram em processo de reabilitação [1],[2], que tenham sofrido incidentes que impeçam ou dificultem de alguma forma a capacidade motora ou cognitiva. Cada vez mais existem estudos de novas formas para apoiar o tratamento de pessoas em reabilitação que utilizam o recurso da realidade virtual e aumentada [3], como ferramenta que pode auxiliar tanto em simulações complexas quanto em tratamentos usando recursos audiovisuais. Assim como, atividades de entretenimento, ou que envolvam movimentos e funções motoras.

O trabalho é uma das pesquisas realizadas no laboratório de reabilitação da instituição, onde existem equipamentos tais como, um eletromiôgrafo, aparelho capaz de captar os sinais do paciente, permitindo um controle de um possível movimento virtual, um conjunto de computadores pessoais, também está disponível um projetor, que pode ser utilizado para o paciente se ver com o auxílio adicional dos projetos na área da realidade aumentada. Além dos equipamentos citados, também está disponível uma luva de movimentos da mão, equipamento que engloba diversos sensores com a finalidade de detectar os movimentos articulados dos dedos da mão.

II. MODELAGEM DAS MÃOS

A medicina é uma das áreas que mais demandam o uso de Realidade Virtual e Aumentada (RVA) como em educação, treinamento, diagnóstico, tratamento [4] e simulação de cirurgia. A RVA, pelas suas características de visualização 3D e de interação em tempo real, permite a realização de aplicações médicas inovadoras, que antes não podiam ser realizadas [5].

Uma das grandes descobertas da biomedicina foi o sinal eletromiográfico que permite o controle de equipamentos de maneira quase que instantânea pelo cérebro de uma pessoa, se controlado de forma precisa. Tal controle é possível graças ao desenvolvimento de equipamentos que podem captar e interpretar os sinais que advém do cérebro passa para a coluna vertebral, e se introduz em um músculo por meio da junção neuromuscular, iniciando uma onda de energia que irradia do centro do músculo para suas extremidades.

Essa energia é captada na forma de tensão, e permite a diferenciação e o reconhecimento do grupo muscular que está sendo utilizado. Por essa perspectiva é indicada a escolha em utilizar esse tipo de controle em algum equipamento que será empregado para auxiliar a execução de alguma atividade física. Um dos focos da reabilitação humana é permitir o retorno do funcionamento natural das funções de movimento do corpo para o paciente que sofreu algum trauma, seja ele afetando alguma região do cérebro ou a região responsável por essa locomoção em si.

A base do corpo humano é o esqueleto, conectado através de articulações que são responsáveis por permitir ou impor restrições ao movimento do corpo. Cada articulação representa um conjunto de possibilidades de movimentos, estes movimentos são restritos em função da forma dos ossos,

ligamentos, cartilagens e fibras que fazem parte de uma determinada articulação. Na modelagem das mãos um aspecto de complexidade está no carpo, onde vários ossos se articulam na mesma região.

Na figura 1 é apresentada uma estrutura de modelagem de mão [6], utilizando uma hierarquia no movimento das falanges dos dedos.

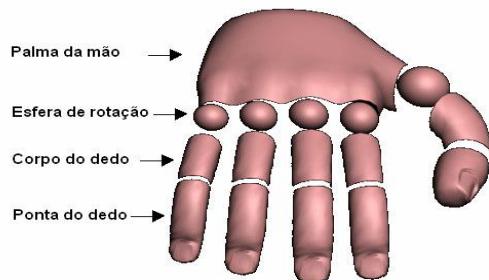


Figura 1- Modelagem da Mão [6]

Uma prótese é um dispositivo que tem como função principal a substituição de um determinado membro. A tecnologia tem avançado nas últimas décadas e é possível observar usuários de próteses participando cada vez mais de esportes e de atividades de recreação, onde são capazes de desempenhar um grande número de atividades. Porém, uma das dificuldades encontradas ao desenvolver próteses de mão é reproduzir a flexibilidade da mão humana, que deve alterar sua forma de acordo com a tarefa que está sendo executada.

A simulação dos movimentos humanos é complexa à medida que cada movimento traz uma carga de significados, seja das características genéticas do indivíduo, de suas emoções e comportamento, assim como a sua relação com ambiente no qual está inserido, até mesmo a cultura, o falar com as mãos. Além desses fatores, outro desafio é trazer realismo na representação de suas características, tal como tipo e cor da pele.

A mão, é um meio de comunicação e de manipulação, é uma ferramenta que permite vários estilos de interação, que utilizam suas funções, permitindo linguagem de gestos, postura das mãos, padrões de movimento, é uma interface de comunicação intuitiva e natural. Além da complexidade inerente da representação dos movimentos humanos, a mão tem suas próprias características que vão além de segurar objetos, ajudar a apoiar ou de pinçar. Também tem suas peculiaridades em suas articulações.

De uma forma geral, as articulações humanas são comparadas com formas geométricas e junções mecânicas, porém o movimento do ser humano é sofisticado e complexo. Um modelo para as mãos, baseado em mecânica, usando a técnica de massa-mola é apresentado em [7], cada dedo, exceto o polegar, é representado como uma cadeia serial de três falanges com quatro juntas. O polegar é modelado com duas falanges com dois graus de liberdade de movimento de uma junta, conforme apresentado na figura 2.



Figura 2: Acima a configuração inicial das mãos e abaixo a flexão de três dedos [7]

No laboratório de reabilitação da instituição foi desenvolvido um primeiro projeto, um projeto mecânico de prótese das mãos, conforme apresentado na figura 3.

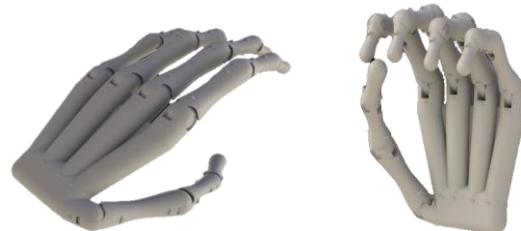


Figura 3- Projeto Mecânico de Prótese das Mãos

Pessoas que sofreram algum tipo de amputação, ou retirada de órgão em alguns poucos casos, costumam sofrer de uma síndrome chamada de órgão fantasma [8], aonde tem a sensação que essa parte que não mais pertence ao corpo continua existindo e se movendo junto com ela. Essa sensação é causada pelo fenômeno da neuroplasticidade, mais especificamente por causa da região do cérebro responsável pelo controle dessa parte.

O homúnculo motor, ou córtex motor, engloba as regiões do cérebro responsáveis pelo controle de cada região muscular. O principal efeito a ser combatido dessa síndrome é a sensação de dor intensa pela qual mais de 50% dos amputados passa a sofrer [9]. O método mais eficaz já

aplicado para isso foi criar terapias, aonde é induzida uma pequena ilusão para o cérebro, indicando que o membro amputado físico continua no mesmo lugar. Novas tecnologias podem ser adaptadas para serem utilizadas como facilitadoras dessas terapias [8].

Realidade aumentada e virtual é uma tecnologia que pode facilitar consideravelmente a execução de operações complexas [10], tornando possível ao usuário maior eficácia na transferência de informação em muitos processos e em muitos ambientes, existe uma vasta gama de utilização na reabilitação a fim de auxiliar na terapia do mesmo.

III. METODOLOGIA DE DESENVOLVIMENTO

Com intuito de tornar a simulação mais precisa, optou-se por utilizar dados reais ao invés de criados de forma artística, adquiridos da mão de uma pessoa via tecnologia de escaneamento 3D. O processo é realizado aplicando técnicas para animação de personagens.

A retopologia de modelos 3D é uma técnica para manipulação e adaptação de modelos complexos, de forma a transferir esse um modelo com grande número de polígonos para outra versão mais simples, é usada em técnicas de escultura para trabalhar a topologia de objetos orgânicos. Tem como função reorganizar as arestas do modelo, para quando a malha estiver deformada, imitar a movimentação, onde cada vetor representa um osso com orientação e comprimento.

A criação do esqueleto é uma hierarquia de vetores que tem a função de imitar a estrutura óssea, no modelo deve sofrer influência da movimentação dos ossos do esqueleto, que nada mais é do que o peso que cada vértice sofre em relação ao movimento dos ossos. O resultado final do processo, o modelo é exportado usando um formato que consiga conter as informações de animação e seja aceito pelas diferentes ferramentas de modelagem.

O realismo anatômico das articulações pode ser tratado pelo menos de duas formas, na criação de modelos genéricos de articulação do corpo humana e por representar uma articulação específica com o máximo de fidelidade possível. Para esse projeto decidiu-se utilizar um modelo real de mão humana. A metodologia de desenvolvimento esta dividida em duas etapas, a primeira utiliza o escâner e pode ser desviada se o usuário optar por dados já disponibilizados, como de um banco de dados fisiológicos, por exemplo; a segunda etapa é realizada em um CAD de modelagem 3D à escolha do usuário. O desenvolvimento consiste em aquisição, tratamento dos dados e a elaboração do modelo.

A. Aquisição

Para uso correto do escâner é necessário conhecer sua operação, para isso foram realizados alguns testes para identificar as funções necessárias para o desenvolvimento de uma aplicação, conforme orientação em [11] que explica a utilização de três escâneres, incluindo o aplicável nesse trabalho. Mesmo após o conhecimento adquirido de como operar o escâner, a realização da aquisição de um objeto

requer experimentação e preparação, no caso a montagem de uma estrutura onde a pessoa escaneada, descansa a mão a fim de evitar movimentos involuntários.

Tal procedimento é importante para que se possa obter um bom resultado [12], traços dessa estrutura podem ser notadas na figura 4. No caso do RevScan [11], necessita-se de alvos, pequenos círculos reflexivos, que o equipamento utiliza para determinar a própria orientação em relação ao objeto escaneado. Esses pontos de referências não podem se mover em relação ao objeto ou entre si, para minimizar esse problema, decidiu-se por seções de aquisição mais curtas, de no máximo de cinco minutos eram ideais, mais do que isso e se reiniciava a aquisição.

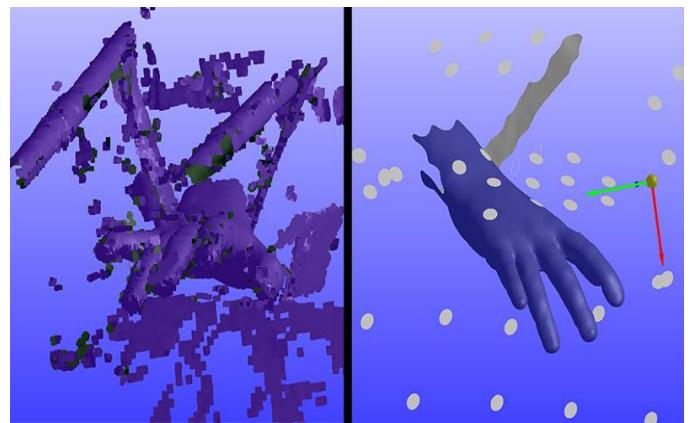


Figura 4. Comparativo dos dados adquiridos pelo escâner

Após uma aquisição satisfatória, é realizado o comando para o software tratar a superfície adquirida e então com o ajuste de filtros, conforme apresentado na figura 4, possibilitou-se corrigir erros antes de se exportar a malha para outro programa. Erros maiores decorrentes da própria montagem utilizada não puderam ser corrigidos, somente foram corrigidos aplicando-se os filtros do software.

B. Tratamento de Dados

A utilização da ferramenta [13] permite praticamente a maior parte do tratamento dos dados escaneados, que é dividido em quatro passos, isto é, a correção de erros, a retopologia, a construção do esqueleto e a identificação das influências, para que ao final deste processo, o modelo possa ser usado para animação. Todos os passos descritos são padrão da indústria de animação. Apesar de a malha ter sido baseada em dados reais, todos estes passos são realizados de forma artística. Cada modelagem 3D possui uma interface diferente e portanto, recomenda-se consultar o manual do software em uso.

a) Correção de Erro

A montagem utilizada introduziu erros na aquisição, na região do pulso, que o software do escâner não pôde corrigir. Para que se faça a correção dos erros, é necessário alterar a malha à descrição do usuário. O pulso com erros é reconstruído e algumas deformações na malha são corrigidas,

como algumas partes dos dedos que estavam mais volumosos do que deveriam. A figura 5 apresenta uma vista da mão escaneada, após a correção e é possível notar a diferença entre a figura 4 e a figura 5, porém esses reparos são baseados somente em observações da mão real.

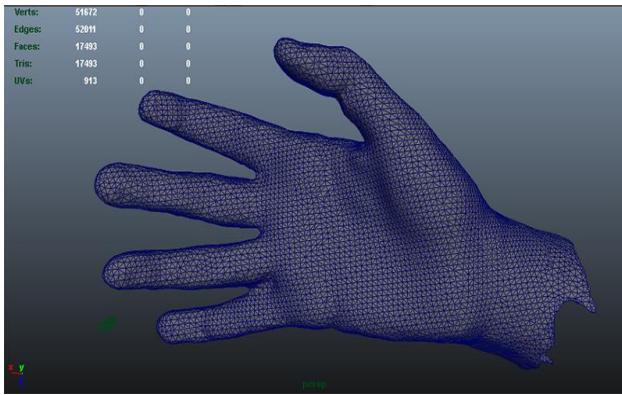


Figura 5. Modelo após ser importado e corrigido

b) Retopologia

A topologia da malha, isto é, a organização das arestas, é fundamental para que a malha deforme corretamente e imite o movimento da pele. A topologia que o escâner gera não é adequada, uma topologia adequada deve formar polígonos concêntricos na região que se contrai ou estende como as regiões coloridas, conforme apresentado na figura 6. A retopologia consiste em refazer a malha com base na superfície existente, assim a nova malha mantém a forma da antiga malha, porém, com as arestas organizadas de maneira otimizada.

C. Elaboração do Modelo

O esqueleto é uma hierarquia de vetores que representam a orientação e comprimento de cada osso, é a parte principal da animação do modelo. Desta forma, quando o modelo for exportado para outro software, cada osso pode ser controlado individualmente realizando o que é chamado de animação procedural [14]. Essa técnica não replica exatamente o movimento real. Uma forma de evitar essa falha é usar recursos que simulam músculos, mas na ferramenta [10] esse recurso é proprietário e de difícil compatibilidade com outros softwares, através de um método complexo [15].

Nesse processo foram utilizadas referências da anatomia humana para analisar a estrutura óssea da mão, porém para acomodar as falhas da técnica utilizada, essa estrutura pode sofrer alterações de acordo com a necessidade.

Com o esqueleto pronto, ainda é necessário anexar os vértices a ele, a definição de influências ou pesos cumprem essa função. O sistema de influência faz com que os movimentos de um osso influenciem a posição de determinados vértices que foram anexados a ele, cada vértice pode sofrer influência de mais de um osso, porém o sistema é multiplicativo e a soma de todas as influências em um vértice devem somar 1, assim vértices nas juntas sofrem influência

dos dois ossos desta, conforme apresentado na figura 7, que apresenta um exemplo da mão, com o esqueleto e influências prontas.

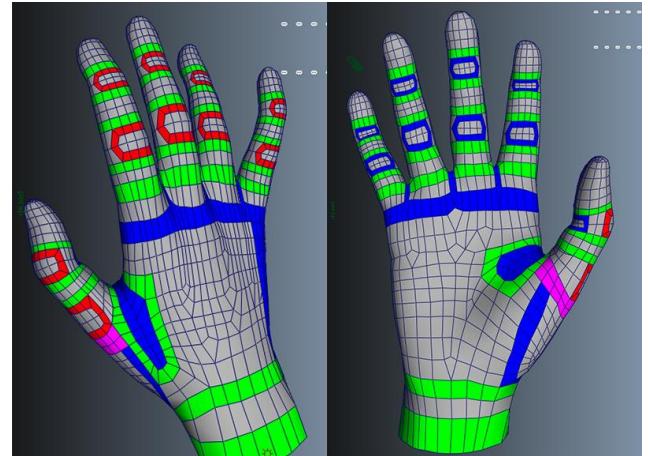


Figura 6. Mão com a retopologia com duas vistas

Tendo como premissa de que um modelo é a simplificação da realidade, uma das dificuldades é decidir que detalhes devem ser incluídos ou excluídos. O desenvolvimento do modelo apresentado na figura 7 tem como base um modelo real, com captura de dados de uma mão humana. Neste desenvolvimento se todos os detalhes fossem incluídos num primeiro momento, o modelo seria de difícil análise.

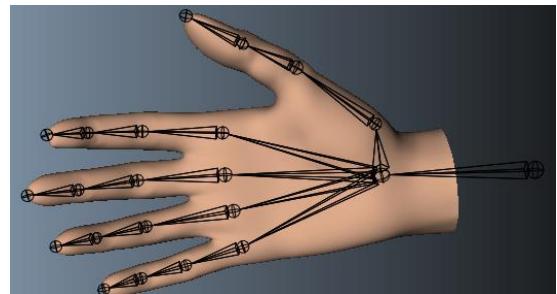


Figura 7. Modelo da mão com o esqueleto

Apesar do processo se iniciar com dados reais adquiridos pelo usuário, é possível usar dados de bancos de dados fisiológicos. E mesmo sendo utilizados dados reais, muitos detalhes artísticos são executados ao longo do processo e, portanto, não se pode afirmar que o resultado final replica com grande precisão a realidade do modelo utilizado.

IV. CONCLUSÃO

O processo não está completo, pois faltam elementos que completam a simulação, tal como a textura do modelo que necessita de mais detalhes. Porém, é possível gerar um modelo que serve de ferramenta para outros processos, como estudo de próteses [12] e realidade virtual [16]. E o estudo desse

trabalho provê uma introdução à computação gráfica e como ela pode ser usada para auxiliar na reabilitação humana.

Como apresentado no tutorial sobre adaptação de músculos para esqueleto [12], o método para obter a precisão dos músculos é complexo e exige animações pré-definidas ao invés de animação procedural, enquanto é possível fazê-lo, isso aumentaria o tempo de implementação do modelo em outros softwares, pois seria necessário que um sistema de mais complexo animação fosse implementado.

Os ambientes virtuais também possibilitam a interação física que ganham cada vez mais espaço na reabilitação. Os exergames, termo usado para jogos de videogame utilizados como forma de exercícios físicos, são ferramentas promissoras para o estímulo e incentivo de exercícios dentro da reabilitação de pessoas com alguma alteração neuromotora [17], com deficiência motora [18], com necessidades especiais [19]

Pela maior facilidade de ser encontrado e pela grande gama de aplicações, existem aplicativos que se configuram como *serious games*, isso é, uma ferramenta desenvolvida que não possui como propósito principal o entretenimento, e sim o de educar e motivar os participantes, em envolver os indivíduos atuantes [20].

A realidade aumentada e virtual também pode ser amplamente utilizada em contextos educacionais, e isso pode se aplicar às pessoas com necessidades especiais (PNEE) [21]. A ICAR (Interface de Comunicação Alternativa Remota) foi desenvolvida com o intuito de propiciar a pessoas impossibilitadas de falar, um meio alternativo para tal por intermédio da realidade aumentada [22]. Também existem aplicativos que unem essa tecnologia a musicoterapia [23] para PNEE, facilitando a utilização do mesmo para o usuário ao mostrar de maneira visual os diferentes instrumentos e reproduzindo seus respectivos sons.

AGRADECIMENTOS

Agradecemos ao Conselho Nacional de Desenvolvimento Científico e Tecnológico pelo apoio no desenvolvimento deste projeto.

REFERÊNCIAS

- [1] Nunes, F.L.S., Machado, L.S.; Moraes, R.M., Evolução da Realidade Virtual e Aumentada em Saúde: Uma Reflexão a partir de 15 Anos de SVR, XVI Symposium on Virtual and Augmented Reality, Salvador, Bahia, Brasil, 2014.
- [2] Dores, A.R., Barbosa F., Marques A., Carvalho I.P., Sousa L., Castro-Caldas A., Realidade Virtual na Reabilitação: Por Que Sim e Por Que Não? Uma Revisão Sistemática, Revista Científica da Ordem dos Médicos; Nov-Dec, 25(6):414-421, 2012
- [3] Damasceno, E.F., Lamounier Junior, E.A., Cardoso, A., Uma avaliação heurística sobre um Sistema de Captura de Movimentos em Realidade Aumentada, Journal of Health Informatics; Julho-Setembro, 4(3): 87-94, 2012.
- [4] Oliveira, I. C., Santos, S. C., Fernandes, F. G., Rodrigues, M. L., Vita, S. S. B. V., Realidade Virtual E Aumentada Aplicada Em Reabilitação Fisioterapêutica Utilizando O Sensor Kinect E Dispositivos Móveis, Conferência de Estudo em Engenharia Elétrica, Universidade Federal de Uberlândia, 2014.
- [5] Fernandes, F. G.; Paiva, L. R. M; Oliveira, L. C.; Rodrigues, M. L. A Tecnologia da Informação Aplicada À Área da Saúde: Investigações de Métodos, Técnicas e Aplicações sobre Realidade Virtual e Aumentada; Simpósio de Excelência em Gestão, 2012.
- [6] Corrêa,C.G, Nunes, F. L. S., Bezerra A. Implementação de Interação de Sistemas Virtuais para Simulação de Exames de Biópsia, Workshop de Informática em Medicina, 2008.
- [7] Sibile L., Teschner M., Srivastava S., Latombe J.C., Interactive Simulation of the Human Hand, In CARS, Proceedings...Paris, France, 2002.
- [8] Regenbrecht, H, Hoermann, S., Ott, C.; Muller, L., Franz, E.; Manipulating the Experience of Reality for Rehabilitation Applications, Proceedings of the IEEE, Volume:102, Issue: 2, 2014.
- [9] Sherman, R. A.; Sherman, C.J.; Parker, L. Chronic phantom and stump pain among American veterans: Results of a survey, Elsevier, Volume 18, Issue 1, January 1984, Pages 83–95, 1984.
- [10] Hincapé, M.; Caponio, A.; Rios, H.; Mendivil, E. G.. An Introduction to Augmented Reality with Applications in Aeronautical Maintenance, 13th International Conference on Transparent Optical Networks (ICTON), 2011.
- [11] Advanced Output Center. **Guide for 3D Scanning**. 2015. Disponível em: <http://crit.artic.edu/aoc/pdfs/aoc_3d_scan_handyscan.pdf>. Acesso em: 30 Julho 2015.
- [12] Bibb, R.J. An investigation of three-dimensional scanning of human body surfaces and its use in the design and manufacture of prostheses. 2000. Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine, 214 (6), pág. 589-594.
- [13] Maya Autodesk. User Guide. Disponível em: <http://download.autodesk.com/global/docs/maya2013/en_us/>. Acesso em: 30 Julho 2015.
- [14] Azevedo, A; Conci, A., Computação Gráfica, Rio de Janeiro: Campus, 2003.
- [15] Burke, N.; Skin Jiggle and Sliding for Game Engines. 2012. Disponível http://natbird.com/NBurke_Thesis_SkinSimulationAndDeformation.pdf Acesso em: 30 Julho 2015.
- [16] Mattioli, F.E.R, et al., Classification of EMG signals using artificial neural networks for virtual hand prosthesis control; 33º Annual International Conference of the IEEE EMBS, Massachussets, EUA, 2011.
- [17] Santos, J. V. S.; Carvalho, L. C.; Bressan, P. A. Physioplay: Um exergame ára reabilitação física aplicando a interatividade do Kinect® como biofeedback visual. IX Workshop em Realidade Virtual e Aumentada, Blumenau, 2012.
- [18] Balista, V. G., PhysioJoy- Sistema de Realidade Virtual para Avaliação e Reabilitação de Déficit Motor, Workshop on Virtual, Augmented Reality and Games, SBC Proceedings of SBGames, 2013.
- [19] Corrêa, A. G. D.; Assis, G. A.; Nascimento, M.; Lopes, R. D., GenVirtual: um jogo musical para reabilitação de indivíduos com necessidades especiais, XVIII Simpósio Brasileiro de Informática na Educação, 2007.
- [20] Faria, V. N. R.; Silva, N. L.; Abrahão, T. A.; Nogueira, K. L.; Junior, E. A. **Serious Game Utilizando Técnicas De Realidade Aumentada** Destinado À Reabilitação De Crianças Com Disabilidades Psicomotoras, Conferência de Estudo em Engenharia Elétrica, Universidade Federal de Uberlândia, 2014.
- [21] Paula, M. M. V.; Ferreira, G. G. A.; Silva, R. A., Uma análise exploratória do uso da realidade aumentada por pessoas com necessidades educativas especiais, Anais do 23º Simpósio Brasileiro de Informática na Educação (SBIE 2012), 2012.
- [22] Lima, E. D.; Forte, C. ICAR: Implementação de uma Ferramenta de Comunicação Alternativa e Remota para Pessoas com Dificuldades Motoras, 2011.
- [23] Ficheman, I. K.; Correa, A. G. D; Nascimento, M.; Lopes, R. D., O Uso Da Tecnologia De Realidade Aumentada No Apoio Ao Processo De Reabilitação Em Sessões De Musicoterapia, Revista Brasileira de Inovação Tecnológica, v.1, n.3, 2011.

Natural Interaction Tool for Data Structure Teaching

Anthony Ferreira La Marca, Marcus Vinicius Brito da Silva e Windelimarcos Ferreira de Borba.
Universidade Federal do Mato Grosso

anthony@computacao.cua.ufmt.br, marcusbrito17@yahoo.com.br, windelimarcos@gmail.com,

Resumo

Com o surgimento de estudos na área de Interação Humano-Computador, foram criadas diversas ferramentas que proporcionavam aos usuários maior facilidade de interação com as máquinas. Devido ao grande sucesso atual do setor de jogos eletrônicos, muitas dessas ferramentas são designadas a esse meio, como o caso do dispositivo Kinect. Tais ferramentas possuem grande potencial, devido as novas tecnologias associadas, tornando-as viáveis para os estudos no meio acadêmico. Desta forma, o trabalho aborda a criação de uma aplicação utilizando o dispositivo Kinect para auxiliar no ensino de estrutura de dados. A ideia da aplicação é aumentar a interação dos alunos para/com o sistema por meio de mecanismos que ofereçam a interação natural, proporcionando uma ferramenta de auxílio para o ensino, afim de minimizar os problemas de desistência e reprovação.

Abstract

With the appearance of studies in Human-Computer Interaction (HCI) area, several tools which provided users ease of interaction with the machines were created. Due to the large current success of the video game industry, many of these tools are designed for HCI, such as Kinect device. These tools have great potential due to new technologies, making them viable for studies in academia. Therefore, this paper presents the creation of an application using Kinect device to assist in teaching data structure. The idea of the application is to increase student interaction with the system via mechanisms that provide natural interaction, offering a assistance tool for teaching, in order to minimize issues such as: either flunk or drop the course.

1. Introdução

O aprendizado não é apenas em uma sala de aula com professores e livros, está em todos os momentos em que o ser humano passa por novas experiências. Em muitos casos, os métodos tradicionais não são tão eficazes, pois cada aluno possui suas próprias características, interesses, capacidades e necessidades de aprendizagem. Desta forma, faz-se necessário desenvolver um meio onde alcance o interesse de todos, para que assim, tenham um melhor aproveitamento do ensino.

Um dos grandes diferenciais observados atualmente no ensino é o de uso de tecnologias computacionais, na qual consegue-se maior atenção dos alunos por oferecerem recursos que aumentem o realismo para/com o ensino. Com o avanço tecnológico, observou-se que a interação homem-máquina aumentou significativamente, proporcionando melhor interação, imersão e navegação entre o homem e sistemas computadorizados. Neste contexto, rastreadores de movimentos “casados” com aplicações de Realidade Virtual, surgiram para trazer inovações ao mundo de pesquisa e de ensino.

Entre os rastreadores disponíveis, o Kinect lançado em 2010, proporcionava aos usuários um comportamento livre, sem a necessidade de controle, apenas utilizando o próprio movimento do corpo, diferentemente de outros lançamentos da época. O que este rastreador de movimento proporciona aos usuários, chamamos de Interação Natural, sendo um sistema capaz de entender ações naturais do ser humano e repassa-las em tempo real para o ambiente tridimensional, modificando-o [10].

Devido o potencial do dispositivo, foi desenvolvida uma ferramenta como alternativa no ensino da disciplina de estrutura de dados, visto que, atualmente, na Universidade Federal do Mato Grosso, campus Araguaia, o número de desistências e de reprovações de alunos é grande, conforme ilustra a Figura 1. Para minimizar o problema, a ferramenta

dispõe do uso de Interação Natural através do dispositivo Kinect para tornar a criação e a manipulação das estruturas mais intuitivas e interativas. Desta forma, deseja-se utilizá-la como objeto de aprendizagem para proporcionar uma alternativa na metodologia de ensino do professor.

Ano	Disciplina	Total de Matriculados	Total Reprovados
2009	ED 1	11	7
2010	ED 1	53	40
2011	ED 1	56	44
2012	ED 1	43	22
2013/1	ED 1	25	12
2013/2	ED 1	23	14

Figura 1 – Índice de Reprovação [10]

2. Interação Humano-Computador (IHC)

Interação Humano-Computador surgiu por voltar de 1980, junto com a criação dos computadores pessoais. É uma área multidisciplinar composta por psicologia, ciência da computação, linguísticas, artes, dentre outras, com o objetivo de criar uma melhor forma de comunicação entre o ser humano e computadores [11].

Com o crescimento na demanda do uso desse equipamento, surgiram novas pesquisas voltadas aos usuários e em como eles reagiam aos computadores. Foram iniciados estudos sobre interfaces gráficas mais intuitivas, dispositivos para uma interação mais rápida e precisa. Segundo Figueiredo, foi devido a esses fatores a criação da área de IHC [3].

Um dos grandes problemas da área de IHC é conseguir uma interface amigável e de fácil compreensão. Para contornar este problema, foram criados estudos em busca de novas interfaces dentro do contexto da área de IHC, chamadas de Interfaces Naturais do Usuário ou Natural User Interface (NUI), com o propósito de possibilitar ao usuário fácil controle dessas interfaces por meio de comandos que já são comuns em seu dia-a-dia.

3. Natural User Interface (NUI)

Segundo Pressman, todos os usuários já tiveram a infelicidade de encontrar interfaces de difícil compreensão e até mesmo frustrantes [9]. Para Barros, a interface pode ser comparada como a embalagem do software, devendo conter características como a de facilidade de aprendizagem, simplicidade de uso e clareza [1].

Uma boa interface ao usuário é essencial para facilitar a compreensão do ser humano das funcionalidades computacionais.

É comum que o ser humano utilize o auxílio da linguagem gestual para se comunicar, por ser uma

forma natural e simples. Desta forma, a aplicação incorporará os benefícios da NUI através do dispositivo Kinect, sendo um dos mais novos modelos para criar interfaces com interação natural, onde procura-se criar sistemas que entendem ações que são naturais aos usuários, como gestos, expressões e movimentos [11]. Segundo Liu, uma NUI oferece maior flexibilidade na interface utilizando da interação natural, diferentemente das interfaces tradicionais, na qual é necessário um nível maior de conhecimento para a sua utilização [6].

Para Valli, o conceito de Interação Natural é definido como uma forma de manipular um ambiente em que é necessária uma aprendizagem mínima para que haja interação, ou seja, qualquer pessoa teria a capacidade de utilizá-lo sem a necessidade de um treinamento rigoroso [11].

Gerar uma aplicação de Interação Natural se tornou de fácil acesso, devido aos diversos dispositivos de baixo custo existentes no mercado e na disponibilização de ferramentas para desenvolvedores dedicadas para esses dispositivos.

4. Dispositivo de Rastreamento - Kinect

Os dispositivos de rastreamentos são utilizados para acompanhar a posição do corpo e os movimentos do usuário, assim como a posição de outros objetos por ele utilizados. Após a captação destes movimentos, eles são repassados ao mundo tridimensional, modificando-o em tempo real, além de um *feedback* para o usuário.

Um destes dispositivos é o Kinect, lançado em 2010 pela Microsoft com o objetivo de mudar a forma de interação das pessoas com os jogos, sendo uma tecnologia onde as pessoas não teriam necessidade de utilizar controles físicos para jogar, mas sim, o próprio movimento do corpo.

O Kinect utiliza uma visão tridimensional para captar os movimentos e a distância dos usuários chamada de *Time-of-Flight*, que diferentemente de outras câmeras, não utiliza a estereoscópica passiva [8].

Além da tecnologia da câmera infravermelha, o Kinect possui uma câmera RGB (Red, Green and Blue) de baixa qualidade e quatro microfones utilizados para o comando de voz. Um dos principais atrativos da tecnologia do Kinect é a sua capacidade de diferenciar seres humanos dos demais objetos no ambiente [8].

O lançamento do Kinect não só teve grande repercussão nas indústrias de vídeo games (X-Box 360), como também em pesquisas acadêmicas [7]. Devido ao grande sucesso do dispositivo na área acadêmica, a Microsoft lançou uma versão específica

para o computador e disponibilizou um Software Development Kit (SDK) com diversas ferramentas prontas para auxiliar os desenvolvedores [8].

O Kinect tornou-se uma ferramenta de grande potencial para aplicações com Interação Natural em diversas áreas. Dentre os âmbitos em que o Kinect está sendo inserido, além do de jogos, há pesquisas na área de medicina, de engenharia, de segurança, e principalmente na de ensino, que é a abordagem deste trabalho [7].

5. Estrutura de Dados

O mundo computacional, desde sua criação durante a segunda guerra mundial, onde surgiram os primeiros computadores conhecidos atualmente, sofre constantes evoluções criando novas perspectivas. Como consequência dessas evoluções, a quantidade de dados a serem manipulados e armazenados teve um aumento extraordinário, tornando a eficiência de como controlá-los um fator de grande importância. Para resolver esse problema, começaram a criar novas técnicas para melhorar o aproveitamento de memória e de processamento para o controle desses dados. Essas técnicas foram chamadas de estrutura de dados, que segundo Laureano, nada mais é, do que formas de armazenamento utilizadas no mundo real adaptadas para o mundo computacional com a finalidade de otimizar a manipulação e o tratamento de dados eletronicamente [5].

Neste trabalho foram utilizadas diversas estruturas de dados: a pilha, a fila, a lista simplesmente encadeada e as árvores: binária, balanceada pela altura (AVL) e de múltiplos filhos (m-ária).

6. Descrição do Trabalho

A aplicação foi desenvolvida utilizando a linguagem de programação C# com o ambiente de desenvolvimento Visual Studio 2012, devido o SDK ter sido disponibilizado oficialmente pela fabricante (Microsoft) para este ambiente de desenvolvimento [8].

As principais tarefas do sistema são inserção e remoção de elementos nas estruturas citadas acima, além de representações gráficas e comandos via interação natural, a fim de proporcionar imersão e envolvimento com o sistema virtual.

Na interface gráfica, todo os objetos devem estar dentro da área de um componente chamado *KinectRegion*, para que possam ser controlados por meio de interações com o Kinect. Além da *KinectRegion*, o SDK fornece o *KinectScrollViewer*s e o *KinectTileButtons*, que respectivamente representam

as barras de rolagem e os botões, ambos acionados pelo dispositivo de rastreamento.

Dentro da área do *KinectRegion*, onde os objetos estão contidos, é capturado automaticamente o posicionamento de uma das mãos do usuário à frente do Kinect. Já nessa aplicação foram feitas alterações na *Dynamic-Link Library* (DLL) ou Biblioteca de Vínculo Dinâmico, responsável pelo fluxo de interação. Desta forma, é possível o rastreio das duas mãos do usuário simultaneamente.

A interação com um botão, ou seja, pressioná-lo é feita por meio do movimento de empurrar a mão em direção ao sensor Kinect, dando a impressão de estar realmente apertando um objeto. A *KinectScrollView*, ou barra de rolagem do Kinect, é manipulada por meio do fechamento da mão. Após fechar a mão é possível arrastar a barra na direção desejada.

Os movimentos utilizados para os comandos devem ser bem distintos, pois a precisão do Kinect não é perfeita, podendo ocorrer confusão para o sistema em definir uma pose ou um movimento.

Há diversas técnicas para a detecção de movimentos, dentre elas a de redes neurais, a de comparação de um gesto previamente armazenado e a de detecção por algoritmos especializados [2].

O trabalho utilizou o algoritmo especializado chamado *keyframes*. O método *keyframe* utiliza quadros-chaves, onde é definido valores a certos parâmetros que variam com o tempo em cada movimento, como posição ou cor. Por meio desses valores é calculada a evolução dos parâmetros em cada quadro de imagem em um determinado tempo [4]. Este limite entre os parâmetros do quadro inicial até o quadro no tempo final é utilizado para identificar a velocidade necessária para que o movimento seja aceito como um comando na aplicação [2].

O movimento de aceno com o braço direito é responsável pela inserção nas estruturas. Esse movimento é composto por uma lista de quadros com três poses, onde é necessário que inicialmente a mão direita do usuário esteja em uma posição maior que a posição do cotovelo direito em um eixo horizontal. Na pose seguinte, a mão e o cotovelo do usuário devem estar em uma mesma posição no eixo horizontal e a mão em uma posição superior em relação ao eixo vertical. Como última pose, a mão direita deve estar em uma posição inferior com relação ao eixo horizontal, conforme ilustra a Figura 2. Já o movimento de remoção é similar, mas deve ser executado utilizando o braço esquerdo. Com relação ao movimento para retornar ao menu principal, basta posicionar os braços à frente do corpo e baixa-los.



Figura 2 – Comando de Inserção

7. Resultado

Foi obtida uma interface simples e de fácil manuseio para o usuário, por meio de interações naturais. Efetuando movimentos diante do Kinect é possível utilizar toda a funcionalidade da aplicação, dispensando a utilização de dispositivos convencionais, como mouse e teclado. A Figura 3 ilustra a interface inicial do sistema.

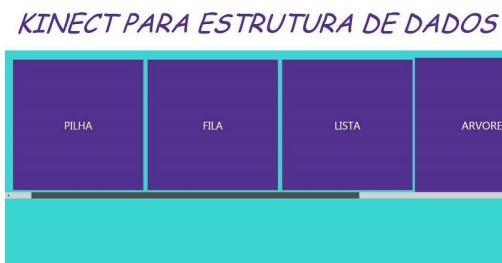


Figura 3 – Tela do Menu Principal

Na interface gráfica do menu principal é disponibilizado botões para selecionar as opções de temas disponíveis relacionados a disciplina de estrutura de dados. Dentre os temas à serem escolhidos estão a pilha, a fila, a lista e as árvores, além de estar disponível a opção de ajuda que identifica todos os comandos possíveis da aplicação e como executá-los.

Em todas as interfaces há a presença de um quadro que contém os números que podem ser inseridos nas respectivas estruturas e um frame que simula a posição do usuário em frente ao dispositivo Kinect. Todos os números são acionados via Interação Natural por meio do Kinect, bastando ir em direção ao número com a mão e realizar o movimento de “pressionar” para selecioná-lo. Após selecionar o número desejado basta pressionar “OK” e o número será renderizado na forma da estrutura escolhida.

A Figura 4 e 5 apresentam as interfaces das estruturas Pilha e Fila, respectivamente, na qual são acessadas a partir da interface principal.

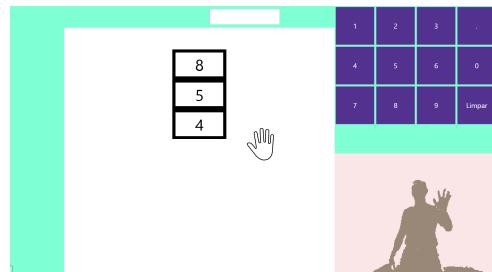


Figura 4 – Interface da Estrutura Pilha

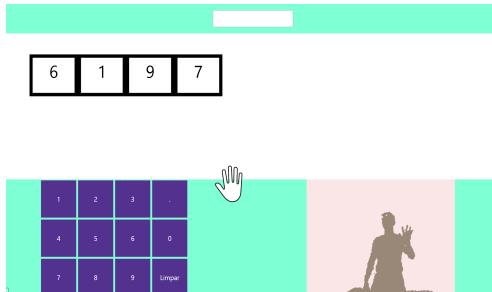


Figura 5 – Interface da Estrutura Fila

A interface da estrutura lista pode ser analisada na Figura 6. Nesta estrutura há uma representação de um ponteiro para o elemento inicial da lista, um indicando o seu elemento sucessor e um para o último elemento da lista, indicando o final da lista com um endereço nulo. Quando há uma remoção de um elemento interno, há um percurso do novo ponteiro do elemento anterior ao que será removido até o elemento posterior, além do elemento removido ficar vazio, indicando endereço nulo, conforme ilustrado na Figura 7.

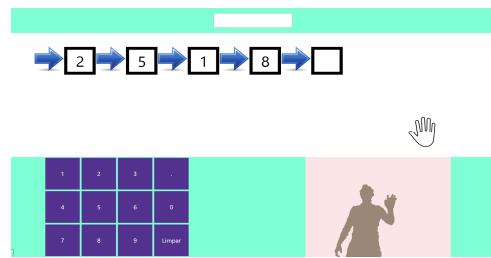


Figura 6 – Interface da Estrutura Lista

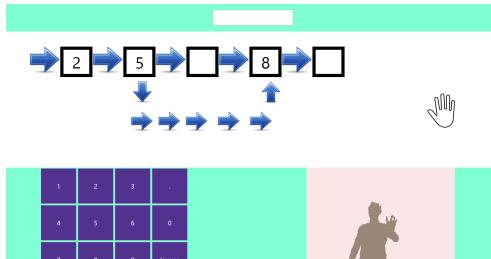


Figura 7 – Remoção de um Elemento

A última estrutura disponível na aplicação é a árvore binária e a de múltiplos filhos, tendo o posicionamento de seus elementos de forma análoga a ramificações de uma árvore, havendo um reposicionamento de acordo com o elemento a ser inserido ou removido. Esse reposicionamento de elementos pode ser vistos ao comparar a Figura 8 com a Figura 9, no caso da árvore binária e a Figura 10 com a Figura 11, no caso da árvore de múltiplos filhos, onde um novo elemento, contendo o valor 110 (cento e dez), é inserido. Com a inserção do valor, a altura da árvore é incrementada e seus elementos reposicionados.

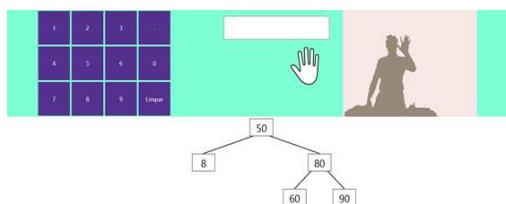


Figura 8 – Interface da Estrutura Árvore Binária

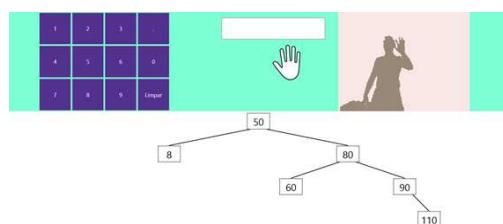


Figura 9 – Redimensionamento após uma Inserção na Árvore Binária

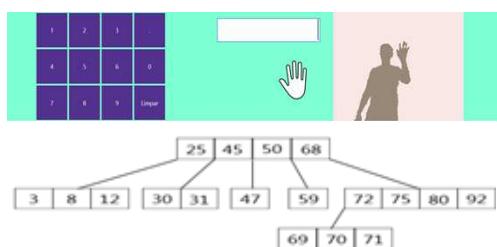


Figura 10 – Interface da Estrutura da Árvore de Múltiplos Filhos

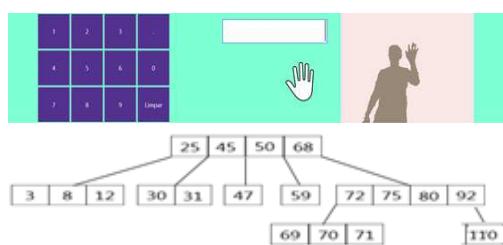


Figura 11 - Inserção na Árvore de Múltiplos Filhos

O sistema permite a remoção de elementos, este sendo nó folha, contendo apenas um filho ou dois filhos. Caso seja nó folha, o elemento é apenas removido, caso contenha elementos, tanto na subárvore esquerda quanto na subárvore direita, o sistema exclui obedecendo a propriedade de remoção de uma árvore binária de busca, ou seja, o maior elemento da subárvore esquerda ou o menor elemento da subárvore direita assume a posição do elemento a ser excluído. A Figura 12 demonstra o estado inicial da árvore binária e a Figura 13, ilustra um exemplo de remoção de elemento.

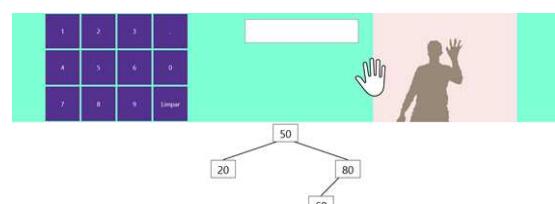


Figura 12 – Estado Inicial da Estrutura Árvore

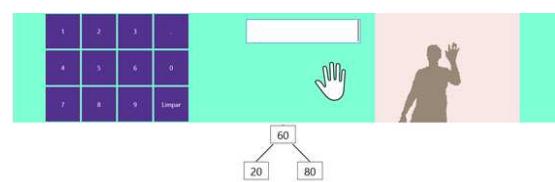


Figura 13 – Remoção do Elemento 50

Na representação da árvore AVL, houve a necessidade de mudança do *layout* para a representação dos nós, devido à aplicação de fatores de balanceamento e rotações. A Figura 14 ilustra a inserção de três elementos com o fator de balanceamento zero, indicando a não necessidade do reposicionamento dos nós.



Figura 14 – Árvore AVL

Após a inserção dos elementos 30 (trinta) e 20 (vinte), houve mudança nas cores dos elementos e dos fatores de carga, indicando o desbalanceamento da árvore, conforme ilustra a Figura 15. A cor vermelha, no fator de carga, indica que o respectivo nó precisa ser

rotacionado. A cor verde clara indica que o nó está em sua posição correta e a cor verde escuro, indica que o nó será reposicionado para a posição do nó de cor azul.

Um alerta é mostrado indicando a necessidade da rotação e qual será aplicada para a árvore manter-se balanceada, conforme ilustra a Figura 15.

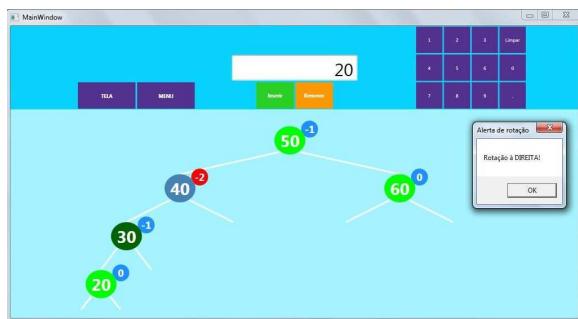


Figura 15 – Árvore AVL desbalanceada

Após o balanceamento, conforme ilustra a Figura 16, observa-se que todos os nós encontram-se verdes claros com seus respectivos fatores de carga azul.



Figura 16 – Árvore AVL balanceada

8. Conclusão

O levantamento bibliográfico foi necessário para o desenvolvimento da aplicação, pois foram apresentadas as justificativas, as estruturas e as tecnologias que proporcionaram o seu desenvolvimento.

A facilidade em compreender a partir de imagens e a possibilidade de interagir com elas, foram os principais fatores para incentivar o desenvolvimento da ferramenta. Uma vez que é possível visualizar e acompanhar a modificação nas estruturas de dados a partir de imagens gráficas, dinâmicas e interativas, em um ambiente integrado e em tempo real, torna sua compreensão mais intuitiva.

Como trabalhos futuros, novas estruturas podem ser incorporadas ao trabalho, devido a sua programação modular, além de animações para melhorar a interação em sua manipulação, principalmente na árvore AVL.

Outro trabalho importante seria aplicar a ferramenta em sala de aula, seguido de análises dos feedbacks de professores e alunos para futuras correções e melhorias.

9. Referência

- [1] Barros, Vanessa t. de o. Avaliação da interface de um aplicativo computacional através de teste de usabilidade, questionário ergonômico e análise gráfica do design. Dissertação (mestrado em engenharia de produção). Universidade Federal de Santa Catarina, Florianópolis – SC. 2003.
- [2] Cardoso, Gabriel s. Microsoft Kinect: crie aplicações interativas com o Microsoft kinect. Casa do código. São Paulo, 2013.
- [3] Figueiredo, Lucas; Souza, Pedro; Simões, Francisco; Roberto, Rafael; Teichrieb, Veronica. Interação natural a partir de rastreamento de mãos. Universidade Federal de Pernambuco. Simpósio de realidade virtual. Rio de Janeiro, 2012.
- [4] Gomide, João v. b; Pacheco, Daniel; Araújo, Arnaldo de a. Captura de movimento e animação de personagens em jogos. Universidade Fumec e Universidade Federal de Minas Gerais. Rio de Janeiro – RJ, 2009.
- [5] Laureano, Marcos. Estrutura de dados com algoritmos e c. Curitiba: Brasport livros e multimídia Ltda., 2012.
- [6] Liu, Weiyuan. Natural user interface - Next mainstream product user interface. in 11th international conference on computer-aided industrial (caidcd) 2010 ieee. p. 203 – 205.
- [7] Medeiros, Anna Carolina Soares. Interação natural baseada em gestos como interface de controle para modelos tridimensionais. Universidade Federal da Paraíba. João Pessoa, 2012.
- [8] Microsoft. Kinect for Windows sdk. Disponível em: <<http://msdn.microsoft.com/en-us/library/hh855347.aspx>>. Acesso em: julho de 2014.
- [9] Pressman, Roger S. Software Engineering: A Practitioner's Approach. 3. ed. McGraw-Hill, 1992.
- [10] Siga. Sistema de Informações de Gestão Acadêmica. Disponível em: <https://sia.ufmt.br/www-siga/htmldelphi/login/LoginUnico/login.htm>. Acesso em: agosto de 2014.
- [11] Valli, a. Notes on natural interaction. 2005.

Utilizando Sensor de Movimentos para o Ensino de Cinemática

Marco A. Barbosa Filho, Thiago H. Reis, Paulo A. Bressan, Artur J. Roberto Junior

Laboratório de Tecnologia Educacional, ICEx, UNIFAL-MG, Alfenas

{marco.antonio; a15040}@bcc.unifal-mg.edu.br, paulo.bressan@unifal-mg.edu.br, arturjustiniano@gmail.com

Resumo—Este artigo descreve um software para o ensino de Física em formato de jogo que se utiliza da temática de futebol. O jogo intitulado CineFut tem como objetivo servir de auxílio ao ambiente de sala de aula no ensino de Física, que propõe uma melhor interação professor-aluno e aluno-conteúdo. Os conceitos abordados são de lançamento parabólico, do ponto de vista da cinemática. Além disso, é utilizado o sensor de movimento Kinect para a captura dos movimentos dos usuários, permitindo assim uma maior imersão ao jogo.

Palavras-Chaves—Cinemática; Sensor de Movimentos; Jogos Sérios; Ensino de Física.

Abstract—This paper describes a software for physics teaching, game based, using a football theme. The game titled CineFut aims to aid the classroom environment on physics teaching, proposing an enhanced teacher-student and student-content interaction. The concepts approached comprise parabolic launch, from the kinematic optic. In addition, the motion sensor Kinect is used to capture the users' movements, allowing more immersion in the game.

Keywords—Kinematic, Motion Sensor; Serious Games; Physics Teaching.

I. INTRODUÇÃO

Durante muito tempo, os computadores eram encontrados apenas em grandes empresas ou universidades, isso se dava ao seu alto preço e ao grande espaço que ocupavam. Devido aos avanços tecnológicos os computadores foram sendo construídos cada vez menores e seu preço foi se tornando cada vez mais acessível às pessoas, podendo assim encontrá-lo em quase todas as casas atualmente. Junto com seu avanço e popularização o mercado de jogos se desenvolveu bastante devido a melhora dos gráficos e da jogabilidade, termo que define a capacidade de interação do usuário com o jogo, incluindo todas as experiências entre jogador e jogo [1].

Durante o avanço tecnológico na área de jogos, foram surgindo tecnologias para melhorar a maneira como as pessoas interagem com os jogos. Uma dessas ferramentas desenvolvidas foi o Kinect da Microsoft®, que é capaz de capturar os movimentos realizados pelo jogador e reproduzi-los na tela.

Neste trabalho serão discutidas as melhorias e a usabilidade do jogo educacional CineFut, desenvolvido com a utilização do sensor Kinect para o ensino da cinemática. O jogo possui a temática de futebol para o ensino de lançamento parabólico. O software será de grande apoio ao ensino da Física, pois, segundo Bruner (Moreira apud Bruner) [2], uma questão

relativa ao desenvolvimento é a representação visual dos conceitos, na qual o indivíduo desenvolve-se por meio da aquisição de formas de representar o que ocorre no seu ambiente. Além disto, ele deve ser capaz de conservar em um modelo a experiência decorrente da estimulação do ambiente e recuperar a informação por através deste mesmo modelo.

Ainda Bruner (Moreira apud Bruner) [2] diz que, se for considerado o processo intelectual do indivíduo, é possível ensinar qualquer coisa a ele, pois o desenvolvimento intelectual acontece num primeiro estágio pela representação ativa no qual o indivíduo estabelece relações entre a experiência e a ação. Após isto, vem à representação icônica, onde é possível utilizar de modelos e representações próximas à realidade, tais como gráficos e figuras, que representam conceitos. Para finalizar, as representações simbólicas apresentam a matéria na forma de proposições derivado de um sistema simbólico regido por normas ou leis, com o intuito de formar ou transformar proposições.

Neste contexto, o CineFut é uma forma de representação icônica do ambiente do aluno que permite a sua apresentação sem que houvesse a necessidade de introdução da totalidade dos conteúdos por ele abordado. Além de existir elementos de representação simbólica, as equações, na interface do jogo. Dessa forma, o professor será capaz de mostrar, por exemplo, como seria o movimento da bola ao se realizar o chute com a gravidade da Lua, mostrando a representação na tela e utilizando as equações para provar o fenômeno ocorrido.

A aplicação foi desenvolvida para ser utilizada por dois usuários, o usuário-professor determina os valores das variáveis e guiará o usuário-aluno na execução dos chutes na frente do Kinect.

Neste trabalho, inicialmente será apresentado um levantamento bibliográfico, onde foram selecionados trabalhos relacionados ao ensino e uso do Kinect em sala de aula. Os trabalhos escolhidos apresentam discussões sobre as vantagens de se usar o sensor de movimentos Kinect para o ensino, e também trabalhos que descrevem a utilização de outras tecnologias, tanto para o ensino da cinemática quanto para o uso em experimentos científicos. Em seguida, na metodologia, serão descritas as ferramentas utilizadas para o desenvolvimento da aplicação. Na descrição do jogo serão apresentados resumos de cada fase e as alterações feitas desde a última versão do software. A seguir, no teste de usabilidade, serão descritos os testes realizados com professores do ensino médio, que participam dos mestrados profissionalizantes do ensino de Física e de História e também os resultados

alcançados com os testes. Por último, a conclusão apresentará uma avaliação das informações apresentadas neste trabalho.

II. TRABALHOS CORRELATOS

Inicialmente foram pesquisados trabalhos que também fazem uso do sensor de movimentos Kinect para o ensino de Física. Como não foram encontrados trabalhos semelhantes, o levantamento foi realizado em artigos de três tipos: aqueles que discutem o uso do sensor de movimento para ensino em outras áreas, aqueles que discutem a utilidade de sensor de movimentos em sala de aula e aqueles que utilizam jogos para o ensino da cinemática.

No trabalho de Hsu [3] é apresentado o potencial do Kinect como ferramenta de ensino nas salas de aula. Este artigo sugere que o Kinect é uma ferramenta capaz de ser usada como suporte ao ensino, com grande potencial e que possibilita o aumento da interação e do interesse do estudante. Entretanto para que o Kinect seja usado de maneira eficiente é necessário que o professor passe primeiro por um treino para se adaptar a tecnologia.

O software desenvolvido por Pereira et al. [4], tem o objetivo de inserir novas tecnologias ao ensino da geometria. Em que são trabalhados conceitos de dualidade e simetria de poliedros. Desenvolvido em formato de jogo para *desktops*, o software utiliza-se do sensor de movimento Kinect para realizar a interação do usuário com as fases criadas, em que cada uma das fases tenta transmitir ao usuário determinado conceito relacionado a geometria.

Com o intuito de aumentar a imersão dos alunos de medicina e de áreas relacionadas, Matsumura e Sonnino [5] desenvolveram o Fusion4D. O software através de um computador de uso convencional e do sensor Kinect captura movimentos e comandos realizados pela voz, permitindo ao usuário manipular modelos 3D de órgãos do corpo humano.

Desenvolvida no Instituto Superior de Engenharia do Porto por Araújo [6], a aplicação Matemática Interativa tem como objetivo dar suporte ao ensino de conceitos de matemática do 3º ano do ensino básico através da captura de movimentos pelo sensor Kinect. Foi apresentado um resultado satisfatório em relação a utilização do sensor de movimento em sala de aula, aumentando a imersão e o interesse dos alunos pela disciplina.

No trabalho de Rodrigues e Carvalho [7] eles utilizam o jogo Angry Birds, desenvolvido pela Rovio Entertainment, junto com o analisador de vídeo Tracker como uma metodologia para o ensino da cinemática aos alunos de diferentes níveis. Embora a metodologia encoraje os alunos a participarem das aulas tornando-as mais interessantes, ela utiliza softwares não planejados para o ensino de cinemática.

O trabalho de Erickson et al. [8] apresenta alguns experimentos de medição de movimentos utilizando os acelerômetros e a conectividade Bluetooth dos controles remotos da Nintendo. O primeiro experimento observou o movimento harmônico com um pêndulo oscilando no plano vertical. Já o segundo experimento observou o movimento circular em uma plataforma giratória. O artigo conclui que as

ferramentas de captura de movimento são poderosas para os experimentos de Física, dando liberdade para se realizar diversas atividades tanto em laboratório quanto em ambientes de diversão infantil.

É apresentado no trabalho de Sun et. al. [9] os efeitos da utilização de jogos comerciais no ensino de conceitos de Física. Foram utilizados dois jogos para este trabalho, Cut the Rope e Angry Birds Space. O experimento realizado dividiu 83 alunos em dois grupos, um para o ensino utilizando jogos comerciais e outro com o ensino tradicional. De acordo com os participantes, houve uma melhora no aprendizado dos conceitos de Física através do jogo Cut the Rope em relação ao ensino tradicional. O mesmo não ocorreu com a utilização do jogo Angry Birds Space. Analisando os testes concluiu-se que os jogos comerciais não oferecem um avanço significativo para o aprendizado devido a sua falta de informação instrucional. Assim é necessário adicionar tutoriais e ferramentas que explicam os princípios básicos da Física para não prejudicar a transferência de experiências perceptivas.

III. METODOLOGIA

Para o desenvolvimento do software foi utilizado o ambiente de desenvolvimento Microsoft Visual Studio com a linguagem de programação CSharp. Além disso, foram utilizados três módulos que se comunicam para realizar as ações do jogo, conforme Figura 1.

A. Kinect SDK

O Kinect SDK (*Software Development Kit*) inclui *drivers* e bibliotecas para a utilização do sensor de movimento Kinect e amostras de código que facilitam a criação de aplicações. O sensor Kinect é responsável pela identificação dos usuários e captura dos seus movimentos. Os *drivers* são responsáveis pela identificação dos pontos importantes do corpo dos usuários. E a biblioteca permite a utilização desses pontos pelos projetos desenvolvidos.

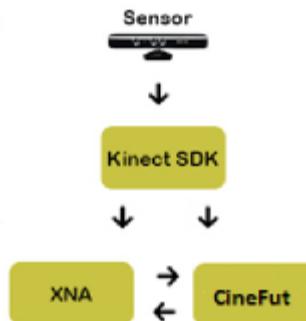


Figura 1. Comunicação entre os módulos.

B. Microsoft XNA

Microsoft XNA (*XNA's Not Acronymed*) é um *framework* utilizado para o desenvolvimento de jogos para computadores com sistema Windows, consoles Xbox 360 e para Windows Phone. Foi criado a partir de um conjunto de bibliotecas .NET para o desenvolvimento de jogos. Também foi utilizado para proporcionar a abstração de tarefas necessárias, porém difíceis de serem implementadas, como carregamento e deslocamento de imagens na tela.

C. CineFut

O aplicativo CineFut foi desenvolvido a partir dos módulos Kinect SDK e Microsoft XNA, e realiza a integração desses módulos recebendo informações do usuário pela janela de opções ou pelo Kinect. Através do reconhecimento do corpo humano temos acesso a dois pontos de interesse:

- Mão e ombros: através desses dois pontos é calculado o ângulo para serem utilizados como parâmetros das funções na aplicação.
- Pé direito: é utilizado para realizar o chute na bola e calcular a força com que ela será chutada. A força do chute é determinada pela altura em que se encontra o pé direito, quanto mais alto o pé estiver maior a força.

Na primeira versão do CineFut, a codificação foi feita apenas em uma classe, todas as condições e ações eram realizadas pela mesma classe, tornando o código confuso para quem trabalhasse. Assim, o código passou por uma refatoração, em que as fases foram divididas em classes possuindo agora várias características de programação orientada a objetos. Essas modificações permitiram um melhor entendimento da implementação e uma maior facilidade para manutenção. Atualmente, o código possui uma classe principal em que as variáveis do jogo são definidas, como por exemplo, ângulo, velocidade e gravidade. Além disso, existe uma classe com a função de atualização e desenho para cada uma das fases. Assim, quando uma fase é selecionada sua classe é chamada para realizar as atualizações de variáveis e apresenta-las na tela durante a execução do jogo.

Também foram feitas alterações na interface de opções para se adequar as exigências verificadas nos testes de usabilidade realizadas no trabalho de Reis et. al. [10]. Além disso, vários erros graves foram encontrados, o que demandou bastante esforço para a correção.

IV. DESCRIÇÃO DO JOGO

O software foi desenvolvido para ser utilizado por duas pessoas, o professor e o aluno. O professor deve escolher quais e em que ordem as fases serão chamadas, além dos valores das grandezas físicas envolvidas. Isso deve ser de acordo com o seu próprio plano de aula, sempre vinculando a fase ao conteúdo em discussão. Embora a ordem das fases possa ser escolhida, as fases do CineFut foram estabelecidas segundo um levantamento feito sobre os conteúdos curriculares do ensino de Física, como pode ser visto em Reis et. al. [10]. Já o aluno deverá realizar a tarefa proposta pelo professor e, considerando o conteúdo discutido em sala de aula, apresentar a resposta para cada fase. Isso será realizado passando valores para o ângulo e a velocidade através do sensor de movimento.

Quando inicializado, o jogo abre a janela de opções e a janela do ambiente virtual. A escolha da fase e dos valores das grandezas físicas é realizada na janela de opções, Figura 2, onde o professor (ou outro usuário) define as informações do jogo sem se posicionar na frente do sensor de movimento. Isso permite ao professor ajustar valores específicos, e também demonstrar aos alunos os movimentos realizados pela bola, em qualquer fase, sem a necessidade do sensor. A

possibilidade de demonstração sem a utilização do sensor proporciona liberdade ao professor durante o momento da aula.

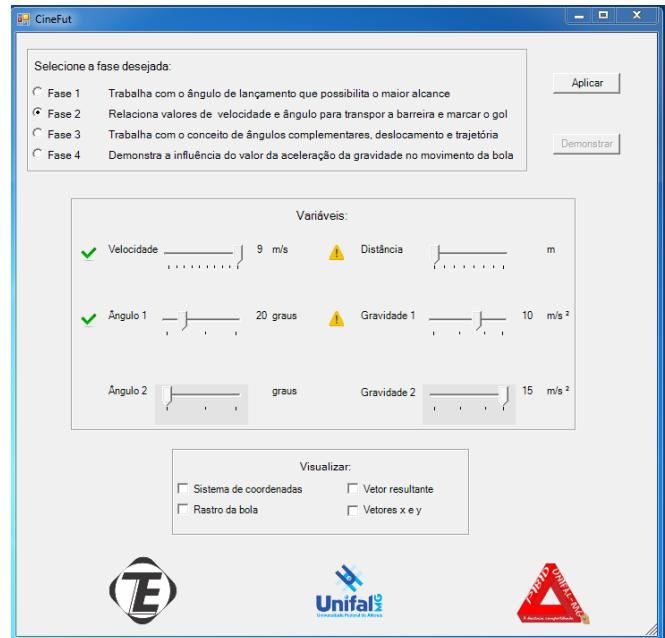


Figura 2. Janela de Opções com as variáveis Velocidade, Ângulo 1 marcadas com símbolo verde e Distância e Gravidade 1 marcadas com símbolo amarelo.

Na janela de opções, depois de selecionada a fase, as variáveis referentes a fase escolhida apresentam um ícone amarelo com uma exclamação ao lado indicando que devem ser definidos valores para elas. Se os valores forem selecionados corretamente o ícone se torna verde. Todos as variáveis referentes a fase devem estar verdes antes de se demonstrar o chute, caso isso não ocorra, a demonstração não acontece. Os valores das variáveis são selecionados através de barras de deslizamento, diferente de versões anteriores em que eram utilizados caixas de texto para a escolha dos valores.

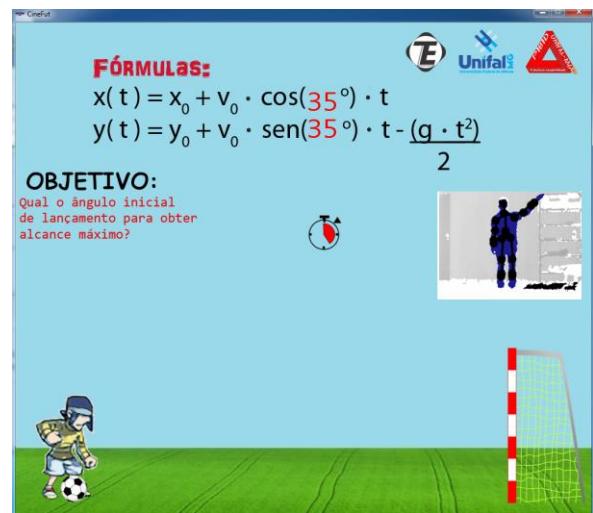


Figura 3. Tela da fase 1 durante a seleção do ângulo inicial, com o cronômetro.

A. Primeira fase

O objetivo dessa fase é realizar um chute em que ocorra o deslocamento máximo da bola. O usuário deve acertar a bola dentro do gol modificando apenas o ângulo inicial através do sensor. Com a mão direita, Figura 3, o usuário pode escolher um ângulo entre 0° e 90° que permita o alcance máximo da bola quando realizar o chute. Além disso, na janela de opções é possível alterar o valor da velocidade do chute que modifica a posição em que o personagem estará na tela, quanto menor a velocidade mais próxima do gol.

B. Segunda Fase

Diferente das outras fases, a segunda fase possui uma barreira, Figura 4. O objetivo dessa fase é fazer com que o aluno trabalhe com o ângulo inicial e velocidade inicial para transpor a barreira e marcar o gol. A barreira possui 1,75m de altura e está posicionada a 3m da bola. Na janela de opções, Figura 2, é possível modificar além do ângulo inicial e velocidade inicial, o valor da distância entre o personagem e o gol e também a gravidade, permitindo diferentes configurações do chute.

A segunda fase apresentava problemas para reconhecer o gol em determinadas situações. Este problema foi solucionado após ser realizada uma refatoração do código.

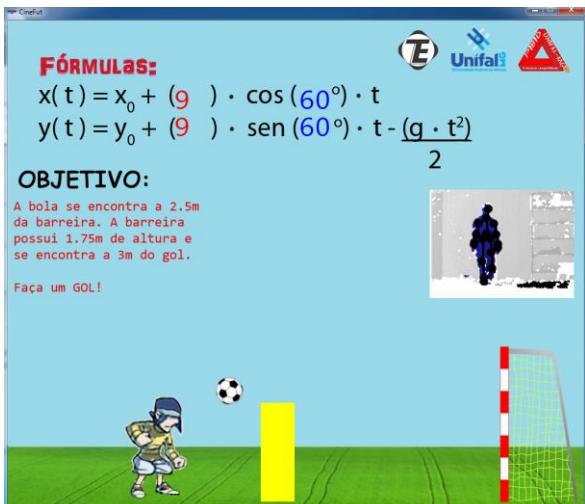


Figura 4. Fase 2 durante a realização do chute, em que é a barreira é apresentada

C. Terceira Fase

O objetivo da terceira fase, Figura 5, é trabalhar com ângulos complementares. O usuário deve realizar dois chutes com diferentes ângulos que resultem no mesmo deslocamento da bola. Primeiro é selecionado um ângulo inicial e realizado o chute, feito isso, aparecerá uma mensagem para que se realize o segundo chute com um ângulo diferente, mas que proporcione o mesmo deslocamento.

Através da janela de opções é possível modificar a gravidade, modificando os movimentos da bola. Além disso, caso o usuário não queira utilizar o sensor, pode-se através da janela de opções selecionar as variáveis utilizadas e clicar em demonstrar, assim os chutes serão realizados automaticamente.

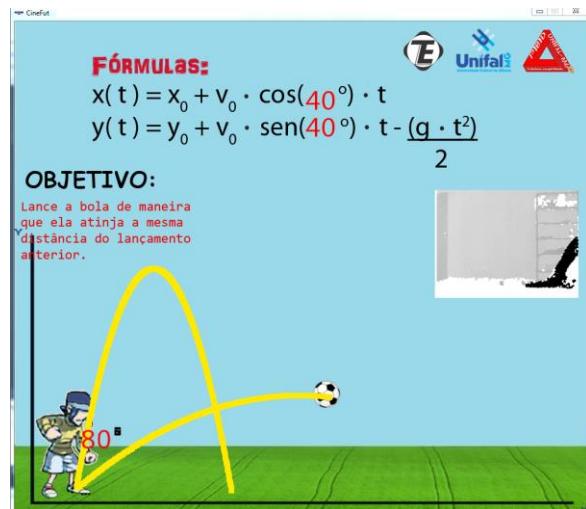


Figura 5. Fase 3 durante a realização do segundo chute com opções “Rastro da bola” e “Sistema de coordenadas” selecionadas.

D. Quarta Fase

Fase muito semelhante a terceira fase, mas com objetivo diferente. Nesta fase, Figura 6, deve-se selecionar duas gravidades diferentes para realizar o chute, com o objetivo de mostrar a influência da aceleração da gravidade no movimento da bola. O usuário deverá utilizar a janela de opções para selecionar as gravidades desejadas, caso isso não seja feito são utilizados os valores padrões definidos para a fase. Se for utilizado o botão demonstrar os dois chutes com gravidades diferentes também serão feitos de maneira automática.

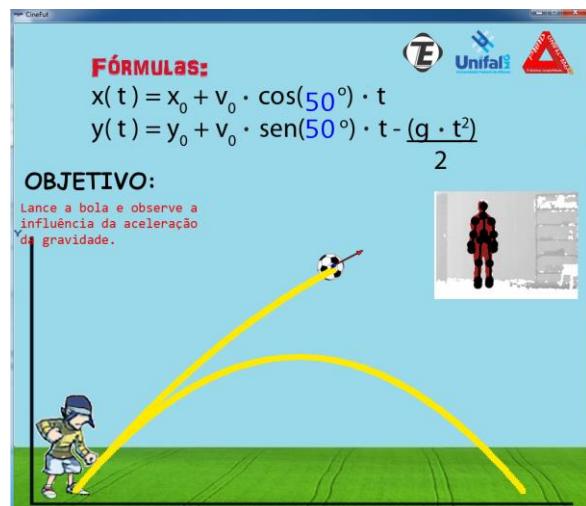


Figura 6. Fase 4 durante a realização do segundo chute com opções “Rastro da bola” e “Sistema de coordenadas” selecionadas.

Em todas as fases, o usuário deve permanecer com a mão direita na mesma posição até que um temporizador (cronômetro) apareça na tela e comece a ficar preenchido. Quando o cronômetro ficar completo o ângulo que estava selecionado é fixado e o usuário pode realizar outros movimentos sem a necessidade de manter a mão direita sempre na mesma posição.

V. TESTES DE USABILIDADE

Foram realizados testes com professores do ensino médio, que são alunos dos mestrados profissionalizantes do ensino de Física e de história. Para isso, foi criado um questionário para que o professor realizasse algumas tarefas do jogo e indicasse os seus níveis de dificuldade. As opções para cada atividade a seguir foram divididas em: muito fácil, fácil, intermediário, difícil e muito difícil.

- Selecione a velocidade em 9 m/s;
- Selecione a aceleração da gravidade em 10 m/s²;
- Selecione “Aplicar”;
- Posicione-se em frente ao Sensor Kinect e execute os movimentos para jogar;
- Selecione “Mostrar Vetor”;
- Selecione “Mostrar vetor resultante”;
- Selecione a posição da bola em 7 m;
- Selecione “Mostrar rastro”;
- Selecione a “gravidade 1” em 10 m/s²;
- Selecione a “gravidade 2” em 14 m/s²;
- Selecione “Demonstrar”.

Após realizar as atividades e marcar o grau de dificuldade, os professores deveriam marcar o grau de satisfação em relação às características abaixo da seguinte maneira: totalmente satisfeito, satisfeito, parcialmente satisfeito, insatisfeito, totalmente insatisfeito.

- Funcionalidade da janela de seleção de fases;
- Forma de selecionar os dados na janela de seleção;
- A parte gráfica do jogo;
- A utilização do Kinect (a forma de obtenção de comando);
- Forma de Abordagem dos conceitos físicos.

Ao final do questionário deixa-se a opção para que os professores possam fazer críticas sobre a aplicação.

Após realizar os testes foi feito uma análise dos resultados, apresentados a seguir em forma de gráficos:

O Gráfico 1 apresenta o nível de dificuldade dos professores em utilizarem a aplicação. Pode-se observar que a maior parte das atividades foram realizadas com facilidade. A maior dificuldade durante a realização dos testes foi em relação ao posicionamento em frente ao sensor. Alguns dos professores não possuíam familiaridade com a tecnologia e apresentaram dificuldades em realizar os movimentos para selecionar o ângulo inicial, por não ser um movimento tão natural para eles.

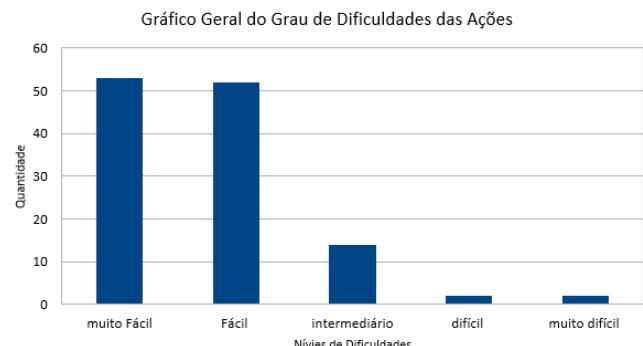


Gráfico 1. Nível de dificuldade em realizar as atividades.

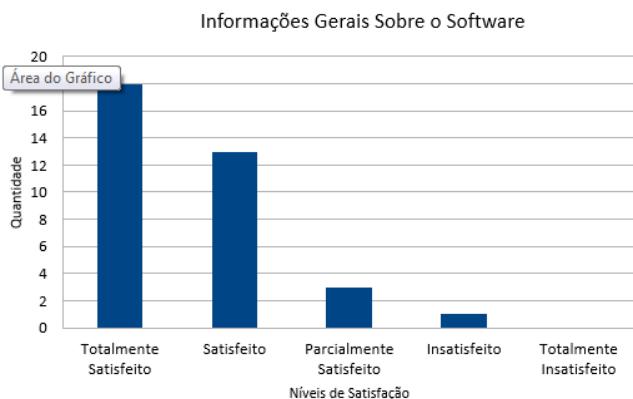


Gráfico 2. Nível de satisfação com algumas características do jogo.

O Gráfico 2 mostra o quão satisfeitos os professores ficaram com algumas características do software após realizarem os testes. Quase todos os pontos citados anteriormente foram assinalados como satisfeitos ou totalmente satisfeitos, mostrando o interesse dos usuários que realizaram os testes pela aplicação. Apenas um professor se apresentou insatisfeito com a maneira em que alguns conceitos foram abordados, pois de acordo com ele deve-se tornar o jogo mais interativo e apresentar o motivo pelo qual o aluno conseguiu alcançar o objetivo de algumas fases. E ainda que ao final das fases deveria ser apresentado um texto explicativo dos conceitos daquela fase, ou realizar perguntas ao aluno, dando liberdade ao professor de questionar os alunos.

O Gráfico 3 apresenta a frequência com que os professores utilizariam o software em suas aulas. Todos mostraram interesse em utilizá-lo para apresentar conceitos de cinemática aos alunos. No entanto, alguns se mostraram receosos em utilizar a aplicação devido à quantidade de alunos em sala de aula e o número de sensores de movimentos disponíveis.

No espaço para críticas ao jogo, os professores se mostraram na maioria dos casos satisfeitos com a aplicação. Foram feitas sugestões, como por exemplo, apresentar a janela de opções ao lado da janela principal do jogo e mostrar uma caixa de diálogo com instruções dos movimentos a serem feitos.

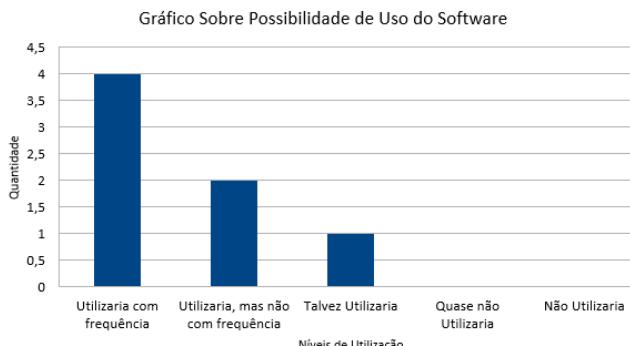


Gráfico 3. Frequência de utilização da aplicação em sala de aula.

Além das análises feitas pelos questionários, durante os testes foi possível observar alguns pontos da aplicação que devem ser abordados para a melhoria do jogo. Como por exemplo, melhorar a forma de retirar as mensagens que aparecem na tela, mudar a forma como o cálculo da posição da mão direita é realizado e também tornar o jogo ainda mais intuitivo. Tudo isso porquê alguns professores não possuem intimidade com o sensor de movimentos, ou mesmo com o computador, deixando o jogo mais difícil de ser utilizado.

VI. CONCLUSÃO

Desde sua última apresentação o CineFut passou por diversas modificações com o intuito de melhorar a forma em que os conceitos são transmitidos aos alunos e também para permitir uma melhorar usabilidade do jogo, facilitando para o professor a sua utilização em sala de aula.

As principais alterações se encontraram na janela de opções que anteriormente era muito confusa para o usuário que iria utilizá-la pela primeira vez. Após as alterações a janela se tornou mais intuitiva e de mais fácil utilização para usuários com pouca experiência. Além disso, foram adicionadas novas funções nas fases, como o travamento do ângulo depois de selecionado, dando maior liberdade ao usuário.

Através dos testes realizados com os professores do ensino médio foi possível observar diversos pontos que devem ser abordados para que o software funcione de maneira mais adequada para os professores e alunos. Principalmente na parte de interação com os usuários, que se mostrou um pouco confusa para usuários que estão utilizando a tecnologia pela primeira vez.

Mesmo com as críticas feitas, os professores se mostraram muito interessados em utilizar o jogo em sala de aula para o ensino da cinematografia.

Portanto, mesmo sendo necessárias melhorias na aplicação, ela já pode ser utilizada para o ensino dos conceitos de Física.

Para trabalhos futuros serão realizados testes com os alunos do ensino médio junto aos professores, e estes poderão opinar sobre o funcionamento e a utilização do mesmo em sala de aula.

REFERÊNCIAS

- [1] H. Vannuchi, G. Prado. “Discutindo o Conceito de Gameplay”, Portal de Periódicos Universidade Federal de Santa Catarina. Volume 5. Número 2. Página 130. 2009.
- [2] M. A. Moreira. Teorias de aprendizagem. 1º. ed. São Paulo: EPU, 1999. 195 p. ISBN 85-12-32140-7.
- [3] H. J. Hsu. “The Potential of Kinect in Education,” International Journal of Information and Education Technology vol. 1, no. 5, pp. 365-370, 2011.
- [4] D. C. Pereira, J. V. L. Pinheiro, C. E. V. Vaz. “I Love Edros – Um Game Educacional Para o Ensino Da Geometria,” Universidade Federal de Santa Catarina. 2013.
- [5] K. K. Matsumura, R. Sonnino. “Fusion4D – Interface Natural e Imersiva Para Manipulação de Objetos 3D”. 2011. 109f. Tese (Graduação em Engenharia de Computação) – Departamento de Engenharia de Computação e Sistemas Digitais, Escola Politécnica da Universidade de São Paulo, São Paulo.
- [6] H. S. Nunes Araújo. “Utilização do Kinect no Apoio ao Ensino/Aprendizagem dos Fundamentos da Matemática.”. 2014. 95f. Dissertação (Mestrado em Engenharia Informática) – Instituto Superior de Engenharia do Porto, Porto.
- [7] M. Rodrigues, P. C. Simeão. “Teaching physics with Angry Birds: exploring the kinematics and dynamics of the game,” Phys. Educ. 48 431, 2013.
- [8] M. Erickson, R. Ochoa, C. Ochoa. “The Wiimote on the Playground,” The Physics Teacher, Volume 51, Issue 5, pp. 272-275 (2013).
- [9] C. Sun, S. Ye, Y. Wang, “Effects of commercial video games on cognitive elaboration of physical concepts,” Computers & Education, Volume 88, Pages 169–181, October 2015.
- [10] T. H. Reis, G. K. V. Bichara, P. A. Bressan, A. J. Roberto Junior. “Ensino Conceitos de Física com Sensores de Movimentos,” XIII Simpósio Brasileiro de Jogos e Entretenimento Digital, 2014, Porto Alegre. XIII Simpósio Brasileiro de Jogos e Entretenimento Digital, 2014. p. 597-600.

Desenvolvimento de um laboratório em ambiente virtual com base na teoria das inteligências múltiplas.

Toni Amorim

Programa de Pós-Graduação em Engenharia Elétrica
Universidade Estadual Paulista "Júlio de Mesquita Filho", Ilha Solteira, SP,
toniamorim@gmail.com

Eduardo de Paula Lima Nascimento

Programa de Pós-Graduação em Computação
Universidade Estadual Paulista "Júlio de Mesquita Filho", S.J. do Rio Preto, SP
eduardonascimento@sjrp.unesp.br

Norian Marranghello

Abstract— Virtual labs are interactive development environments to create and conduct simulated and controlled experiments whose results depend on the user-defined input data. This article describes the development of a virtual laboratory in 3D environment using the OpenSim software, integrated with Moodle. The activities proposed for testing in the 3D environment have been developed based on the theory of multiple intelligences (TIM), combined with information and communication technologies (ICT). The possibility of involving the application of ICTs TIM aims to enable greater scope to the Theory of Multiple Intelligences, specifically in the field of education, from the availability of tools that assist in stimulating one or more multiple intelligences. This paper also presents the results obtained from the use of laboratory for students of Unesp and Unemat as well as the activities to be performed in later research.

Keywords— 3D Virtual Environments, Learning Management Systems, OpenSim, Sloodle.

I. INTRODUCÃO

Laboratórios virtuais são ambientes de simulação que proporcionam ao aluno executar experimentos em condições pré-estabelecidas. Os laboratórios podem ser acessados em uma rede local ou pela rede mundial de computadores.

O desenvolvimento de laboratórios virtuais caracteriza-se como uma alternativa viável no processo de formação e aperfeiçoamento de profissionais em várias áreas. Dentre as tecnologias que podem ser utilizadas para o desenvolvimento desses laboratórios encontram-se os ambientes virtuais 3D, também conhecidos como mundos 3D (MV3D).

Os mundos virtuais 3D (MV3D) são ambientes interativos capazes de simular algumas características do mundo real, tais como gravidade, topografia e locomoção, ações em tempo real e comunicação interpessoal.

O uso de ambientes virtuais 3D na educação tem sido demonstrado em pesquisas como as de Ritzema e Harris [18], Baldi e Lopes [3], Ridgewell *et al.* [17], mas há algumas perguntas que precisam ser respondidas, em especial as que abordam a metodologia de ensino e avaliação dentro desse ambiente.

Universidade Estadual Paulista "Júlio de Mesquita

Filho", S.J. do Rio Preto, SP

norian@ibilce.unesp.br

Alexandre C.R. Silva

Universidade Estadual Paulista "Júlio de Mesquita

Filho", Ilha Solteira, SP

acrsilva@dee.feis.unesp.br

Aledir S. Pereira

Universidade Estadual Paulista "Júlio de Mesquita

Filho", S.J. do Rio Preto, SP

aledir@ibilce.unesp.br

Uma alternativa pedagógica para essas questões é a teoria das inteligências múltiplas, proposta por Howard Gardner [8]. A união da tecnologia de laboratórios virtuais à teoria proposta por Gardner pode propiciar uma forma diferenciada de apoio a atividades desenvolvidas em sala de aula.

Na pesquisa relatada neste artigo, descreve-se a criação de um laboratório virtual em ambiente 3D utilizando-se o software OpenSim [14]. Neste laboratório virtual, busca-se associar a tecnologia dos mundos virtuais 3D à teoria das inteligências múltiplas, visando desenvolver ou aprimorar nos alunos, algumas das inteligências definidas por Gardner [7,8] dentre elas a Lógico-matemática, Lingüística, Espacial, Corporal-cinestésica, Interpessoal, Intrapessoal e a Musical.

A pesquisa descrita neste artigo tem como foco investigar a Teoria das Inteligências Múltiplas de Howard Gardner e propor uma metodologia de ensino que possa contribuir para ampliar as possibilidades de aprendizado a partir da oferta de conteúdos elaborados com base na teoria das inteligências múltiplas, utilizando as tecnologias da informação e comunicação em um ambiente virtual 3D.

Assim, foi desenvolvida uma metodologia de ensino, baseada na teoria das inteligências múltiplas, que engloba o uso de AV3D por meio do desenvolvimento de um laboratório virtual, associado a um sistema de gerenciamento de aprendizagem, especificamente o MOODLE.

Tal metodologia foi desenvolvida através da aplicação da TIM como forma de auxiliar esse aprendizado no âmbito da disciplina de sistemas digitais, por meio de atividades realizadas no âmbito de um ambiente virtual.

Este artigo está estruturado em 7 seções e está organizado da seguinte forma: Na seção 1, apresenta-se a introdução, os objetivos gerais e específicos, bem como a metodologia utilizada e o escopo da pesquisa.

Na seção 2, descrevem-se os laboratórios virtuais e o gerenciamento de acesso por meio do LMS (Learning Management System). Na seção 3 descreve-se os metaversos, em especial o OpenSim, e como é feita a integração do OpenSim ao Moodle, por meio do uso do plugin Sloodle, relatando-se ainda um breve estado da arte e a Teoria das Inteligências Múltiplas.

Na seção 4, descreve-se o processo de desenvolvimento do laboratório, abordando-se as ferramentas pedagógicas que o compõem.

Na seção 5, são descritos os testes e implementados com base nos conteúdos abordados. Na seção 6, apresentam-se os resultados obtidos na fase de testes realizados.

Na seção 7 apresentam-se as considerações finais e as atividades futuras a serem executadas.

II. LABORATÓRIOS VIRTUAIS

Laboratórios Virtuais são espaços de trabalho e colaboração a distância usados para experimentação e pesquisa ou outra atividade criativa que visem gerar e entregar resultados usando a tecnologia de informação e comunicação [9].

Esses ambientes de simulação remotos propiciam um espaço de experimentação no qual não há preocupações de segurança relacionadas a equipamentos reais, podendo este ser acessado via web ou por estações locais.

A funcionalidade de laboratórios virtuais é frequentemente reforçada por cursos *on-line* e instruções passo-a-passo para acompanhar os trabalhos laboratoriais [9].

O controle de acesso a estes laboratórios pode ser feito por LMS (*Learning Management System*), cujo objetivo é realizar o acompanhamento e registro das atividades desenvolvidas pelo aluno nesse ambiente.

Um dos sistemas mais utilizados para esse fim é o Moodle (*Modular Object-Oriented Dynamic Learning Environment*). O MOODLE é uma plataforma gratuita e de código aberto desenvolvida por Martin Dougiamas, que possibilita acompanhar o progresso do aluno em cursos presenciais ou semipresenciais [6,11,12].

III. MUNDOS VIRTUAIS 3D

Os mundos virtuais 3D ou metaversos são ambientes de imersão que propiciam realizar simulações características de um ambiente real. Os mundos virtuais mais conhecidos são o **Second Life**, o **Active Worlds**, o **There** e o **Opensim** (OS).

O OpenSim é um *software* de código aberto mantido por uma comunidade de desenvolvedores e possui as mesmas características do Second Life (SL) [14], apesar de possuir algumas limitações quando comparado com o SL.

Porém, o fato de ser um *software* de código aberto que não impõe limites à criação de objetos ou de regiões, facilita o seu uso para fins educacionais.

O SL, apesar de permitir a inscrição e a interação do usuário de maneira gratuita, a criação de objetos permanentes e grupos é limitada e muitas ações executadas nesse ambiente requerem pagamento, o que limita a possibilidade de uso da ferramenta na educação.

Por sua vez, o OS oferece compatibilidade com SL e protocolos de comunicação para outras plataformas, sendo utilizado por universidades e grandes empresas como a IBM e a Microsoft.

A integração do ambiente virtual 3D ao LMS (*Learning Management System*) pode ser realizada com a utilização do

plugin SLOODLE (*Simulation Linked Object Oriented Dynamic Learning*), cujo objetivo é disponibilizar aos usuários ferramentas que proporcionam a realização de tarefas dentro do ambiente virtual, registrando essas ações no Moodle [11], [21].

A - Laboratórios virtuais e metaverso: estado da arte

O desenvolvimento de ambientes virtuais de ensino utilizando ambientes 3D é objeto de pesquisa nas mais diferentes áreas. A aplicação dessa tecnologia permite criar ambientes como laboratórios virtuais restritos a um único prédio virtual até *campi* inteiros.

Guo *et al* [10] desenvolveram o *campus* e os laboratórios virtuais da universidade de Jiliang China, para tal foi combinada a estrutura de rede física do campus com o ambiente virtual.

O sistema utiliza técnica de *blended learning*, metodologia de formação onde a maior parte dos conteúdos é transmitido em curso a distância, normalmente pela internet, entretanto inclui necessariamente situações presenciais [2], para realizar a combinação de recursos disponíveis no ambiente real e virtual.

A idéia básica da rede do sistema experimento virtual baseada em campus virtual é combinar as funções dos laboratórios virtuais com as armazenadas no servidor de rede do campus real.

O objetivo foi proporcionar aos alunos um ambiente virtual semelhante ao ambiente físico, onde os alunos pudessem executar experiências com base nos conhecimentos adquiridos na disciplina de redes.

O laboratório virtual desenvolvido por Tarouco *et al* [19], utilizando o *software* OpenSim, tem por objetivo o ensino da disciplina de cálculo nos cursos de engenharia. O laboratório foi desenvolvido em 4 estágios que têm início com a escolha de métodos teóricos para o ensino de cálculo, e perpassa pela elaboração de objetos de aprendizagem para o ambiente virtual.

Rico et al [16] desenvolveram uma arquitetura composta por um conjunto de laboratórios virtuais que contemplam biotecnologia, química, eletrônica e operação de uma usina nuclear.

Cada laboratório representa um módulo dentro do ambiente e é operado de acordo com o nível de conhecimento de cada usuário.

B - O Sloodle

Para a integração do ambiente virtual 3D (AV3D) ao Moodle é utilizado o SLOODLE (*Simulation Linked Object Oriented Dynamic Learning*) [11], que disponibiliza aos usuários ferramentas para a realização de tarefas dentro do ambiente virtual, registrando essas ações na plataforma Moodle [21].

O Sloodle possui as seguintes funcionalidades [11]: Web-intercom; Registration booth; MetaGloss; Quiz tool; 3D Drop Box; Multi-function Toolbar; Presenter; QuizChair.

C - A Teoria das Inteligências Múltiplas

A Teoria das Inteligências Múltiplas desenvolvida por Howard Gardner tem por objetivo apresentar um novo conceito de inteligência, que até então se restringia aos

resultados de testes de QI (Quociente de Inteligência) criados em 1920 por Alfred Binet [3].

Para Gardner, a inteligência é a habilidade de resolver problemas ou criar produtos que sejam significativos em um ou mais ambientes culturais sendo a inteligência uma combinação entre a herança genética e as influências desses ambientes [8].

Em sua teoria, Gardner sugere a existência de nove inteligências, sendo elas [9]: Inteligência Lingüística, Musical, Lógico-matemática, Espacial; Corporal-cinestésica, Intrapessoal, Interpessoal; Naturalista e Existencial.

Na pesquisa relatada neste artigo, são estudadas apenas 7, pois as inteligências existencial e naturalista não se demonstram inviáveis de associar a TICs.

IV. O DESENVOLVIMENTO DO LABORATÓRIO

No desenvolvimento do ambiente virtual 3D associam-se tecnologias de informação e comunicação (TICs) com metodologias de ensino propostas por Piaget e Gardner visando uma melhor assimilação de conteúdos ministrados pelo professor, referentes a sistemas digitais.

A proposta didática deste laboratório envolve o ensino de conteúdos de uma disciplina de sistemas digitais, através da execução de atividades, dentro do ambiente virtual, que são desenvolvidas com base nas sete inteligências definidas por Gardner.

Os conteúdos a serem abordados na disciplina são [5]: Mapa de karnaugh, Introdução aos sistemas discretos e contínuos, Definição de sistemas binários e digitais, Representações de sinais binários, Diagrama de estados, Operações lógicas, Portas lógicas básicas (E, OU e NÃO), Conexão de portas; Tabela verdade.

Dada a integração do ambiente virtual à plataforma MOODLE as tarefas poderão ser realizadas com o uso de objetos disponibilizados no ambiente 3D ou quando o uso desse objeto não for possível, a atividade poderá ser realizada dentro da plataforma.

Cada inteligência está associada à execução de uma ou mais atividades, assim a inteligência linguística é estimulada através da participação do aluno em fóruns, *chats* e pela leitura e interpretação das avaliações a serem realizadas por meio de atividades relacionadas a perguntas de múltipla escolha.

A inteligência lógico-matemática é estimulada através da construção de circuitos representados dentro do ambiente virtual. A construção desses objetos visa também propiciar o uso de sua inteligência corporal cinestésica uma vez que exige do acadêmico habilidade quanto ao uso do *mouse*.

A inteligência musical é estimulada pelo uso de vídeos, *chat* de voz e outras formas de sons que podem ser importadas para o ambiente de acordo com a necessidade de cada tarefa.

As inteligências visual espacial, e a corporal-cinestésica são estimuladas também através da realização de atividades que envolvam a criação de objetos no ambiente virtual, uma vez que para a execução dessas atividades o aluno deve ter uma percepção visual para dimensionar e posicionar os objetos.

Outro fator que auxilia na estimulação dessa inteligência é o fato do usuário ter que se locomover dentro do ambiente, o que pode ser feito de três formas: andando, correndo ou voando.

As inteligências interpessoal e intrapessoal serão estimuladas de maneira conjunta e são estimuladas por atividades que devem ser realizadas em grupo disponibilizadas tanto no ambiente virtual quanto na plataforma MOODLE.

O laboratório possui 3 ambientes, sendo eles, 1 sala onde serão ministrados conteúdos teóricos via chat e texto, e uso de objetos 3D e 2 salas para a realização de testes elaborados com base nos conteúdos apresentados aos alunos. Todos os ambientes podem ser configurados de acordo com o perfil da atividade a ser realizada.

A Figura 1 apresenta a imagem frontal do laboratório, onde estão dispostos 2 elementos virtuais que gerenciam o acesso do acadêmico ao espaço interno.



Figura 1. Imagem frontal do laboratório

O token de acesso (Figura 1) representado pela imagem de uma cabine, quando acionado pelo aluno, direciona-o para a tela de login do MOODLE. O segundo elemento representa o SLOODLE.

O MOODLE foi instalado e configurado para que pudesse rodar o *plugin* do SLOODLE, conforme demonstrado na Figura 3. Os objetos de aprendizagem do SLOODLE são ligados ao ambiente virtual 3D através de um recurso denominado *SLOODLE Controller* [13], que é vinculado a um curso que deve ser criado na plataforma. Esse recurso ativa as funcionalidades descritas na seção III item B.

Por sua vez a Figura 3 demonstra a integração do laboratório com o MOODLE.



Figura 2. tela de acesso do MOODLE

Após acessar o laboratório virtual o aluno deve realizar seu *login* também na plataforma MOODLE, uma vez que o acadêmico pode ter nomes diferentes, para o seu avatar e para seu usuário na plataforma. Quando o avatar toca o objeto virtual, abre-se dentro do ambiente virtual a página de *login* do MOODLE, conforme Figura 2.



Figura 3. Cenário de integração do Sloodle

Os objetos de aprendizagem do SLOODLE foram adicionados ao ambiente 3D do OpenSim através da função denominada cenário, que deve ser adicionada pelo administrador do OpenSim.

Esse recurso ativa as funcionalidades descritas na seção funcionalidades do SLOODLE, presente na subseção 3.1. O laboratório foi configurado de modo que os alunos possam efetuar o login ao sistema MOODLE [13].

O cenário é constituído de módulos nos quais os objetos são criados através da função denominada populate que está vinculado à plataforma Moodle por meio da função denominada *sloodle controller*. No laboratório desenvolvido, cada módulo se refere a uma disciplina.

Os objetos, após terem sido criados no cenário, devem ser renderizados para que possam ser visualizados no ambiente 3D. Essa ação é realizada a partir do botão que possui a grafia *Rezz*, que após ser selecionado altera seu texto para *Rezzed*, o que indica que o objeto está pronto para ser utilizado no ambiente 3D.

A relação entre objetos que poderão ser utilizados no ambiente 3D, integrados ao MOODLE através do Sloodle são apresentados na Figura 4.



Figura 4. Relação dos objetos de aprendizagem integrados. Foram desenvolvidas algumas atividades a serem executadas pelos alunos, sendo elas distribuídas em diferentes ambientes do laboratório. A Figura 5 representa

um dos ambientes do laboratório, o qual foi dimensionado equipado como uma sala de aula real, com elementos dispostos de modo a garantir acesso a algumas das atividades elaboradas.

A distribuição das atividades por diversos ambientes do laboratório foi realizada de modo a estimular as inteligências visual-espacial e corporal-cinestésica, uma vez que os alunos precisam se deslocar a fim de encontrar e executar as atividades.

As atividades disponibilizadas inicialmente nesse ambiente são acionadas pelo toque do avatar do aluno no objeto virtual. As atividades habilitadas na primeira fase foram o *chat*, o glossário de termos que pode ser consultado pelo aluno e a cadeira de testes onde o aluno é avaliado por meio de perguntas de múltipla escolha.

E posteriormente foram inclusas atividades referentes a máquina de estados de Mealy e máquina de estados de Moore, portas lógicas, tabela verdade, mapa de Karnaugh e scoreboard.

Essa atividade foi desenvolvida para estimular 3 inteligências, sendo elas, a linguística, sendo estimulada uma vez que o aluno deve ler o texto formulado para questão, bem como o texto referente às respostas para posteriormente escolher a resposta correta.

A inteligência lógico-matemática é estimulada uma vez que o aluno deve executar algumas operações lógicas características do conteúdo abordado, e posteriormente escolher a resposta correta. Já a inteligência interpessoal é estimulada uma vez que os alunos criam uma competição entre eles em busca de uma maior pontuação no *score board*.

O *score board* permite apresentar pontos para os alunos em um curso e caso o professor queira zerar a pontuação, pode fazê-lo através da criação de um novo processo denominado rodada. Como citado anteriormente, a atividade relacionada a este objeto de aprendizagem visa estimular a inteligência interpessoal.

Acredita-se que não só esta inteligência é estimulada, como também a inteligência intrapessoal, uma vez que o aluno, influenciado pelo espírito de competição, pode ser levado a conhecer o seu próprio limite, o que segundo Gardner caracteriza essa inteligência.

O objeto denominado *vending machine* permite que os alunos manipulem objetos 3D disponibilizados pelo professor. Essa manipulação de objetos tem por objetivo estimular a inteligência visual espacial e a inteligência corporal-cinestésica.

Outra atividade desenvolvida foi o *chat*, que permite aos alunos interagir tanto com o professor quanto com os demais alunos. Esta atividade tem por objetivo estimular a inteligência lingüística e a inteligência interpessoal, uma vez que exige que os alunos interajam entre si, criando fóruns de discussão sobre o conteúdo.

Os objetos disponibilizados para todas as atividades desenvolvidas só podem ser visualizados pelo aluno no ambiente 3D caso o mesmo esteja vinculado a algum curso ministrado na plataforma MOODLE e habilitado pelo plugin do SLOODLE através da ferramenta cenário.

V. A SALA DE TESTES

Foram desenvolvidas duas salas de avaliação, onde os alunos podem realizar atividades que irão identificar quanto o aluno assimilou dos conteúdos estudados. Uma delas é apresentada na Figura 6.

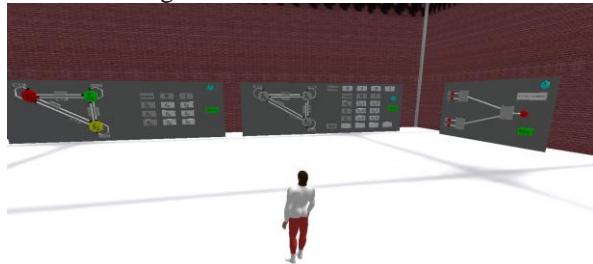


Figura 6. Imagem da sala de testes

O objetivo de desenvolver duas salas para a realização das avaliações foi o de diminuir o tempo gasto na realização das mesmas e para que todo o conjunto de atividades pudesse ser executado em no máximo 1 hora, que é o tempo médio de duração de uma aula.

Dentro das salas de avaliações, que possuem a mesma estrutura e atividades, encontram-se dez objetos 3D e um painel de perguntas e respostas, o objetivo é que todos os alunos façam todas as atividades nos objetos e respondam as perguntas do painel. Os objetos possuem um botão que fornece informações do exercício proposto.

VI. OS TESTES REALIZADOS

Os testes foram realizados com alunos dos cursos de ciência da computação da Unesp/Ibilce de São José do Rio Preto e da Unemat, campus regional de Alto Araguaia de Alto Araguaia.

A população avaliada era constituída de um total de 90 acadêmicos matriculados no 1º e 2º semestres do curso de ciência da computação e do 7º semestre do curso de licenciatura em computação.

Os testes tiveram a duração de média de 1 hora e 30 minutos e durante a realização dos mesmos os acadêmicos puderam realizar atividades que envolviam o uso do ambiente 3D. Dentre as atividades realizadas, foi testada a ferramenta denominada *prim drop*, que permite aos alunos enviar objetos criados no ambiente 3D diretamente para o Moodle.

Os conteúdos abordados durante os testes compõem a ementa das disciplinas de Circuitos Lógicos, nos cursos de bacharelado e de arquitetura de computadores no curso de licenciatura. Foi realizada uma introdução aos conteúdos abordados, disponibilizada por meio de slides.

Ainda em relação ao conteúdo, foi disponibilizado a *quiz chair*, as quais foram vinculadas questões sobre multiplexadores.

Foi solicitado ainda aos participantes que respondessem a um questionário contendo seis questões divididas em dois grupos, que abordavam o uso do ambiente e o método de aprendizagem.

Foi aplicado um questionário *on line* contendo 6 questões divididas em 2 grupos que abordavam o uso do ambiente e a aprendizagem no a partir do uso de diferentes tecnologias.



Figura 7. Resultado do grupo 1 de perguntas

Fonte: Os autores

Os resultados apresentados por meio da Figura 7 demonstra que a aceitação dos alunos em relação ao uso do laboratório, em sua grande maioria foi boa, o que nos leva a concluir que em se tratando da parte estrutural do ambiente, da locomoção e realização das atividades são de modo geral satisfatórios.

Em relação às demais respostas, mediana, fraca e ruim as quais estas duas últimas não foram selecionadas pelos alunos, deve-se a alguns ajustes e melhorias que devem ser realizadas no ambiente, a ser realizados em uma outra fase da pesquisa.

A Figura 4 apresenta os resultados obtidos quanto à aprendizagem utilizando ambiente 3D. De modo geral os alunos estão divididos em relação ao uso do ambiente 3D para ensino, isso talvez se deva a fase inicial da pesquisa onde nem todas as ferramentas pedagógicas foram implementadas no ambiente.



Figura 8. Resultado do grupo 2 de perguntas

VII. CONSIDERAÇÕES FINAIS

Nesta seção são apresentadas algumas considerações e também, algumas perspectivas de trabalhos futuros são elencadas.

O processo de desenvolvimento do laboratório virtual em ambiente 3D, com base na teoria das inteligências múltiplas, proposto neste trabalho, com base nas etapas executadas até o momento nos leva as seguintes conclusões:

Desenvolver um laboratório virtual não é uma tarefa que exija tanto esforço, porém a associação do ambiente a uma teoria pedagógica representa um desafio, pois a relevância pedagógica do uso de ambientes virtuais 3D e que passa necessariamente pela compreensão das possibilidades do uso da tecnologia 3D na área da educação.

No caso específico deste artigo conclui-se que essa associação demandou primeiramente identificar quais inteligências eram possíveis de serem estimuladas e posteriormente em como fazê-las, buscando relacionar cada atividade ou objeto de aprendizagem a uma ou mais inteligências.

A conclusão a que chegamos ao final desta pesquisa, é a de que a forma de estimular as inteligências requer o uso de objetos desenvolvidos e codificados dentro do ambiente, de modo que o comportamento dos mesmos possa ser definido pelo desenvolvedor.

Concluiu-se que para trabalhos futuros uma das atividades será conciliar a construção de objetos com a realização das demais atividades.

REFERÊNCIAS

- [1] Ali, H.B. and Bushro Z. (2008) Multimedia mathematics coursewarebased on the multiple intelligences model (mi-maths). In Information Technology, 2008. ITSim 2008. International Symposium on, volume 2, pages 1–5, 2008.
- [2] Arbaugh, J. B.; Hwang, A (2013). Uses of Multivariate Analytical Techniques in Online and Blended Business Education An Assessment of Current Practice and Recommendations for Future Research. Journal of Management Education, v. 37, n. 2, p. 229-260, 2013.
- [3] Baldi P. and Lopes C. (2012) The universal campus: An open virtual 3-d world infrastructure for research and education. eLearn, 2012(4):6,2012.
- [4] Binet A. (2004) L'étude expérimentale de l'intelligence. Editions L'Harmattan, 2004.
- [5] Costa-Neto, A.; Marranghello, N.; Pereira, A. S. Application of the theory of multiple intelligences to digital systems teaching. 2009.
- [6] Dougiamas, M. (2013) Moodle: Developers manual.[online][consultado em 17.09.2013], 2013.
- [7] Gardner H. A multiplicity of intelligences. Scientific American, 9(4):19–23, 1998.
- [8] Gardner H. Estruturas da mente: A teoria das inteligências múltiplas. Porto Alegre, 1997.
- [9] Gomes L. and Bogosyan S. (2009) Current trends in remote laboratories. Industrial Electronics, IEEE Transactions on, 56(12):4744–4756, 2009.
- [10] Guo, T.T. and Guo, L. and Wang, Z. And Lin, S. and Pan, J. H. (2009) A networked virtual experiment system based on virtual campus. In Education Technology and Computer Science, 2009. ETCS '09. First International Workshop on, volume 3, pages 884–888, 2009.
- [11] Konstantinidis A, Tsatsos T, Demetriadis S., and Pomportsis A. (2010) Collaborative learning in opensim by utilizing sloodle. In Telecommunications (AICT), 2010 Sixth Advanced International Conference on, pages 90–95. IEEE, 2010.
- [12] Kumar, S. Gankotriya, A. K. and Dutta, K. A (2011) comparative study of moodle with other e-learning systems. In Electronics Computer Technology (ICECT), 2011 3rd International Conference on, volume 5, pages 414–418. IEEE, 2011.
- [13] Nunes, F.B. and Voss, G. B. (2013) Tutorial sloodle.Techical report, GRECA - Grupo de Redes e Computação Aplicada -UFSM, 2013.
- [14] Overte Foundation. (2013) Opensimulator: Virtual world, 2013.
- [15] Qingsheng, L. and Aimin, W. Jinyu, K. and Jipeng, W. (2010) Research on multimedia intelligence course and intelligence classroom based on multiple intelligence theory. In Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2010 Sixth International Conference on, pages 398–401, 2010.
- [16] Rico, M. Ramírez, and J. Riofrío, and D. Berrocal-Lobo, and Antonio, A. (2012) An architecture for virtual labs in engineering education. In Global Engineering Education Conference (EDUCON), 2012 IEEE, pages 1–5. IEEE, 2012.
- [17] Ridgewell, W. and Kumar, V. And Lin, O. and K Kinshuk. (2011) Opensim virtual worlds as a platform for enhanced learning concepts. In Advanced Learning Technologies (ICALT), 2011 11th IEEE International Conference on, pages 623–624. IEEE, 2011.
- [18] Ritzema, Tim; Harris, Billy. The use of Second Life for distance education. Journal of Computing Sciences in Colleges, v. 23, n. 6, p. 110-116, 2008.
- [19] Tarouco, L., Gorziza, B., Correa, Y., Amaral, E. M., and Muller, T. (2013, March). Virtual laboratory for teaching Calculus: An immersive experience. InGlobal Engineering Education Conference (EDUCON), 2013 IEEE (pp. 774-781). IEEE.
- [20] Vernaza, A., Armuelles, V. I.; Ruiz, I. Towards to an open and interoperable virtual learning enviroment using Metaverse at University of Panama. In: Technologies Applied to Electronics Teaching (TAEE), 2012. IEEE, 2012. p. 320-325.
- [21] Yasar O. and Adiguzel, T. (2010) A working successor of learning management systems: Sloodle. Procedia-Social and Behavioral Sciences, 2(2):5682–5685, 2010.

Serious Game para Auxílio na Reabilitação de Crianças com Deficiência Física utilizando Interfaces Naturais

Serious Game to Help in Rehabilitation of Children with Disabilities using Natural Interfaces

Flávia Gonçalves Fernandes, Alexandre Cardoso

Programa de Pós-Graduação em Engenharia Biomédica – PPGEB

Faculdade de Engenharia Elétrica – FEELT

Universidade Federal de Uberlândia - UFU

Uberlândia – MG – Brasil

flavia.fernandes92@gmail.com, alexandre@ufu.br

Abstract — The use of virtual rehabilitation through games aims to simulate real situations. Thus, it is observed that the use of this moves the patient focus of pain or discomfort and also enhances the functionality of the affected limbs and leads them to resume the activities in the areas of occupational performance. In this work, it was used the technology of Virtual Reality together with the wearable device Myo, applied to the context of serious games for people with physical disabilities, with the goal of helping them in rehabilitation.

Keywords - natural interfaces; rehabilitation; serious game.

Resumo — O uso da reabilitação virtual através de jogos visa simular situações reais. Assim, percebe-se que o uso desta afasta o paciente do foco da dor ou do incômodo e também melhora na funcionalidade dos membros acometidos e o leva a retomar as atividades nas áreas de desempenho ocupacional. Neste trabalho, foi utilizada a tecnologia de Realidade Virtual juntamente com o dispositivo vestível Myo, aplicados ao contexto de serious games para pessoas com deficiência física, com o objetivo de auxiliá-las na reabilitação.

Palavras-chave - interfaces naturais; reabilitação; serious game.

I. INTRODUÇÃO

A Tecnologia Assistiva, no Brasil, de acordo com o Comitê de Ajudas Técnicas da Secretaria Especial dos Direitos Humanos da Presidência da República, diz respeito a produtos, recursos, metodologias, tecnologias, estratégias, práticas e serviços que objetivam promover a funcionalidade, relacionada à atividade e participação, de pessoas com deficiência, incapacidades ou mobilidade reduzida, visando sua autonomia, independência, qualidade de vida e inclusão social [1].

Os recursos de Tecnologia Assistiva, quando combinados aos recursos tecnológicos, podem atender a uma grande

diversidade de usuários com deficiências. Com o avanço tecnológico e com a proliferação de tecnologias de rede sem fio, os usuários estão, principalmente, interessados em serviços avançados que tornam o ambiente altamente inteligente e facilitam significativamente suas atividades [2].

A deficiência física se torna um grande obstáculo na vida das pessoas com necessidades especiais que são privadas de realizar até mesmo suas atividades do dia-a-dia sem a necessidade de ajuda de outras pessoas. As pessoas que possuem deficiência possuem necessidades especiais, sendo que usuários com deficiências físicas apresentam problemas de mobilidade, tais como dificuldades de acesso ao seu destino final. Essas pessoas com deficiências e as pessoas idosas possuem a necessidade de encontrar o “melhor” caminho de navegação em ambientes externos (outdoor) e em grandes ambientes internos (indoor), tais como hospitais e shoppings centers [3].

Recentemente, observa-se que o desenvolvimento da tecnologia da informação vem auxiliando inúmeras práticas na área da saúde, em atividades como diagnóstico, terapia, gerenciamento e educação, o que exige a necessidade de mudanças e desenvolvimento de novas habilidades pelos profissionais das áreas envolvidas [4].

Neste campo, a Realidade Virtual (RV), definida como uma abordagem entre o usuário e uma interface computadorizada, que envolve a simulação em tempo real de um determinado ambiente, cenário ou atividade e oferece interação ao usuário através de múltiplos canais sensoriais, vem trazendo novas possibilidades para criação, modelagem, visualização, interação e simulação tridimensional de imagens, proporcionando interfaces avançadas capazes de gerar a imersão do usuário em ambientes com os quais pode interagir e explorar [5].

Desse modo, a aplicação da RV na área da saúde, ou mesmo para o desenvolvimento humano, tem sido avaliada de forma intensiva nos últimos anos e merecido destaque, pois representa novos desafios e potencialidades, com a inserção de informações complementares e/ou relevantes ao cenário real [6].

Os jogos digitais deixaram de ser vistos como uma forma de entretenimento prejudicial à saúde [7]. Games tornam-se uma ferramenta importante para melhorar o tratamento dos pacientes que vão desde aqueles que estão atravessando uma grave enfermidade, como o câncer, por exemplo, até os que demandam procedimentos mais leves, como a fisioterapia [8].

Nessa perspectiva, este trabalho apresenta o desenvolvimento de um *serious game* para auxiliar na reabilitação de indivíduos com deficiência física utilizando Realidade Virtual e o *wearable Myo*. Logo, este sistema visa motivar os pacientes a realizar o tratamento de maneira interativa e lúdica, tornando o tratamento mais atrativo e dinâmico.

II. TRABALHOS RELACIONADOS

Dentre os trabalhos disponíveis na literatura, alguns deles apresentam desenvolvimento, conceitos, referências teóricas e frameworks relacionados a *serious games*.

PhysioPlay é um exergame para reabilitação física, como mostra a Figura 1. Seu cenário é um galpão fechado que foi projetado em um ambiente virtual tridimensional para que a integração sensorial e motora do paciente capturada pelo sensor Kinect possibilite a interação do corpo e dos objetos do ambiente como biofeedback visual, contribuindo, assim, para o desenvolvimento de habilidades motoras, além da atenção e concentração como terapia ocupacional [9].



Figura 1. Cenário do exergame PlaysioPlay [9].

MoVER é um jogo da categoria *serious game* que simula movimentos fisioterapêuticos por meio de desafios para a realização de tarefas virtuais com o uso do corpo humano. Tem como objetivo apresentar o desenvolvimento, a implementação e uma aplicabilidade prática na reabilitação física de um *serious game* para reabilitação motora em ambiente virtual, de maneira a auxiliar profissionais da área de reabilitação a definir e avaliar exercícios para o paciente [10]. Segundo os autores, o mesmo oferece apoio para estimular o paciente na execução de exercícios de reabilitação de membros superiores.

Na atividade do jogo, são colhidas informações referente a posição do objeto e das mãos do paciente, tempo percorrido para atingir o objetivo e distância entre

deslocamento da mão ao objeto. Conforme mostra a Figura 2, o semáforo tem a função de corrigir a postura do jogador, evitando que o mesmo se antecipe ao movimento. Assim, a execução do jogo é liberada somente quando o usuário se encontra na postura correta, que é indicada pelo semáforo na luz verde.

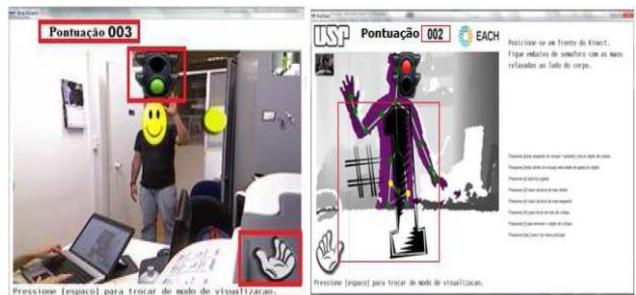


Figura 2. Cenário do serious game MoVER [10].

Os *serious games* *PhysioPlay* e *MoVER* são utilizados para reabilitação fisioterapêutica. Porém, eles possuem limitações quanto à quantidade de exercícios realizados no contexto do jogo. Já o *Serious Game* para auxílio às crianças que não possuem o antebraço, apresentado neste trabalho, possui mais benefícios, uma vez que simulam as atividades cotidianas com o uso de prótese, com o intuito de motivar essas crianças a utilizar a prótese na vida real.

Torres desenvolveu um jogo para auxiliar no treinamento de exame de biópsia mamária. O aprendiz manipula um instrumento médico usando um dispositivo com retorno haptico executando os procedimentos necessários para treinar o exame de biópsia [11], segundo pode ser visto na Figura 3.



Figura 3. Sistema para realização de biópsia da mama [11].

Investigou-se sobre a possibilidade do dispositivo *Kinect* ser utilizado no tratamento de pacientes com distúrbios neurológicos de maneira segura e coerente. Concluiu-se que o dispositivo torna a rotina do tratamento mais dinâmica e alivia o estresse do procedimento convencional, mas os exercícios precisam ser acompanhados pelo profissional capacitado a fim de que possa ser observada a evolução do procedimento empregado [12].

Os trabalhos citados confirmaram que a utilização dos jogos digitais ajuda a estimular os usuários a realizarem atividades do mundo real. Além disso, sabe-se que o uso de Interfaces Naturais é recente e estimulante, uma vez que são

elementos naturais humanos, quase imperceptíveis quando o usuário está imerso na aplicação. Por isso, as Interfaces Naturais (NUI) constituem o foco do presente trabalho.

III. TECNOLOGIAS UTILIZADAS

A. Blender

Blender é um programa de computador de código aberto, desenvolvido pela *Blender Foundation*, para modelagem, animação, texturização, composição, renderização, edição de vídeo e criação de aplicações interativas em 3D, tais como jogos, apresentações e outros, através de seu motor de jogo integrado, o *Blender Game Engine* [13].

O programa é multiplataforma, estando, portanto, disponível para diversos sistemas operacionais. Em modelagem, o Blender pode ser utilizado em qualquer área que seja necessária a geração de modelos tridimensionais, geração de imagens renderizadas, animação e jogos. Podemos citar aplicações em arquitetura, design industrial, engenharia, animação, produção de vídeo, e desenvolvimento de jogos, graças ao seu motor de jogo embutido. Esta característica pode ser ampliada e agilizada com o uso de scripts em *Python*.

B. Unity 3D

Unity, também conhecido como Unity 3D, é um motor de jogo 3D proprietário, criado pela Unity Technologies. Unity é similar ao Blender, Virtools ou Torque Game Engine, em relação a sua forma primária de autoria de jogos: a sua interface gráfica [14].

O Unity possui duas versões principais: Unity Pro, que é paga, e a versão gratuita, simplesmente chamada Unity, que pode ser usada tanto para fins educacionais, quanto para fins comerciais. Suas principais características são [14]:

- Suporte para o uso de shaders;
- Programação em Boo, C# ou JavaScript;
- Suporte ao PhysX, incluindo detector de colisão, soft body e ragdoll;
- Compatibilidade com Blender, 3ds Max, Maya, Cinema 4D, Cheetah 3D, Softimage, modo, ZBrush, Lightwave, Photoshop, Fireworks e Substance.
- Geração de jogos para Web e Mobile: compatibilidade com os principais browsers (via o plugin Unity Web Player): Internet Explorer, Firefox, Safari, Opera, Google Chrome e Camino.

C. Linguagem C#

O C# é uma linguagem de programação visual dirigida por eventos e totalmente orientada a objetos. Permite um novo grau de intercâmbio entre linguagens (componentes de software de diferentes linguagens podem interagir). Os desenvolvedores podem empacotar até software antigo, para trabalhar com novos programas C#. Além disso, os

aplicativos C# podem interagir pela Internet usando padrões do setor, como SOAP (Protocolo de Acesso a Objetos Simples) e XML (Linguagem de Marcação eXtensível).

O C# tem raízes em C, C++ e Java, adaptando os melhores recursos de cada linguagem e acrescentando novas capacidades próprias. Ele fornece os recursos que são mais importantes para os programadores, como programação orientada a objetos, strings, elementos gráficos, componentes de Interface Gráfica com o Usuário (GUI), tratamento de exceções, múltiplas linhas de execução, multimídia (áudio, imagens, animação e vídeo), processamento de arquivos, estruturas de dados pré-empacotadas, processamento de banco de dados, redes cliente/servidor com base na Internet e na World Wide Web e computação distribuída [15].

D. Wearable Myo

Dispositivos portáteis ou vestíveis, também conhecidos como *wearables*, são pequenos dispositivos eletrônicos que podem ser facilmente acoplados ao corpo do usuário. Eles possibilitam a presença da Realidade Virtual em diversas aplicações, visto que é uma tecnologia de interface avançada entre um usuário e um sistema computacional [16].

O *Myo* é um *wearable* com formato de bracelete, também chamado de braçadeira. Este dispositivo pode controlar aplicações por meio de gestos e interagir com computadores e outros meios digitais similares, reconhecendo impulsos elétricos nos músculos do usuário. Além disso, não exige câmeras para rastrear os movimentos da mão ou braço e possuem baixo custo [17].

O *Myo* é utilizado no braço ou antebraço do usuário, conforme pode ser observado na Figura 4. É necessário realizar uma calibração da braçadeira para cada usuário específico, pois cada um possui atividades e contrações musculares específicas. Após efetuada a sua calibração, o *wearable* possibilita controlar softwares e outras aplicações por meio de gestos e movimentos. Seu propósito é controlar computadores, telefones e outros dispositivos, enviando os dados capturados por ele via *Bluetooth* [17].



Figura 4. Wearable Myo.

IV. VISÃO GERAL DA APLICAÇÃO

A aplicação, na forma de um *serious game* suportado por RV e interfaces naturais (NUI) visa auxiliar na reabilitação de crianças com deficiência física, inicialmente, com idade entre oito e quatorze anos, do sexo masculino ou feminino que nasceram sem o antebraço.

Com relação aos aspectos metodológicos e tecnológicos, para a implementação desta aplicação, utilizou-se RV por meio do desenvolvimento de ambientes virtuais, incluindo interações e animações, com uso da linguagem de programação C#. Para a modelagem dos objetos tridimensionais, utilizou-se o software Blender, e para a criação dos cenários do jogo optou-se pelo uso do motor de jogo Unity 3D, pois ambos são ferramentas gratuitas e muito populares para o desenvolvimento de *games*.

Na Figura 5, apresenta-se o diagrama de casos de uso do *serious game*.

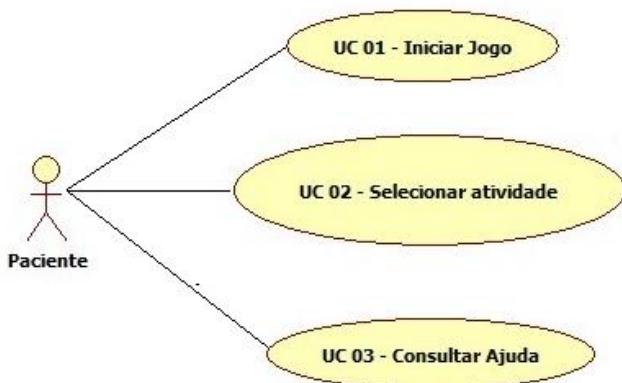


Figura 5. Diagrama de casos de uso.

O *serious game* apresentado neste trabalho é persuasivo e possui três fases, com desafios incrementais, em que a criança joga utilizando o dispositivo *wearable Myo*.

O personagem do jogo é uma criança que utiliza uma prótese para realizar as atividades propostas no *serious game*, e que ganha pontos e poderes conforme avança de nível, além de permitir a troca de prótese. Assim, o jogo tem o propósito de retratar a realidade da criança.

Na Figura 6, é apresentada a arquitetura da aplicação, mostrando que a criança, ao utilizar o dispositivo *wearable Myo*, joga o *serious game*, tentando atingir cada vez mais pontos e avançar nas fases do jogo.

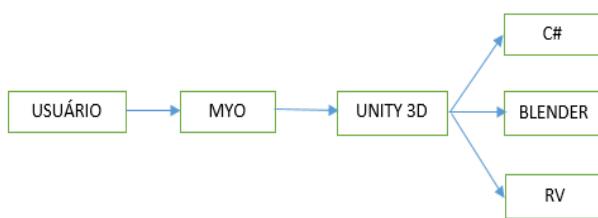


Figura 6. Arquitetura da aplicação.

Primeiramente, a criança coloca o dispositivo *wearable Myo* em seu braço para realizar os movimentos necessários durante a execução do *serious game* e se posiciona em frente a um computador. A calibração do Myo no braço humano pode ser visualizada na Figura 7.

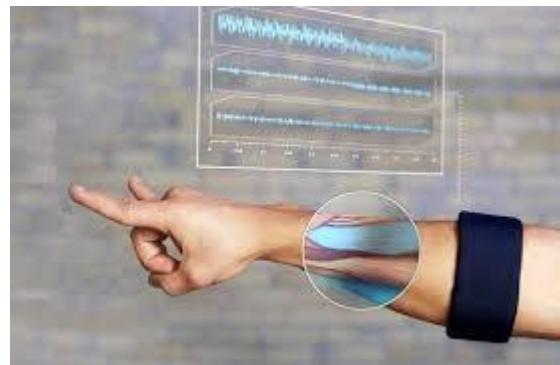


Figura 7. Calibração do Myo no braço humano.

O *wearable Myo*, ao invés de reproduzir diretamente os movimentos, como o Kinect, tem formato de braçadeira e fica posicionado no braço do usuário; através do contato com a pele, percebe a atividade elétrica dos músculos do braço, traduzindo os mínimos movimentos para um computador com a capacidade de controlar uma variedade de aparelhos.

O Myo possui a habilidade de controlar aparelhos, reagindo a partir de movimentos sutis e filtrando movimentos inúteis que braço e mão realizam. A variedade de aparelhos que podem ser controlados é muito grande, sendo possível jogar games, controlar vídeos, passar slides e manipular aeromodelos apenas com os movimentos do braço e da mão.

Após a calibração do Myo no braço da criança, ela seleciona o seu sexo (menino ou menina) e escolhe com o qual personagem quer jogar.

Os símbolos apresentados na Figura 8 são os principais gestos realizados pelo usuário durante a execução do *serious game* controlado pelo *wearable Myo*, representando os comandos necessários para a boa continuidade do jogo, os quais são, respectivamente: “Iniciar Jogo”, “Ir para a direita”, “Ir para a esquerda” e “Pausar Jogo”.

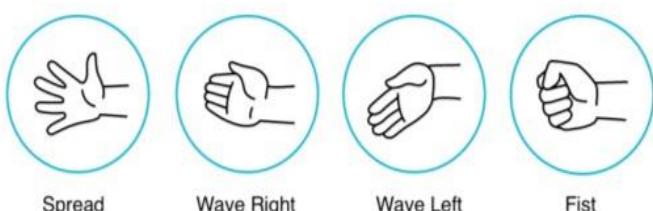


Figura 8. Gestos executados pelo usuário com o Myo.

Na primeira fase, há a prótese que auxilia a pegar objetos. Foram inseridos no jogo objetos que fazem parte do cotidiano da criança, como brinquedos, bolas, bonecas,

carrinhos, conforme o gênero da criança. Desse modo, quanto mais objetos a criança conseguir pegar, mais pontos ela adquire. E ao atingir determinada pontuação, o personagem avança o nível do game.

Na Figura 9, é exibido um esquema da Primeira Fase do *Serious Game* no software Blender.

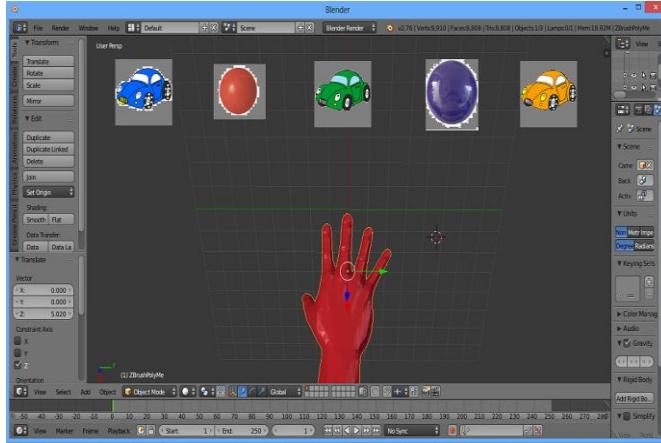


Figura 9. Primeira Fase do *Serious Game* no Blender.

Na segunda fase, há a prótese que ajuda o personagem do jogo a comer. Para este nível, foram utilizados alimentos saudáveis, como frutas, verduras, legumes. Dessa forma, quanto mais o personagem se alimentar de maneira saudável e usando a prótese de modo adequado, mais pontos ele consegue e avança o nível do jogo.

Na Figura 10, é apresentado um esquema da Segunda Fase do *Serious Game* no software Blender.



Figura 10. Segunda Fase do *Serious Game* no Blender.

Na terceira fase, há a prótese que ajuda o personagem do jogo a utilizar dispositivos tecnológicos, como celulares, tablets, computadores, notebooks, por meio dos principais comandos de uso destes. Quanto mais acertos o usuário conseguir realizar na execução destes comandos, maior será sua pontuação.

Na Figura 11, é mostrado um esquema da Segunda Fase do *Serious Game* no software Blender.

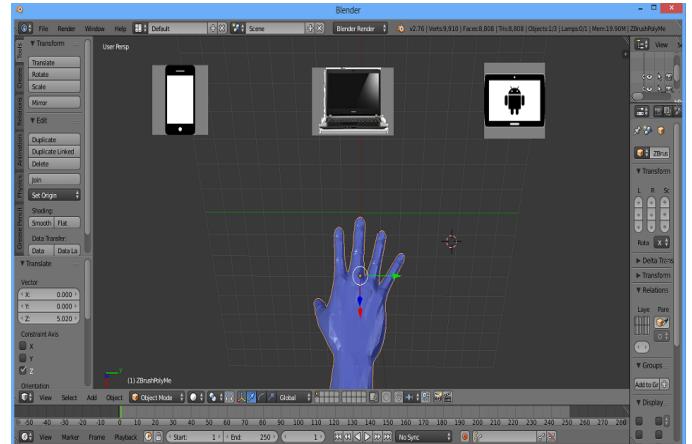


Figura 11. Terceira Fase do *Serious Game* no Blender.

Na Figura 12, pode ser visualizado o cenário do *serious game* no software Unity 3D.

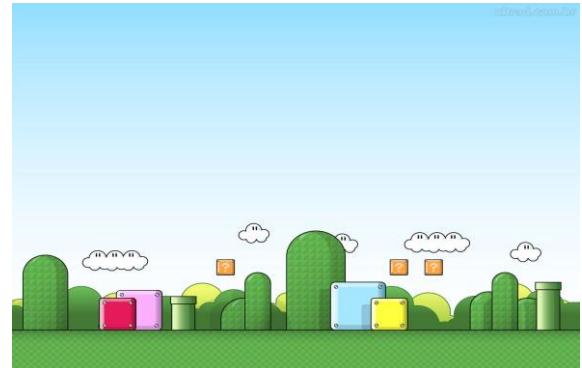


Figura 12. Cenário do serious game no Unity 3D.

Logo, ao concluir as três etapas do *Serious Game*, a criança vence o jogo e torna-se mais motivada para utilizar a prótese na vida real em seu cotidiano diário, conforme pode ser observado na Figura 13.



Figura 13. Criança utilizando prótese na vida real.

Myo se destaca em um mar de tecnologias de controle por gesto, muitas das quais dependem de câmeras ou exigem hardware volumoso para reconhecer os gestos do usuário e traduzi-los em ações em uma tela. Além de ocupar

espaço, esses sistemas podem precisar ser calibrados ou exigir uma certa quantidade de luz para operar, todos fatores que podem limitar onde e como poder usá-los. E sabe-se o quanto os consumidores querem se desfazer de mouse, teclados e telas sensíveis ao toque para o controle por gesto.

Uma vez que a braçadeira Myo interpreta os impulsos elétricos gerados pelos movimentos musculares em seu braço ou antebraço, ele não precisa nem de luz, nem de uma câmera para funcionar. Isso, juntamente com o seu tamanho relativamente pequeno, pode torná-la mais fácil de usar em salas escuras ou à luz solar intensa e pode oferecer aplicativos móveis interessantes, como permitir controlar os recursos em um relógio inteligente.

V. CONCLUSÕES

Serious games baseados em Realidade Virtual e *wearables* são inovações tecnológicas inseridas aos poucos no cotidiano das pessoas. A sua principal utilidade é promover interesse e motivação aos usuários ao realizar certas atividades, como de reabilitação e tratamento de doenças.

Myo pode diferenciar entre diferentes movimentos dos dedos e identificar rotações e movimentos da mão, medindo os diferentes padrões de impulsos elétricos, que os movimentos do usuário geram e usam um sensor de inércia para entender os movimentos. Com a braçadeira no braço, pode-se fazer coisas, como imitar disparar uma arma para controlar uma arma de fogo em um vídeo game, ou passar a mão para passar os slides de uma apresentação. Esta informação é enviada a um processador na braçadeira, e um algoritmo traduz os comandos, que são enviados via Bluetooth de baixo consumo energético para o *gadget* que o usuário está tentando controlar, como um *smartphone*.

As principais dificuldades encontradas na realização deste trabalho foram a modelagem para a construção do jogo e o reconhecimento dos movimentos pelo dispositivo *wearable*, devido à sua complexidade e à diversidade de variáveis envolvidas neste processo.

Como trabalhos futuros, pretende-se prosseguir com o desenvolvimento de mais cenários no jogo com a finalidade de melhorar o incentivo ao uso de prótese em crianças com deficiência física nos membros superiores. Também implementar o *serious game* para outros membros do corpo humano. E também realizar o desenvolvimento de um módulo de supervisão da aplicação, no qual o profissional habilitado poderia visualizar gráficos para acompanhar a evolução de determinando paciente, e também comparar os resultados obtidos por diferentes pacientes na execução do jogo.

AGRADECIMENTOS

Os autores agradecem à CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) pelo apoio

financeiro que permitiu a realização deste trabalho, e também ao Grupo de Realidade Virtual e Aumentada (GRVA) da Universidade Federal de Uberlândia (UFU) pelo incentivo.

REFERÊNCIAS

- [1] Brasil. Subsecretaria Nacional de Promoção dos direitos da pessoa com deficiência - CORDE. 2007. Disponível em: <<http://portal.mj.gov.br/corde>>. Acesso em: 10 abr. 2015.
- [2] L. M. C. Santarosa; D. Conforto; L. D. O. Basso. Eduquito: ferramentas de autoria e de colaboração acessíveis na perspectiva da web 2.0. In: Revista Brasileira de Educação Especial, v. 18, n. 3, 2012. ISSN 1413-6538.
- [3] V. Tsetsos et al. Semantically enriched navigation for indoor environments. International Journal of Web and Grid Services (IJWGS), v. 2, n. 4, p. 453-478, 2006. ISSN 1741-1106. R. M. Costa; M. W. Ribeiro. Aplicações de realidade virtual e aumentada. Porto Alegre: SBC, 2009. 146 p.
- [4] C. Kirner; R. A. Tori. Fundamentos de Realidade Virtual e Aumentada. In: IX Symposium on Virtual and Augmented Reality, 28. 2007, Petrópolis. Anais.Petrópolis: LNCC, 2007. p. 02-21.
- [5] M. Ribeiro; E. Zorزال. Realidade Virtual e Aumentada: Aplicações e Tendências. In: XIII Symposium on Virtual and Augmented Reality. 23. Uberlândia-MG: Sociedade Brasileira de Computação – SBC, 2011.
- [6] L. S. Machado; R. M. Moraes; F. L. S. Nunes. Abordagens Práticas de Realidade Virtual e Aumentada. In: XI Symposium on Virtual and Augmented Reality. 25. Porto Alegre - RS: Pontifícia Universidade Católica – PUCRS, 2009.
- [7] R. S. Dias; I. L. A. Sampaio; L. S. Taddeo. A Introdução do Lúdico no processo de reabilitação de pacientes em tratamento fisioterapêutico. In: VIII Brazilian Symposium on Games and Digital Entertainment. 4. Rio de Janeiro-RJ, 2009.
- [8] R. S. Monteiro Junior et al. Efeito da reabilitação virtual em diferentes tipos de tratamento. Revista Brasileira de Ciência da Saúde, Rio de Janeiro, n. 29, p. 56-63, jul/set. 2011.
- [9] J. V. S. Santos; L. C. Carvalho; P. A. Bressan. Physioplay: um exergame para reabilitação física aplicando a interatividade do Kinect como biofeedback visual. In: IX Workshop de Realidade Virtual e Aumentada (WRVA), Paraná, 2012
- [10] V. D. Sousa Junior et al.. MoVER: Serious Game aplicado à reabilitação motora usando sensor de movimento Kinect. In: Congresso da Sociedade Brasileira de Computação (CSBC), Maceió, 2013.
- [11] R. S. Torres; H. H. Biscaro; L. V. Araujo; F. L. S. Nunes. ViMeTGame: A serious game for virtual medical training of breast biopsy. SBC Journal on 3D Interactive Systems: 3, p. 12-19, 2012.
- [12] P. R. Rocha; A. H. Defavari; P. S. Brandão. Estudo da viabilidade da utilização do Kinect como ferramenta no atendimento fisioterapêutico de pacientes neurológicos. In: XI Simpósio Brasileiro de jogos e Entretenimento Digital, Brasília – DF, 2012.
- [13] Blender. Disponível em: <<http://blendertotal.wordpress.com/o-que-e-blender/>>. Acesso em: 01. Jun. 2015.
- [14] Unity 3D. Disponível em: <www.unity3d.com>. Acesso em: jun. 2015.
- [15] Introdução à linguagem C# e .NET Framework. Disponível em: <www.microsoft.com/pt-br/library/vstudio/z1zx9t92.aspx>. Acesso em: jun. 2015.
- [16] S. Mann. Computação Wearable. In: Soegaard, Mads e Dam, Rikke Friis (eds.). "Encyclopedia of Intereração Humano-Computador". Aarhus, 2012. Dinamarca: A Fundação Interaction-Design.org.
- [17] Thalmic Labs - Myo Gesture Control Armband. Disponível em: <<https://www.thalmic.com/myo/>>. Acesso em: 30 jun. 2015.

Prototipação de uma Ferramenta de Interface Natural com *Finger Tracking* para Ensino Multisensorial

Eduarda Maganha de Almeida, Adriane Carla Anastácio Silva, Luciano Henrique Quintino, Eduardo Filgueiras Damasceno

Universidade Tecnológica Federal do Paraná

Departamento de Computação - Laboratório de Computação Visual

email: {duda.maganha, luciano.hq}@hotmail.com, {damasceno, anastacio}@utfpr.edu.br

Resumo—This work presents a prototype of a tool developed with the reality technology increased to stimulate of way lúdica the user to learn, to associate and to understand what is doing, and the natural interaction to track the movements of the user's fingers through colored markers. The prototype possesses characteristics of artificial intelligence, implemented with an algorithm of net neural. When evaluating the capacity of the software application and the success tax and as that process it can interfere in the knowledge establishment for the students. The tests were accomplished in two children with learning disturbances. As result, we presented the advantages of her application and to discuss the coming of this.

Keywords—*Prototype, Augmented Reality, Natural Interaction, Learning Disorders;*

I. INTRODUÇÃO

Atualmente, apesar de estarmos vivenciando o uso e abuso de tecnologias computacionais continuamos a nos deparar com problemas do dia a dia na Educação. A Tecnologia Educacional está a apoiar docentes, pais e psicopedagogos em uma abordagem de alfabetização de crianças com distúrbios de linguagem, em alguns casos, percebemos que existe um *gap* entre a tecnologia e a práxis pedagógica. Os educadores alfabetizadores confrontam-se com desafios no processo de ensino e aprendizagem de crianças e jovens com distúrbios de linguagem, observa-se da escassez de recursos à ausência de tecnologias que auxilie na aprendizagem [8].

Os distúrbios ou transtornos de linguagem são analisados como uma dificuldade para decodificar tanto a palavra escrita como a oral, o que prejudica a compreensão do material lido e/ou ouvido, alterando e comprometendo o aproveitamento escolar [15]. Estes devem ser diagnosticado, preferencialmente, por especialistas, como os psicólogos, os fonoaudiólogos, os psicopedagogos e em alguns casos um neuropediatria [18].

Existem diversos métodos empíricos que apoiam o tratamento de tais transtornos, todavia, os mais aceitos métodos de ensino são o fônico e o multissensorial [5]. O método fônico utiliza-se dos sons fônicos das letras associados a desenhos lúdicos para que o aluno faça a associação correta [20]. No entanto, o método multissensorial o aluno observa as letras e/ou palavras impressas em uma folha de papel e pronuncia e/ou desenha o movimento de sua escrita [17]. Os dois métodos não dispensam a atenção do terapeuta ou do psicopedagogo,

que muitas vezes exige uma melhor especificação para a execução do processo e da tecnologia para promove-lo.

Nos últimos anos houve uma maior conscientização das autoridades escolares em relação as dificuldades de portadores de distúrbios de aprendizagem. De acordo com a Lei nº 9394 (20/12/1996), de Diretrizes e Bases da educação nacional (LDB), o sistema de ensino deve assegurar a seus educandos com necessidades educacionais especiais um processo que envolva uma práxis docente com intuito de assegurar seu direito de inclusão [4].

Entretanto, na maioria das vezes os métodos praticados e a obrigatoriedade da lei não são suficientes para que se obtenha um resultado satisfatório na processo de aprendizagem do educando. Necessitando que as práticas de alfabetização utilizadas sejam revistas e adaptadas de forma a atender melhor o aluno com esta necessidade especial.

Acredita-se então que desenvolver uma aplicação multimídia que combine o artefato visual, a interação gestual, a animação aliados à técnicas de Realidade Aumentada, será possível auxiliar os profissionais no tratamento de distúrbios de aprendizagem.

II. FUNDAMENTAÇÃO TEÓRICA

A. Práticas na alfabetização de crianças com distúrbios de aprendizagem

É muito comum encontrar educadores que praticam o método fônico nas escolas com abordagens que incentivam os educandos a desenvolver habilidades de rima, segmentação fonêmica, discriminação de sons, e ensinam as relações entre as letras e os sons [13]. Contudo, encontramos na língua portuguesa uma dificuldade maior do que na língua inglesa de aprendizado, considerando a irregularidade das relações grafofonêmicas da ortografia, com equivalências imprevisíveis entre letras e sons [17].

Diferente método e muito praticado no cotidiano dos educadores é o da escrita multissensorial, uma combinação de distintas modalidades sensoriais no ensino da linguagem escrita. Para Capovilla [5] este método facilita a leitura e a escrita a partir da conexão entre aspectos visuais, a desenho ortográfico da letra ou da palavra; os aspectos auditivos, a configuração fonológica; os aspectos táteis e cinestésicos da grafia, movimentos necessários para escrita das letras e

palavras; e aspectos cinestésicos da articulação, movimentos e posições para pronunciar sons e palavras .

No Brasil, um procedimento multissensorial bastante difundido e adotado é o fonovisuoarticulatório, conhecido como método das boquinhas. Da Silva [6] garante que o este método indicado tanto para a alfabetização de crianças quanto na reabilitação dos distúrbios da leitura e escrita, a partir de estratégias fônicas (fonema/som), visuais (grafema/letra) e articulatórias (articulem/boquinhas); e em casos de fracasso escolar de crianças mais velhas [20]. Nesta linha outras abordagens vem ganhando espaço, a partir da combinação de métodos que envolvem a repetição oral e reforço da aprendizagem [21]. Estas são focadas na repetição da letra e de seu som, acrescenta unidades as palavras e, em seguida, as frases.

B. Tecnologia Educacional para distúrbios da aprendizagem

No universo da educação não podemos deixar de citar a tecnologia como um marco de progresso e avanço, aplicada a métodos de ensino e aprendizagem. A cada dia novas tecnologias aplicadas à educação têm influenciado as novas formas de aprendizagem e ampliado oportunidades aos alunos que sofrem distúrbios da aprendizagem.

Um exemplo é o método das boquinhas, aprovado como tecnologia educacional pelo MEC, em dezembro de 2009, uma proposta multissensorial, uma nova oportunidade oferecida ao aluno no processo de aprendizagem [2]. A tecnologia educacional vem transformando a tecnologia em recursos de aprendizagem, disponibilizando aos alunos em sala de aula, redes sociais, games, plataformas e aplicativos interativos.

Os softwares interativos de ensino como o CBE (*Computer Based Education*), permite desenvolver um processo de aprendizagem mais interessante por parte do aluno e a transmissão de conhecimentos por parte do professor [10]. Esses softwares são desenvolvidos na maioria das vezes utilizando tecnologias multimídia, que ganham cada vez mais espaço no ambiente educacional. Para Andre [2], essas plataformas interativas são ferramentas de apoio, incentivo, motivação e desenvolvimento das habilidades do aluno.

Outros conceitos tecnológicos igualmente estão sendo inseridos no ambiente escolar, como o conceito de gamificação e realidade aumentada. No primeiro caso, estão sendo desenvolvidos jogo educativos que permitem ao aluno realizar diferentes atividades, estimulando o desenvolvimento das habilidades cognitivas e emocionais [11].

No campo da realidade aumentada, destaca que as potencialidades da tecnologia são limitadas pela imaginação, uma vez que completa a percepção e interação com o mundo real, a partir de imagens e objetos tridimensionais virtuais, permite ao utilizador aderir a um ambiente de virtual gerado por computador [7]. Azuma [23] acredita ser é necessário levar em consideração qual tecnologia específica será ideal para atingir os objetivos do indivíduo ou grupo que irá utilizá-lo. No entanto, a escolha da tecnologia deve ser consciente, refletindo favoravelmente no desenvolvimento da aluno com necessidade educacional especial, como abordado, de distúrbios de aprendizagem, por meio da interação serão capazes de mitigar suas dificuldades e desenvolver suas potencialidades.

C. Interação Natural como tecnologia para apoiar o Ensino

A interação natural é a maneira dos seres humanos se comunicarem por meio de gestos, expressões faciais, movimentos corporais e da fala, na percepção do mundo a sua volta, ouve e manipula materiais físicos sem, ou quase sem, mediadores tecnológicos, como por exemplo: joysticks, mouses, teclados e sensores vestíveis, luvas hapticas e *head mounted displays* [22].

Um canal natural de comunicação direta entre o usuário e o ambiente digital, convergindo para a mesma ação de comunicação do usuário dentro ou fora do ambiente. Seria um ambiente ideal para gerar soluções realmente interativas, por meio da combinação integrada todos estes dados fornecidos pelo usuário. Contudo, mesmo com os atuais avanços na tecnologia da informação, há lacunas a serem exploradas, como de uma aplicação a partir da combinação destes dados, ampliando o nível de sofisticação e processamento.

Na área de interação natural uma das mais relevantes linhas de pesquisa é da interpretação de movimentos corporais. Esta interpretação ocorre a partir do rastreio, processamento e reconhecimento de padrões de movimento, que possibilita uma das mais brilhantes técnicas de interação, a Humano-Computador.

Uma das possibilidades exploradas na técnica de interação, a Humano-Computador é o uso de dispositivos e acessórios anexos ao corpo para auxiliar o rastreamento. Neste sentido, há dispositivos que possuem sensores iniciais associados como acelerômetros e giroscópios e assim conseguem medir o deslocamento e rotação de partes do corpo ou partes dele [7]. Há técnicas baseadas em captura de movimento por câmeras de vídeo, que por meio de marcadores afixados no corpo fornecem as informações necessárias a uma unidade de captura e técnicas que dispensam o auxílio de acessórios, deixando o corpo do usuário mais livre .

Uma das técnicas de rastreamento por meio de sensores de profundidade, como o Microsoft Kinect [12]. Contudo, este dispositivo ainda está longe de ser encontrado nas escolas brasileiras, dificultando o uso de aplicações com interação natural por meio deste sensor de movimento. Na maioria das escolas percebe-se que existe um outro dispositivo mais acessível e comum, a webcam [9], foco de nosso trabalho.

Para realizar o rastreamento de um objeto pode ser definido como o problema de estimar a trajetória de um objeto no plano da imagem enquanto se move na cena [19]. Observando a cena em nosso projeto encontramos alguns desafios, como por exemplo: a diferenciação entre as áreas de fundo e o primeiro plano; a diferenciação entre os objetos monitorados e outros objetos em movimento e extração de características relevantes ao rastreamento; as mudanças na iluminação, provocando mudanças nos objetos monitorados; a normalização dos objetos que mudam com a distância da câmera e as possíveis oclusões entre os objetos presentes. E, para iniciar a solução de nosso desafio, primeiramente é requisito o entendimento do tipo de tratamento que realizado a cada imagem para identificar o que está em movimento ou não.

D. Realidade Aumentada para auxílio da escrita

Azuma [3] definiu realidade aumentada como um sistema que apresenta três características: combina o real com o virtual; é interativa em tempo real; e ajusta os objetos virtuais no ambiente. A realidade aumentada mantém o usuário no seu ambiente físico e transporta o ambiente virtual para o espaço do usuário, por meio de algum dispositivo tecnológico., sendo assim, a interação do usuário com os elementos virtuais ocorre de maneira natural e intuitiva, sem necessidade de adaptação ou treinamento [23]. A realidade aumentada combina recursos de multimídia e realidade virtual, para apresentar elementos misturados de boa qualidade e prover interação em tempo real. Como a realidade aumentada mantém o senso de presença do usuário no mundo real, há uma forte tendência em usar recursos tecnológicos invisíveis ao usuário para deixá-lo livre em seu ambiente. Recursos como rastreamento ótico, projeções, interações multimodais, etc., estão sendo usados cada vez mais em aplicações de realidade aumentada. Atualmente as aplicações de Realidade Aumentada têm alcançado diversas áreas, sendo elas: saúde, educação, comercial, artística, musicais entre outras. [7]. Comprovando essa afirmação, destacamos a área da saúde com o projeto ReabRA, que significa Reabilitação Cognitiva através da Realidade Aumentada, esta aplicação desenvolvida por Nascimento, capaz de prover elementos de estimulação da atenção da memória, em um programa de tratamento de pessoas com diferentes lesões cerebrais.[7]. A nossa ferramenta visar auxiliar no desenvolvimento da escrita, e para isso foi adicionado à tela uma imagem lúdica 3D que faz associação com a letra que sera escrita. Como pode ser observado na ilustração da ferramenta na Figura 1, pois a partir do *Finger Tracking* por coloração das pontas dos dedos, a figura irá aparecer no canto superior da tela.

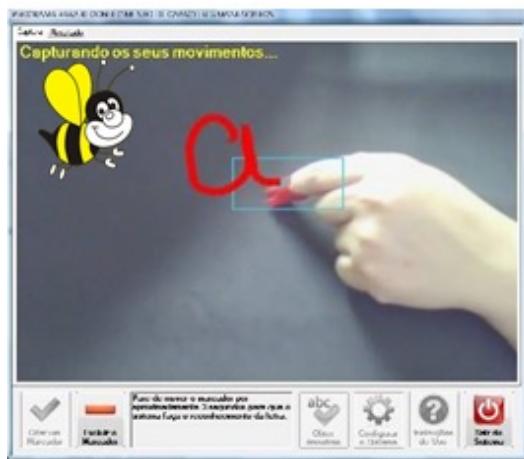


Figura 1. Eventos Associados

E. Identificando o movimento na imagem em tempo real

A captura de movimentos baseado em marcadores coloridos é uma aplicação da técnica de processamento de imagem. Esta

técnica realiza a análise e a variação de cor de um conjunto de pixels na imagem, em função do tempo. A identificação de um padrão de cor em formato RGB e a variação deste padrão pode ser rastreado, basicamente, uma identificação dos bloco de imagem usado como *template*.

O rastreamento pode ser implementado de forma recursiva, deve ser inicializado com a posição de origem do objeto rastreado e a sua cor. É comum adotar um detector para identificar o objeto de interesse na imagem de entrada, chamada de *Region Of Interest*, e fornece a informação indicativa da inicialização do algoritmo de rastreamento. De forma a ilustrar este processo, apresentamos a Figura 2, uma ilustração da nossa ferramenta a partir de 3 eventos associados.



Figura 2. Eventos Associados

O evento A ocorre na inicialização do software que realiza a marcação do objeto de interesse na imagem. Foi utilizada uma fita na cor vermelha para identificar a região da imagem que é associada como um ponto marcador. No evento B se destaca o ponto da imagem identificado para rastreio; e por fim, no evento C é apresentado o objeto sendo identificado e rastreado. No desenvolvimento e testes dos algoritmos, as imagens foram capturadas utilizando o espaço de cores RGB, com resolução de 352x288 e com uma taxa de 21 quadros por segundo, e no rastreio foi utilizado o algoritmo CAMSHIFT (*Continuously Adaptive Mean-Shift*).

O algoritmo desenvolvido é capaz de rastrear um objeto baseando-se apenas em sua localização inicial e histograma. O rastreamento utiliza informações de cor representada no espaço de cores HSV. No projeto foram adotadas as técnicas de rastreamento por coloração, considerando as características do software; e de ensino multissensorial, para estimular o usuário a interagir com os dedos e sentir a escrita da letra, associando com o som e o mundo lúdico simulador, gerado pela multimídia.

F. Redes Neurais Artificiais

Foi utilizada nos projeto uma Rede Neural Artificial (RNA) para realizar o reconhecimento do letra escrita com o dedo. Entretanto, para este reconhecimento foi necessário relacionar uma entrada de escrita à um dos caracteres pertencente a um conjunto, denominado alfabeto. Um único caractere de um alfabeto está sujeito a inúmeras variações quando é manuscrito. Então, podemos inferir que para cada caractere existe uma forma abstrata, não apenas uma imagem, mas um conjunto de padrões que a representam [14].

A forma abstrata precisa ser simples e única para permitir que qualquer variação possa ser reconhecida facilmente. A

escolha do classificador ocorreu por seu destaque no mercado e possuir capacidade de separar classes não lineares. Por decisão, optou-se pelo uso de apenas uma camada oculta na rede neural; capaz de classificar um grande número de problemas. O treinamento foi realizado com algoritmo de retro-propagação do erro com 20 entradas de dados.

De maneira a clarificar o processo de entrada de dados, a Figura 3 mostra as etapas do rastreio do marcador dedal em função do tempo de escrita da vogal A. O tempo para a captura do movimento foi configurado para disparar a execução do evento, a cada 3 segundos o algoritmo captura e processa a imagem em ciclos que denominamos de T.

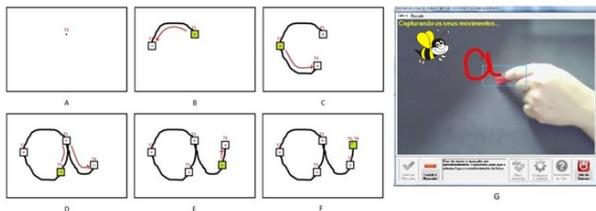


Figura 3. Pontos de Reconhecimento

No primeiro ciclo T1 (parte A), o software vai registrar o ponto em que o marcador está naquele momento, considerando ser a primeira execução do evento, nada mais será verificado. No segundo ciclo T2 (parte B), é verificada se houve alteração da imagem em relação ao ciclo anterior (T1). Caso exista o algoritmo registra novamente o ponto em que o marcador estava no momento da execução do evento T2 e prossegue. Nos ciclos seguintes essa verificação será refeita. O marcador só vai parar de se mover quando ele não tiver saído da pequena área ao redor do ponto registrado no ciclo, imediatamente anterior. Os demais ciclos seguem até a parte F, o algoritmo verifica se o marcador permaneceu dentro da área delimitada ao redor do ponto capturado no ciclo T5. Nesse caso, o marcador está parado, se ele ficar parado por mais de 3 segundos em algum ponto próximo, o algoritmo entende que a letra já foi escrita. Deste modo, começa a realizar o processo de reconhecimento pela rede neural. Para melhor entendimento a Figura 4, apresenta o Fluxograma da ferramenta, desde o processo de *Finger Tracking* até o momento de reconhecimento da letra.

Nota-se inicialmente a câmera é ativada e consequentemente o Touchless que irá avaliar onde o marcador estará se movendo, gerando um evento que envia a coordenada do marcador para que seja possível realizar o rastreamento do movimento. Essa coordenada é convertida em um objeto, que transforma esse movimento em um bitmap. A partir do momento que esse bitmap é gerado é realizada a verificação se foi gerada uma imagem ou um ruído, caso seja gerado uma imagem, esta é enviada para rede neural que verifica suas amostras e exibe uma resposta, no caso exibe uma imagem e emite áudio (pois o mesmo deve ser associado a letra escrita pelo usuário, para que seja feita a correlação de som e imagem) da letra reconhecida, caso seja gerado não seja gerado uma imagem, é enviado um ruído para rede neural, que não reconhece e volta a etapa de

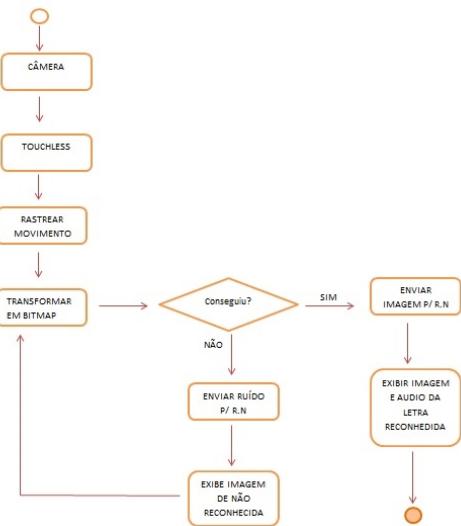


Figura 4. Fluxograma da Ferramenta

transformação do movimento em bitmap.

III. METODOLOGIA DE APLICAÇÃO

Com intuito de ser utilizado por crianças com dificuldades de leitura e escrita, as aplicações computacionais podem ser fortes aliadas no desenvolvimento de habilidades, como visto anteriormente, em tais atividades, reforçando a aprendizagem de forma estruturada e multissensorial [16]. As aplicações computacionais permitem o aluno revisar constantemente um assunto num trabalho individualizado no seu ritmo. Assim sendo, é um ambiente favorável a aplicação da tecnologia, que fornece um conforto maior para o aluno e facilita a quebra de barreiras tecnológicas e sua aceitação. Neste estudo, primeiramente, realizamos uma pesquisa bibliográfica, fundamentada por material já produzido sobre o assunto e pautado em reflexões que permitem repensar as práticas do professor com os alunos com problemas de leitura e escrita, disléxicos. Seu objetivo é oferecer elementos que possam contribuir com discussões posteriores e permear novas ações no software produzido.

No intuito de proporcionar um trabalho individual mais efetivo, escolhemos os materiais e métodos mais simples para obtenção dos resultados do experimento. Para a preparação do local do experimento usamos uma webcam, um anteparo para colocar a webcam a uma altura de 75 cm do foco da mesa. Recomendamos que seja colocado uma folha de EVA para deixar a mesa mais confortável para os alunos que a utilizarão. A webcam está posicionada em cima da mesa com ângulo de 45° para baixo, para que a mesma possa capturar os movimentos do usuário, e por fim a mesa adaptada, que está revestida por um EVA preto, onde será projetada o software e o que o usuário está fazendo.

Considerando que o uso do software em ambientes de ensino escolar e clínicas de fonoaudiologia ainda não foi realizado, os testes foram feitos com dez usuários desconsiderando gênero,

idade, classe social e considerando apenas se o mesmo é ou não alfabetizado. Para estabelecer os parâmetros de escrita cursiva para o treinamento da rede neural. Após este teste inicial de treinamento passamos a contar com 2 usuários reais, sendo estes portadores de algum扰urbo de linguagem, porém já alfabetizados. Para realizar os testes, procuramos apoio em escolas que indicaram estes alunos.

A. hipóteses testadas

Os testes foram realizados em uma escola com 2 alunos com dislexia que apresentam dificuldades na leitura e escrita, estes alunos possuem idade entre 7 e 9 anos. Os testes envolvem a realização de tarefas de leituras de texto aplicando diferentes níveis de dificuldade.

A ferramenta foi apresentada aos usuários de maneira simples, clara e objetiva, sendo colocada uma fita adesiva de coloração vermelha na ponta do dedo indicador de cada um, informado que este seria o lápis para escrever a letra, posteriormente fez uma apresentação de como funciona a ferramenta e partir disto cada um dos usuários realizou o teste. Nos testes realizados até o momento, os alunos destacaram a importância das funcionalidades implementadas e que a ferramenta pode auxiliar suas tarefas do dia-a-dia. As hipóteses que foram testadas neste trabalho estão dispostas foram as seguintes:

H1 A Escrita rastreada pode ser reconhecida com uma taxa de erro que não interfira no percepção do aluno;

H2 A projeção de uma animação/imagem inserindo o lúdico no ambiente favorece o aprendizado e fixação da letra e do som;

H3 O tempo médio de reconhecimento da letra pela ferramenta e a resposta do usuário são favoráveis a aplicação do software;

H4 O Aluno percebe a letra desenhada antes ou depois da apresentação da imagem lúdica;

IV. RESULTADOS E DISCUSSÃO

Nesta seção expressamos os resultados encontrados de nossa pesquisa. Assim, na Figura 5 percebemos que o tempo de resposta do tempo de reconhecimento de todas as letras testadas (vogais) tendem a permanecer em 2,9 (s), e a função de erro usada para rede neural encontrar a entropia do algoritmo de padrões de letras se comporta como uma função logística apresentada. Tendo em vista apenas o treino da rede neural.

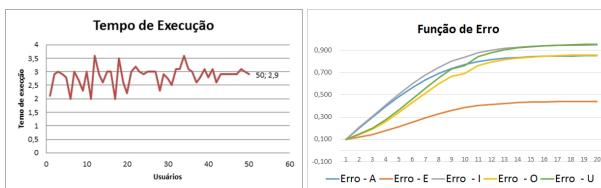


Figura 5. Tempo de Execução e Função de Erro

Apesar deste treino a Figura 7, observamos que mesmo a extenuante fase de treino com todas as letras vogais, ao aplicarmos os testes com os alunos, nossa rede neural ainda

está aquém do volume de acertos esperada. Sendo necessário uma reestruturação no tipo de algoritmo escolhido e aumentar o número de treinos da rede.

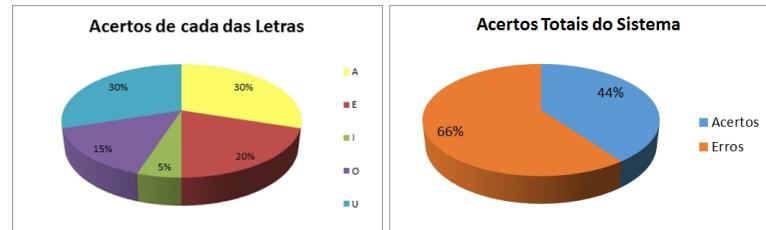


Figura 6. Taxa de Erros e Acertos do Algoritmo

De maneira a responder a hipótese **H1**, chegamos ao ponto de poder afirmar que é necessário a aplicação de outro algoritmo de inteligência artificial. Mesmo a rede neural sendo um ferramenta muito poderosa quando se trata de reconhecimento de padrões deve-se tomar um extremo cuidado, pois ela pode facilmente aprender a reconhecer apenas os dados alimentados na fase de treinamento. Este é um dos principais problemas na configuração da rede uma vez que, durante a fase de treinamento, pode-se observar altas porcentagens de acerto enquanto na fase de execução este índice pode cair consideravelmente se a rede não tiver generalizado corretamente, ou seja, aprendido apenas as características inerentes à letra desenhada. Portanto este hipótese ainda não está confirmada por este resultado.

Assim para responder **H2** apresentamos a ferramenta especificando que cada usuário deveria escrever as cinco vogais respectivamente e repetidas vezes, e como todos eram alfabetizados não foi necessário apresentar cada letra separada, apenas um quadro com as cinco vogais juntas como observado na Figura 7 , os alunos escolhidos apresentaram uma absorção da informação e a retenção do processo de uso do ferramenta de maneira muito rápido, todavia, percebemos que mesmo com seu entusiasmo os alunos apresentaram uma dificuldade inicial a ferramenta, uma retração, talvez pela tecnologia, e em seguida foram experimentando o ambiente. Portanto, confirmamos esta hipótese mediante a este resultado. Ressaltando que o conhecimento da teoria que está por trás dos resultados é decididamente essencial. De nada adianta utilizar uma nova ferramenta de trabalho se a consequência for o desinteresse do aluno em compreender e refletir sobre aquilo que ele próprio está fazendo [1]. Ao observar a Figura 6, podemos considerar respondida da **H3** pelo fato de na média termos um reconhecimento de em até 3 segundos de amostragem. Desta mesma forma, notamos o aluno percebe a letra ao desenhá-la, todavia acreditamos que este conhecimento já está assentado, porém não resgatado a tempo, pois a sensação de algo novo apresentado à tela prendeu a atenção de ambos os alunos.

V. CONSIDERAÇÕES FINAIS

Apesar de considerarmos que a solução proposta possa ser eficiente e ser de baixo custo, sua utilização nas escolas, requer avaliação de outros fatores, que ultrapassam o escopo deste trabalho; salientamos que muitas particularidades ainda não foram avaliadas e serão posteriormente. Acreditamos que

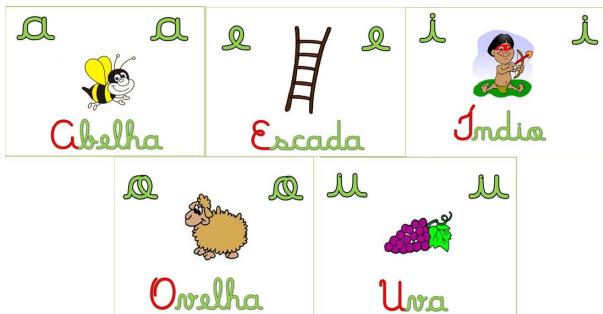


Figura 7. Apresentação das Vogais aos usuários

apesar de ser apoiar em diversas metodologias para auxiliar no ensino e facilitar a aprendizagem dos alunos, portadores de distúrbios de linguagem, ainda estamos longe de encontrar uma solução final e completa. Devido aos resultados encontrados nos testes do sistema, será necessário um estudo aprofundado em técnicas distintas de reconhecimento de movimento. O modelo de Rede Neural também será alterado, passando de uma camada simples para múltiplas camadas. O reconhecimento das letras deixará de ser apenas das vogais, e sim para todo o alfabeto e posteriormente a junção silábica, e os numerais. Acreditamos então que com a incorporação de novas técnicas e novos testes deixará a ferramenta mais robusta e completa para auxiliar no tratamento de distúrbios da linguagem. Existem muitos trabalhos que discorrem sobre ferramentas computacionais de apoio, no entanto, as tecnologias de interação natural ainda se expandem para esta aplicação. Não podemos restringir os estudos ao uso nesta proposta, contudo, permitir o avanço demarcado deste trabalho. Os estudos acerca em trabalhos que farão a combinação de tecnologias de captura de movimento mais acessíveis e mais eficientes; redução no tempo de desenvolvimento de produtos aplicados ao ambiente escolar e os distúrbios de aprendizagem. Destacamos, aqueles que refletem na aprendizagem da leitura e da escrita, para tornar a atividade mais atraente e envolvente, na busca de facilitar a absorção do conceito e do referencial simbólico do desenho de cada letra. Por fim, acreditamos que nosso software poderá ser utilizado com ferramenta de apoio, para auxiliar os psicopedagogos e fonoaudiólogos, no apoio ao ensino e tratamento dos sintomas do distúrbio de linguagem. A ferramenta apresentada trabalha com a criança da mesma maneira que a professora em sala de aula, pode ser utilizado como exercício adicional, uma vez que reforça o som, a partir da associação de uma personagem e ao desenho correspondente da sua letra.

VI. REFERÊNCIAS

REFERÊNCIAS

- [1] A. D. Alves, A. L. F. Bueno, and S. P. Bean, “Uma abordagem lúdico-computacional dos conteúdos de matemática: novas experiências,” *Extensão: Revista Eletrônica de Extensão*, vol. 2, no. 2, 2005.
- [2] C. F. André, “Guia de tecnologias educacionais,” *Brasília: Ministério da Educação, Secretaria de Educação Básica*, 2009.
- [3] R. T. Azuma *et al.*, “A survey of augmented reality,” *Presence*, vol. 6, no. 4, pp. 355–385, 1997.
- [4] S. F. Brasil, “Lei de diretrizes e bases da educação nacional (lei n. 9394/96),” 1996.
- [5] A. G. S. Capovilla, F. C. Capovilla *et al.*, *Alfabetização: método fônico*. Memnon, 2004.
- [6] M. R. P. da Silva, “O processo da leitura e escrita na alfabetização e séries iniciais: O desafio do pedagogo no desenvolvimento do aprendiz com dislexia,” *Revista da Escola de Administração Pública do Amapá*, vol. 2, no. 1, pp. 150–166, 2014.
- [7] E. F. Damasceno, E. A. L. Junior, and A. Cardos, “Uma avaliação heurística sobre um sistema de captura de movimentos em realidade aumentada,” *Journal of Health Informatics*, vol. 4, no. 3, 2012.
- [8] S. das Dores Rodrigues, M. J. M. G. de Castro, and S. M. Ciasca, “Relação entre indícios de disgraphia funcional e desempenho acadêmico,” *Revista CEFAC*, vol. 11, no. 2, pp. 221–227, 2009.
- [9] D. S. de Melo and H. N. Schneider, “Um computador por aluno: Mesclando de experimentos sobre a integração da tecnologia no âmbito escolar,” *International Journal of Knowledge Engineering and Management (IJKEM)*, vol. 3, no. 7, pp. 87–110, 2014.
- [10] J. M. M. Duarte *et al.*, “A multimédia na dislexia: tecnologia multimédia na reeducação da dislexia,” 2013.
- [11] C. Ellwanger, C. P. Santos, and G. J. Moreira, “Padrões de interface, gameificação e mobilidade no desenvolvimento de aplicações educacionais,” in *Proceedings of the 13th Brazilian Symposium on Human Factors in Computing Systems*. Sociedade Brasileira de Computação, 2014, pp. 365–368.
- [12] J.-M. Gottfried, J. Fehr, and C. S. Garbe, “Computing range flow from multi-modal kinect data,” in *Advances in Visual Computing*. Springer, 2011, pp. 758–767.
- [13] T. G. d. Medeiros and E. R. C. Oliveira, “A influência da consciência fonológica em crianças alfabetizadas pelos métodos fônico e silábico,” *Rev. CEFAC*, vol. 10, no. 1, pp. 45–50, 2008.
- [14] C. S. Menezes, L. L. Almeida, F. A. d. Silva, M. A. Pazoti, and A. O. Artero, “Redes neurais classe modular aplicadas no reconhecimento de caracteres manuscritos,” in *Colloquium Exactarum*, vol. 6, no. 2, 2014, pp. 170–183.
- [15] R. Mousinho and M. Goldfeld, “Desenvolvimento da leitura e escrita e seus transtornos,” *Goldfeld M. Fundamentos em fonoaudiologia: linguagem*, vol. 2, pp. 39–59, 2003.
- [16] A. Ramos, C. Tomás, C. Cunha, F. Machado, S. Miranda, and A. Osório, “Integração curricular das tics: Ponto de vista dos professores,” *Osório, A., e Puga, M.(Coords.) (eds.) As TIC na Escola. Universidade do Minho, Braga*, pp. 33–47, 2007.
- [17] A. G. Sebra and N. M. Dias, “Métodos de alfabetização: delimitação de procedimentos e considerações para uma prática eficaz,” *Revista Psicopedagogia*, vol. 28, no. 87, pp. 306–320, 2011.
- [18] S. Shaywitz and V. Figueira, *Entendendo a dislexia: um novo e completo programa para todos os níveis de problemas de leitura*, 1st ed. Artmed, 2006.
- [19] A. W. Smeulders, D. M. Chu, R. Cucchiara, S. Calderara, A. Dehghan, and M. Shah, “Visual tracking: An experimental survey,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 36, no. 7, pp. 1442–1468, 2014.
- [20] M. Soares, “A reinvenção da alfabetização,” *Presença pedagógica*, vol. 9, no. 52, pp. 15–21, 2003.
- [21] S. M. Stinchfield, *Speech disorders: A psychological study of the various defects of speech*. Routledge, 2013, vol. 20.
- [22] A. Valli, “Notes on natural interaction,” retrieved from on Jan, vol. 5, no. 2012, p. 80, 2005.
- [23] M. Young, “The pinhole camera,” *The Phys. Teacher*, pp. 648–655, 1989.

Considerações de Usabilidade para Interação por Voz em Aplicações Móveis de Realidade Aumentada

Usability Considerations for Voice Interaction in a Mobile Augmented Reality Application

Tiago Araújo*, Brunelli Miranda†, Carlos Santos*, Nikolas Carneiro*, Marcelle Mota*†, Nelson Neto*†, Bianchi Serique*†

*Programa de Pós-Graduação em Ciência da Computação - PPGCC

†Faculdade de Computação – FACOMP

Universidade Federal do Pará, Belém, Pará, 66075-110

Resumo – O aumento da popularidade de aplicativos de realidade aumentada móvel tornou possível cobrir novos modos de interação com o seu conteúdo. O uso de linguagem natural torna possível interagir dentro de um aplicativo de uma forma mais intuitiva. O uso da voz, embora não popular entre os pesquisadores está emergindo. Este artigo discute a inclusão de um serviço de reconhecimento de voz e as implicações deste em um aplicativo de realidade aumentada móvel.

Abstract—Increased popularity of mobile augmented reality applications made possible to cover new modes of interaction with the its content. The use of natural language makes possible interact within an application in a more intuitive way. The usage of the Voice although not popular among researchers is emerging. This paper discusses the inclusion of a voice recognition service and the implications of this in a mobile augmented reality application.

Keywords—realidade aumentada, voz, usabilidade, aplicação móvel, interação, Android, reconhecimento de voz

I. INTRODUÇÃO

A Realidade Aumentada Móvel (RAM) ganhou um importante destaque em múltiplas áreas de aplicação, dentre elas: engenharias, saúde, educação turismo, automação e outras. Este destaque está relacionado a vários fatores, como: o desenvolvimento tecnológico dos dispositivos móveis, aumento da popularidade e comercialização de *smartphones* e *tablets*, e a inovação e facilidades da Realidade Aumentada (RA).

Apesar do desenvolvimento tecnológico e da crescente popularidade da RAM serem motivações para sua ampla utilização, são os estudos na usabilidade das ferramentas RAM que podem manter uma boa experiência de uso, consolidando assim a longevidade da utilização dessas ferramentas nas mais diversas áreas.

Martínez *et al.* [1] apresenta alguns desafios para o desenvolvimento de aplicações de RAM. Dentre os desafios apresentados, pode-se destacar a escassez de padrões de

desenvolvimento e o pouco espaço para a apresentação de informações.

A escassez de padrões de desenvolvimento pode ser considerada tanto escassez de padrões para codificação dessas aplicações de RAM, assim como para a organização e *design* das interfaces gráficas. E o pouco espaço para apresentação das informações se dá pelo fato de que a RA necessita mostrar ao usuário o ambiente real acrescido de informações virtuais em sincronia [5], e isto, em uma tela de dispositivo móvel como um *smartphone*, que pode variar de 3 a 8 polegadas aproximadamente.

Ao dispor de um espaço limitado os projetistas de aplicações de RAM devem trabalhar arduamente para mostrar ao usuário a cena misturada (ambiente real acrescido de informações virtuais) e ainda apresentar-lhe as possíveis interações e *feedbacks* do sistema.

Inevitavelmente, o projetista terá que optar por deixar algumas informações escondidas (geralmente em menus que podem apresentar várias opções) infringindo alguns princípios de usabilidade, como a primeira heurística de Nielsen [2], em troca de uma melhor visibilidade do conteúdo principal e deixando a tela do usuário com menos elementos gráficos desnecessários. Entretanto, esconder funcionalidades em menus pode acarretar em um maior esforço cognitivo para encontrar e utilizar tais funcionalidades.

Um importante recurso que pode contribuir com a organização da interface gráfica no pouco espaço de tela é a utilização de *feedbacks* e comandos voz na qual o usuário pode interagir e receber informações utilizando a comunicação oral.

Outro ponto importante relacionado a utilização de aplicações móveis é o uso das mãos na aplicação [3]. Aplicações de RAM que necessitam que o usuário aponte sua câmera para uma direção específica, como é o caso das aplicações de localização de pontos de interesse por GPS, podem demandar que o usuário segure seu dispositivo com uma mão fixando o dispositivo em uma direção e tenha que interagir com a outra.

A situação de segurar o dispositivo com uma mão e interagir com a outra pode ser incômoda a medida que o dispositivo é maior e mais pesado, como os *tablets*. As interações por comando de voz podem inibir esse problema, fazendo com que o usuário consiga interagir com a aplicação mesmo que segurando o dispositivo com as duas mãos.

A utilização de comandos de voz em aplicações de RAM ainda é rara e há uma notável carência de estudos da utilização de voz em aplicações de RAM em dispositivos como *smartphones* e *tablets*.

Em uma pesquisa realizada nos principais indexadores de trabalhos acadêmicos da computação (IEEE *Xplore*, *Science Direct* e ACM *Digital Library*) foi constatada uma carência de estudos que contém RAM e comandos e *feedbacks* de voz.

Ao pesquisar com as seguintes palavras chaves: “*Mobile*”, “*Augmented*”, “*Reality*”, “*Voice*” e “*Speech*”, (sendo que *voice* e *speech* foram pesquisados separadamente por serem utilizadas como sinônimos) foram encontrados um total de 42 artigos, sendo 29 na IEEE, 13 na ACM e nenhum na *Science Direct*. Após a leitura dos resumos, remoção de artigos repetidos e remoção de artigos que não estão relacionados com assunto proposto, resultou um total de 12 artigos, sendo 9 na IEEE e 3 na ACM. A Figura 1 (a) mostra o resultado desta busca.

Após uma leitura mais aprofundada dos artigos selecionados, verificou-se que os artigos são sobre aplicações que utilizam comandos de voz, mas não há um estudo mais detalhado sobre os impactos de sua utilização.

A Figura 1 (b) mostra um gráfico dos artigos selecionados por área em que a voz e a RAM foram utilizados. Embora o número de aplicações com voz seja restrito existe uma ampla variedade de aplicações que vai do ensino e aprendizagem até as engenharias e robótica. A Tabela 1 apresenta as citações dos artigos selecionados na pesquisa por área.

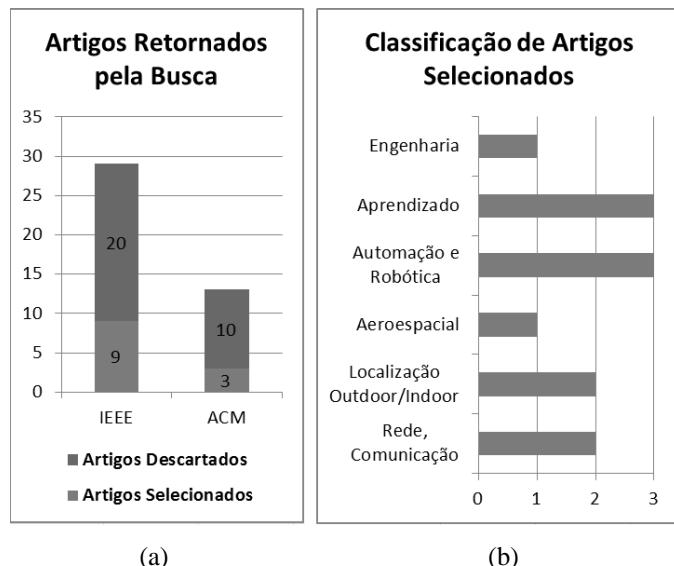


Fig. 1. Informações sobre os artigos pesquisados.

Rede, Comunicação	2	[4], [5]
Localização	2	[6], [7]
Outdoor/Indoor		
Aeroespacial	1	[8]
Automação e Robótica	3	[9], [10], [11]
Aprendizado	3	[12], [13], [14]
Engenharia	1	[15]

TABELA 1. ARTIGOS ENCONTRADOS

Ao realizar a mesma pesquisa, porém sem as palavras *voice* e *speech* foram retornados um total de 1927 resultados, que se comparado aos 42 resultados que incluíram a voz representam bem a ausência de estudos de RAM em conjunto com a voz.

Este trabalho apresenta algumas considerações da utilização de voz em uma aplicação de RAM, e as implicações arquiteturais e interface gráfica provenientes da utilização de voz em uma aplicação de RAM. A aplicação utilizada como estudo de caso chama-se ARGuide e trata-se de uma aplicação de localização e exploração de pontos de interesse por Navegador de RA, Mapa geográfico e leitura de códigos QR Code.

II. REALIDADE AUMENTADA MÓVEL

A RA trata da inserção de elementos virtuais no cenário real, visando aumentar a percepção do usuário sobre os elementos que o cercam [16]. Com a inovação de dispositivos móveis, que agora possuem processadores e componentes mais avançados, capazes de suportar melhor a renderização de objetos 3D e o processamento de informações, surgiu uma boa oportunidade para o desenvolvimento da RAM.

A RAM tem despertado interesse tanto no desenvolvimento de aplicações na área, quanto para a utilização dessas aplicações. Como mostrado em [17], já existem algumas bibliotecas e Kits de Desenvolvimento de Software (SDK), que oferecem funcionalidades apropriadas à RA e facilitam o desenvolvimento de aplicações RAM.

Na pesquisa realizada em [17], usuários foram convidados a utilizar várias aplicações RAM em diversos contextos, para citar alguns, ambientes *indoor*, *outdoor*, com vários pontos de interesse e em grandes áreas urbanas. Resultados mostraram que há interesse por parte dos usuários em aplicações RAM, mostrando uma aceitação positiva, porém, ainda são necessários estudos em como incluir essa tecnologia em serviços relevantes para o dia a dia do usuário.

III. INTERAÇÃO POR VOZ

A voz é um dos meios de comunicação mais utilizados pelos seres humanos, estando entre as três formas de comunicação (voz, gestos e expressões faciais) mais utilizadas no dia-a-dia de uma pessoa [18]. A voz, quando utilizada em uma aplicação, está inserida no conceito de linguagem natural.

A interação por linguagem natural pode ser definida como a comunicação entre humano e máquina utilizando uma linguagem familiar ao humano [19]. A área de Interação Humano-Computador (IHC) atribui papel importante para esse tipo de interação, devido aos benefícios proporcionados ao usuário, entre eles, o fato das ações serem mais intuitivas, portanto minimizando o esforço cognitivo e permitindo que o usuário se concentre na tarefa e não na interação em si [20].

O uso de voz em *smartphones* e *tablets* também apresenta essas vantagens para o usuário. É cada vez mais comum que as aplicações móveis ofereçam mais funcionalidades ao usuário. Porém, esse aumento de funcionalidades exige que as aplicações usem menus mais complexos, que podem acabar escondendo algumas funcionalidades do usuário ou demandar um tempo considerável até que a funcionalidade desejada seja encontrada. Nesse cenário, a interação por voz pode facilitar a busca do usuário por uma funcionalidade, ou até mesmo servir como atalho para a execução da funcionalidade em questão [21].

Apesar da utilização de comandos por voz poder variar de acordo com o tipo de aplicação, de maneira geral, o reconhecimento de voz nas aplicações segue um modelo padrão, que pode ser observado na Figura 2.

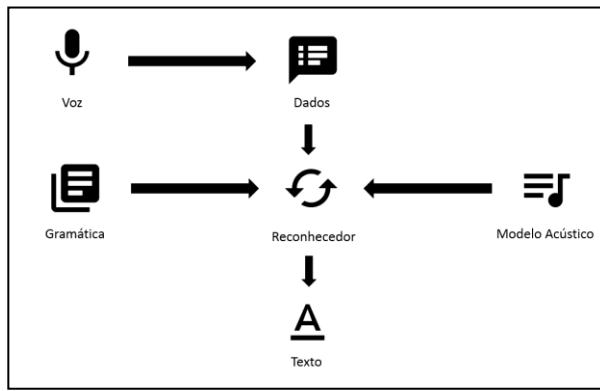


Fig 2. Funcionamento padrão de um modelo de reconhecimento de voz

No modelo apresentado, o usuário entra com a informação em forma de comando de voz, que é repassada para o reconhecedor, onde ela é analisada juntamente com a gramática e modelo acústico empregados na aplicação. O resultado desse processo é então retornado ao usuário em forma textual.

A interação por voz apresenta uso prático em diversos cenários, porém ainda há poucos estudos que a utilizam juntamente com a RAM, portanto, apresenta necessidade de estudos a seu respeito nesse área.

IV. ARGUIDE

O ARGuide é uma aplicação de RAM que utiliza conteúdos georeferenciados como base para a navegação do usuário [22]. A aplicação permite que o usuário possa explorar pontos de interesse (*Points of Interest - POI*) alternando entre várias visões coordenadas para acessar o conteúdo disponível. O conteúdo associado a cada POI pode ser multimídia.

As visões podem ser entendidas como diferentes formas de ver os POIs, sendo elas coordenadas significa que as

alterações realizadas em uma visão refletem em todas as outras, dando essa sincronia ao usuário faz com que ele perceba a mesma informação por diferentes perspectivas. As visões disponíveis na aplicação são as seguintes:

1) **Mapa**: permite a visualização em um mapa geográfico dos POI da aplicação;

2) **Navegador RA**: utiliza a posição do POI para definir marcadores na câmera do aplicativo correspondentes a posição no mundo real;

3) **Lista de POI**: Uma visão que organiza os POI em forma de lista;

4) **Visualização de POI**: mostra o conteúdo do POI em uma janela, organizando o texto e as mídias disponíveis.

Neste trabalho, o ARGuide recebeu um serviço de reconhecimento por voz para ajudar na navegação e visualização de conteúdo.

V. CONSIDERAÇÕES SOBRE A USABILIDADE DA UTILIZAÇÃO DE VOZ NO ARGUIDE

Incluir uma nova forma de interação em uma aplicação de RAM pode acarretar em mudanças no modo em que essa aplicação exibe o conteúdo para o usuário. Lee e Grice [23] identificaram os problemas relacionados à integração de comando de voz em aplicações móveis, e sugerem que esses problemas devem ser tratados desde o início da integração da aplicação com um serviço de reconhecimento de voz.

Para definir quais mudanças devem ser feitas para inserir um serviço de reconhecimento de voz no ARGuide, foi definido um processo para a obtenção de das mudanças que devem ser feitas.

Juntamente com o estudo da literatura, para dar apoio na inserção de interações por voz nessa aplicação, a ajuda de uma especialista em IHC foi necessária. Entrevistas foram realizadas para gerar mudanças no desenvolvimento de forma em que as funções de interações por voz estivessem sendo adicionadas. Foram realizadas duas entrevistas, e entre cada entrevista uma nova versão do ARGuide foi gerada para passar por nova avaliação. A Figura 3 mostra o processo de obtenção das modificações a serem feitas no ARGuide para melhorar a usabilidade quando forem introduzidas interações por voz.



Fig. 3. Processo de inserção de interação por voz no ARGuide

A. 1^a Entrevista

Sem componentes funcionais de reconhecimento de voz, essa versão recebeu avaliação baseada em mudanças de interface simples que facilitariam o uso do reconhecedor de voz. A primeira entrevista foi conduzida com explicação prévia da ferramenta e observação da especialista enquanto utilizava a aplicação, mostrando as funcionalidades no decorrer da entrevista.

As considerações feitas na primeira entrevista foram com elementos de interface. A especialista notou que vários

componentes de interface estavam sem indicação de que poderiam ser acionados por um comando de voz. Elementos como os marcadores do mapa ou ícones de navegação tiveram que ser transformados para elementos textuais.

Um exemplo de retorno obtido através da entrevista está na Figura 4. Essa figura mostra dois aspectos na interface que sofreram alteração em função da primeira entrevista. A primeira observação é a troca dos ícones para texto. Os ícones que representavam ações da aplicação foram substituídos por palavras que podem ser usados como comandos de voz. Os marcadores utilizados para identificar os POI foram trocados rótulos de texto.

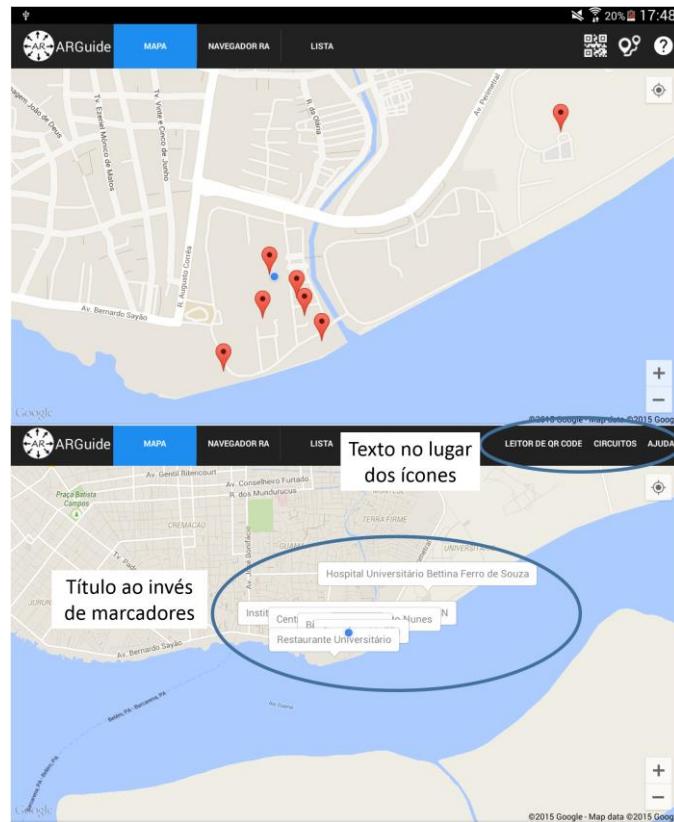


Fig. 4. Alterações realizadas após a primeira entrevista.

B. 2^a Entrevista

Para a segunda entrevista, foi implementado um serviço de reconhecimento de voz, a implementação utilizou o reconhecedor de voz padrão do Android [24]. É um reconhecedor de voz que funciona online, a recepção e transmissão dos dados é feita no aparelho e enviada para o reconhecedor, que retorna o texto correspondente ao que foi falado. Com o conhecimento da ferramenta, a especialista focou a avaliação no reconhecedor de voz e uso da gramática escolhida.

Novas melhorias puderam ser realizadas com o reconhecedor funcionando. As considerações feitas nessa entrevista foram relacionadas a interpretação e reconhecimento de voz. Comandos base tiveram que ser modificados para acomodar facilidades da linguagem natural,

como relaxar comandos que eram somente verbos no infinitivo para substantivos e conjugações mais comuns. Como “sair”, que pode ser pronunciado “sai” ou “saia” que o reconhecedor entende como o mesmo comando. Outra mudança relevante é o controle de legendas laterais, para que o usuário saiba que comandos pode fazer com voz em determinada visão. A Figura 5 exemplifica uma transição com relação à legenda das ações mapeadas para certa visão.

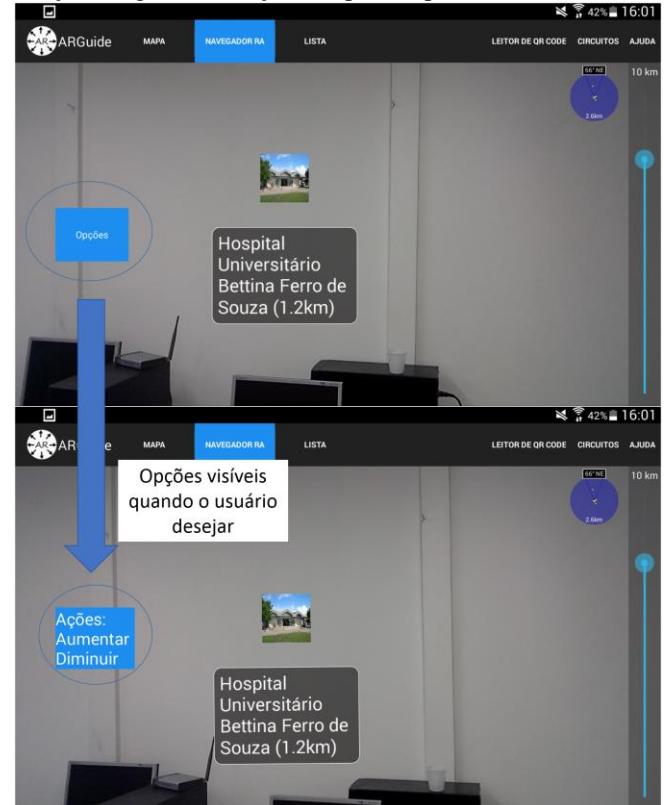


Fig. 5. Alterações realizadas após a segunda entrevista.

O usuário sempre tem disponível uma caixa de texto que abre com um comando de voz para que ele possa visualizar as ações naquela visão. Nesse exemplo, aumentar e diminuir o alcance de pontos de interesse no navegador RA. Em ações muito comuns, como mover uma lista, ou escolher um POI, os próprios marcadores visuais padrão podem indicar esses comandos.

VI. IMPLICAÇÕES DAINSERÇÃO DE VOZ NA APLICAÇÃO

O serviço de reconhecimento de voz no ARGuide afetou dois pontos principais a aplicação. Um dos pontos já foi discutido, no decorrer das entrevistas algumas mudanças visuais foram necessárias. O outro ponto afeta a arquitetura. A arquitetura do ARGuide é uma variação do modelo MVC, que tem possui uma View com várias visões coordenadas, coordenação essa feita por um Coordinator que intermedia as informações e mídias de determinado POI no Model em uma visão. A Figura 6 mostra a mudança na arquitetura do ARGuide com a adição de comandos de voz.

A inclusão de um serviço de voz é uma interface de coordenação, já que além de acessar os POIs, também

comanda outras visões. Esses comandos são paralelos aos comandos do toque do usuário, então a interação de voz passa a ser coordenada como se fosse o toque.

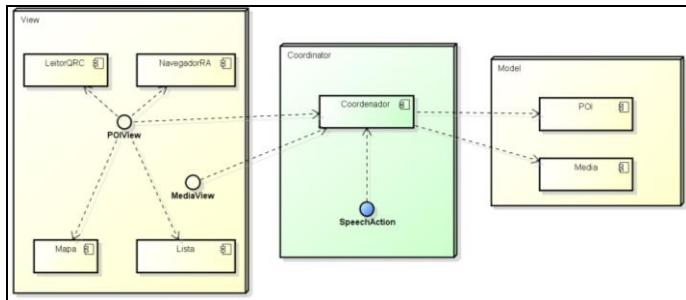


Fig. 6. Mudanças realizadas na arquitetura para acomodar o serviço de voz.

O componente SpeechAction é um conjunto de classes, interfaces e serviços que é responsável pelo conhecimento de voz, ele é o componente novo na arquitetura do ARGuide, e coordena variadas ações do aplicativo.

VII. SÍNTESE DOS RESULTADOS

As entrevistas realizadas para gerar a implementação das interações de voz com o ARGuide trouxeram um conjunto de considerações que apesar de se aplicarem a uma parte da aplicação podem ser estendidos para toda aplicação.

A Tabela 2 define as principais considerações que foram geradas através desse método. As colunas seguintes mostram em que sentido, se pela Interface Gráfica, ou pela Voz esse procedimento deve ser atendido.

TABELA 2. CONSIDERAÇÕES REALIZADAS PARA INTERFACE E VOZ

Nº	Considerações	Interface Gráfica	Voz
1	Legendas sempre presentes	X	
2	Reconhecedor identificando variações		X
3	Legendas controladas pelo usuário	X	X
4	Ações de configuração sempre acessíveis		X
5	Ações comuns claras para o usuário	X	X
6	Elementos de conteúdo identificados através de texto	X	
7	Ícones devem ser trocados por palavras	X	
8	Ações gerais da aplicação devem ser identificados por palavras já conhecidas pelo público alvo		X

Cada consideração pode ser definida como:

1) **Legenda presente:** A legenda para ações gerais sempre deve estar presente para o usuário.

2) **Reconhecedor:** O reconhecedor deve ter uma gramática que aceita variações, não somente uma palavra em cada comando.

3) **Controle de legenda:** O usuário deve controlar a visualização de legendas em cada visão da aplicação.

4) **Ações de configuração:** Para comandos de voz, ações de configuração devem ser sempre acessíveis.

5) **Ações comuns:** Ações comuns da aplicação devem ser facilmente identificados por elementos de interface gráfica.

6) **Identificar conteúdo:** Todo conteúdo disponível por toque deve ter identificação para acessar com voz.

7) **Ícones por palavras:** Substituir ícones por palavras que serão comandos.

8) **Identificar ações gerais:** Ações gerais da aplicação devem ter comandos padrão ou fáceis de memorizar.

A Figura 7 mostra o efeito das considerações em outra parte da aplicação, de maneira mais geral.



Fig. 7. Alterações realizadas após gerar definições gerais.

A troca de ícones que foi feita para as opções foi estendida para a visão de conteúdo do POI. A barra de rolagem sempre presente define a ação de subir ou descer o conteúdo que está sendo mostrado, que serve como exemplo, que mesmo sem legenda é uma ação comum para o usuário mover uma lista para cima ou para baixo, identificada esta ação pela barra de rolagem.

CONSIDERAÇÕES FINAIS

O uso de voz em aplicações RAM se mostra muito relevante no futuro, já que o uso de linguagem natural para interação com conteúdo é um caminho que está se tornando mais simples com a popularização de aplicações e serviços que utilizam esse tipo de interação.

A contribuição gerada por esse trabalho são as considerações extraídas da implementação do serviço de reconhecimento de voz em uma aplicação de RAM. Essas considerações podem ser estendidas em futuras iterações do

ARGuide, com mais visões ou funções que poderão integrar a aplicação.

A continuidade desse trabalho será feita com testes de usuário e a avaliação dos resultados desse teste para mais uma interação da ferramenta. O teste com outros reconhecedores de voz também está previsto para avaliar a acurácia do reconhecedor na aplicação e como podem afetar e experiência do usuário.

REFERENCES

- [1] H. Martínez, D. Skournetou, J. Hyppola, S. Laukkanen, and A. Heikkilä, "Drivers and bottlenecks in the adoption of augmented reality applications," *Journal of Multimedia Theory and Application*, 2014.
- [2] NIELSEN, J. "Heuristic Evaluation". Em: NIELSEN, J.; MACK, R. L. *Usability Inspection Methods*. New York, NY: Katherine Schowalter, 1994. Cap. 2
- [3] DE SÁ, M.; CHURCHILL, E. F. "Mobile Augmented Reality: A Design Perspective". Em: HUANG, W.; ALEM, L.; LIVINGSTON, M. A. *Human Factors in Augmented Reality Environments*. New York, NY: Springer, 2013. Cap. 6, p. 139-164.
- [4] Tino Pyssyalo, Tapio Repo, Tuukka Turunen, Teemu Lankila, and Juha Röning. 2000. "CyPhone—bringing augmented reality to next generation mobile phones." Em *Proceedings of DARE 2000 on Designing augmented reality environments (DARE '00)*. ACM, New York, NY, USA pg. 11-21
- [5] Pulli, P.; Pyssyalo, T.; Metsavainio, J.-P.; Komulainen, O., "CyPhone-experimenting mobile real-time telepresence," in *Real-Time Systems*, 1998. Proceedings. 10th Euromicro Workshop on , vol., no., pp.10-17, 17-19 Junho 1998
- [6] Buti Al Delail, Luis Weruaga, and M. Jamal Zemerly. 2012. "CAViAR: Context Aware Visual Indoor Augmented Reality for a University Campus". Em *Proceedings of the The 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology Volume 03 (WI-IAT '12)*, Vol. 3. IEEE Computer Society, Washington, DC, USA, pg. 286-290
- [7] K. G. Srinivasa, Satvik Jagannath, P. S. Akash Nidhi, S. Tejesh, and K. Santhosh. 2014. Augmented Reality Application: Cloud based Augmented Reality Android application to "know your world better". In *Proceedings of the 6th IBM Collaborative Academia Research Exchange Conference (I-CARE) on I-CARE 2014 (I-CARE 2014)*. ACM, New York, NY, USA, , Article 15
- [8] Markov-Vetter, Daniela; Staadt, Oliver, "A pilot study for Augmented Reality supported procedure guidance to operate payload racks on-board the International Space Station," in *Mixed and Augmented Reality (ISMAR), 2013 IEEE International Symposium on* , vol., no., pp.1-6, 1-4 Oct. 2013
- [9] Giesler, B.; Salb, T.; Steinhäus, P.; Dillmann, R., "Using augmented reality to interact with an autonomous mobile platform," in *Robotics and Automation*, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on , vol.1, no., pp.1009-1014 Vol.1, 26 April-1 Maio 2004
- [10] Tang, L.Z.W.; Kian Sin Ang; Amirul, M.; Bin Mohamed Yusoff, M.; Chee Keong Tng; Bin Mohamed Alyas, M.D.; Joo Ghee Lim; Kyaw, P.K.; Folianto, F., "Augmented reality control home (ARCH) for disabled and elderlys," in *Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2015 IEEE Tenth International Conference on* , vol., no., pp.1-2, 7-9 Abril 2015
- [11] Igarashi, T.; Inami, M., "Exploration of Alternative Interaction Techniques for Robotic Systems," in *Computer Graphics and Applications, IEEE* , vol.35, no.3, pp.33-41, Maio-Junho 2015
- [12] Doswell, J.T., "Augmented Learning: Context-Aware Mobile Augmented Reality Architecture for Learning," in *Advanced Learning Technologies, 2006. Sixth International Conference on* , vol., no., pp.1182-1183, 5-7 Julho 2006
- [13] Doswell, J.T.; Blake, M.B.; Butcher-Green, J., "Mobile Augmented Reality System Architecture for Ubiquitous e-Learning," in *Wireless, Mobile and Ubiquitous Technology in Education, 2006. WMUTE '06. Fourth IEEE International Workshop on* , vol., no., pp.121-123, 16-17 Novembro 2006
- [14] Doswell, J.T., "Context-Aware Mobile Augmented Reality Architecture for Lifelong Learning," in *Advanced Learning Technologies, 2006. Sixth International Conference on* , vol., no., pp.372-374, 5-7 Julho 2006
- [15] Goose, S.; Sudarsky, S.; Zhang, X.; Navab, N., "SEAR: towards a mobile and context-sensitive speech-enabled augmented reality," in *Multimedia and Expo, 2002. ICME '02. Proceedings. 2002 IEEE International Conference on* , vol.1, no., pp.849-852 vol.1, 2002
- [16] Cawood, S., Fiala, M., "About Augmented Reality", *Augmented Reality: A Practical Guide*. Pragmatic Bookshelf, USA 2008, cap 1.
- [17] T. Olsson and M. Salo. "Online User Survey on Current Mobile Augmented Reality Applications". *Proceedings of the 2011 10th IEEE International Symposium on Mixed and Augmented Reality*. Washington: IEE Computer Society. 2011. pp 75-84.
- [18] A. Teixeira et al. "Speech-Centric Multimodal Interaction for Easy-To-Access Online Services – A Personal Life Assistant for the Elderly". 5th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion – DSAI, 2013.
- [19] B. Schneiderman. "Designing the User Interface: Strategies for Effective Human-Computer Interaction", 3rd Edition, pp 293-295.1998.
- [20] N. Vidakis, M. Syntychakis, G. Triantafyllidis and D. Akoumianakis. "Multimodal Natural User Interaction for Multiple Applications: The Gesture – Voice Example". *International Conference on Telecommunications and Multimedia – TEMU*, 2012.
- [21] L. Xia, K. Kai, W. Xiaochun and W. Dan. *Research and Design of the "Voice-Touch-Vision" Multimodal Integrated Voice Interaction in the Mobile Phone*, 2010.
- [22] Lee, K. B.; Grice, R. A., "The Design and Development of User Interfaces for Voice Application in Mobile Devices," Em *2006 IEEE International Professional Communication Conference*, pp.308-320, 23-25 Outubro 2006
- [23] C. Santos, N. Carneiro, B. Miranda, B. Serique. "Uma Aplicação de Realidade Aumentada Móvel para Ambientes Indoor e Outdoor". XV Workshop de Realidade Virtual e Aumentada, Marília – SP, pp 120-126, November 2014.
- [24] Google Inc., "Google Now", <https://www.google.com/landing/now/>. Dezembro 2014.

New Approaches of Feature Extraction from Fruits to Augmented Reality Systems

Priscila Alves Macanhã, Wilson Estécio Marcílio Junior, Danilo Medeiros Eler,

Rogério Eduardo Garcia, Ronaldo Celso Messias Correia

Faculdade de Ciências e Tecnologia, UNESP – Univ Estadual Paulista,

Presidente Prudente, Departamento de Matemática e Computação

Presidente Prudente/SP, Brazil

priscila.macanha@gmail.com; wilson_jr@outlook.com; {danieloeler, rogerio, ronaldo}@fct.unesp.br

Resumo—Augmented Reality is a technology that has been widely employed to provide details on demand. This is used in museums, mobile devices, TV programs, among others. Currently, with the evolution of both hardware and computer vision techniques, it is possible to develop Augmented Reality systems that do not require fiducial markers. Therefore, it is necessary to conduct research in computer vision area to improve the recognition of objects, which are used as if they were markers. This paper presents comparisons of feature extraction approaches from fruit in order to identify the best approach capable of identifying a fruit in a scene. Therefore, to propose a new feature extraction approach, this paper compared in the context of fruit recognition some feature extractors found in the literature for handwritten. Thus, this comparison led to new approaches of feature extractors combination resulting in a correct classification rate of 97.5% with a classifier Multilayer Perceptron. Finally, an Augmented Reality system is presented for the recognition of fruit and details demand are shown in the scene.

I. INTRODUÇÃO

O avanço da tecnologia tem permitido a inclusão de novas maneiras de como a informação pode ser transmitida para as massas. Uma dessas tecnologias é a Realidade Aumentada, permitindo que informações adicionais sejam inseridas em imagens reais para complementar a informação a ser transmitida, ou seja, a Realidade Aumentada combina o real com o virtual para potencializar e facilitar o entendimento da informação. Como as pessoas estão se adaptando a essa nova realidade que combina o real com o virtual, é notável o seu crescimento, por exemplo, programas de televisão tem utilizado esse recurso para facilitar a divulgação de informações sobre esporte, clima, entre outros.

O uso da Realidade Aumentada também tem crescido em diversas áreas de pesquisas, procurando facilitar a interpretação de conjuntos de dados [1], [2]. Geralmente, marcadores são utilizados para identificar o local onde a imagem virtual deve ser inserida na imagem real. Entretanto, alguns trabalhos não utilizam tais marcadores [3], [4], o que necessita de algoritmos de visão computacional e de reconhecimento de padrões para efetuar identificação e a classificação dos objetos de interesse.

Este trabalho tem como objetivo principal propor uma nova abordagem de extração de características de frutas para apoiar sistemas de Realidade Aumentada direcionados para o contexto de apresentar informações sobre frutas. Para atingir esse objetivo, foram pesquisadas e comparadas técnicas já existentes na literatura para o reconhecimento de caracteres

manuscritos, entretanto, neste trabalho, elas serão empregadas no contexto de reconhecimento de frutas. A principal contribuição deste trabalho são novas abordagens de extração que combinam técnicas já existentes. Uma dessas novas abordagens atingiu 97,5% de taxa de classificação correta em um conjunto de imagens de frutas. Finalmente, também é apresentado um sistema de Realidade Aumentada proposto para o reconhecimento de frutas utilizando essa nova abordagem de extração de características.

O restante do trabalho está organizado da seguinte maneira. A Seção II apresenta trabalhos relacionados ao reconhecimento de frutas. A Seção III apresenta abordagens e técnicas para extração de características de caracteres manuscritos, as quais serão empregadas neste trabalho no contexto de extração de características de frutas. A Seção IV apresenta as abordagens de extração de características utilizadas neste trabalho, as quais combinam as técnicas de extração apresentadas na Seção III. A Seção V apresenta os experimentos realizados com as abordagens propostas neste trabalho. A Seção VI apresenta o sistema de Realidade Aumentada proposto neste trabalho, o qual utiliza o melhor resultado obtido nos experimentos. Finalmente, A Seção VII apresenta as conclusões e trabalhos futuros.

II. TRABALHOS RELACIONADOS

Seng e Mirsaee [5] utilizaram características baseada em cor, em formato e no tamanho das frutas, em um sistema de reconhecimento automático. O sistema foi capaz de reconhecer as imagens das frutas, resultando em uma taxa de precisão de 90%.

Um outro sistema proposto por Ganeshan et al. [6] é baseado nas características de cor e textura de frutas. A taxa de precisão do reconhecimento somente por cores obteve uma média de 45,49%, enquanto o reconhecimento somente por texturas obteve uma média de 70,85% e o reconhecimento por cores e textura combinadas apresentou resultado médio de 86%.

Zawbaa et al. [7] também propuseram o reconhecimento de frutas baseado nas cores e no formato. Os melhores resultados foram obtidos para maçãs e laranjas, resultando em 85% de precisão para maçãs e 65% para laranjas.

Sengupta e Lee [8] propuseram um sistema de reconhecimento de frutas cítricas verdes em árvores frutíferas em ambientes naturais utilizando forma e textura. Os resultados apresentaram uma taxa correta de classificação de 81,7%.

Neste trabalho empregaremos extratores de características comumente utilizados para o reconhecimento de caracteres manuscritos com o intuito de aumentar a taxa de acerto dos sistemas de reconhecimentos de frutas. Esses extractores são apresentados na próxima seção.

III. EXTRATORES DE CARACTERÍSTICAS

Diante da infinidade de técnicas de extração de características existentes na literatura, este trabalho selecionou duas técnicas para embasar a pesquisa, as quais são descritas em seguida.

A. Contagem de pixels ativos

O algoritmo de zoneamento [9] consiste na contagem dos pixels ativos em uma dada região. Os pixels ativos neste contexto são aqueles que constituem o caractere. O resultado do algoritmo é um número inteiro correspondente a quantidade de pixels ativos encontrados. A técnica em questão foi abordada em sua definição padrão descrita a seguir e foram propostas outras duas variações da mesma.

1) Contagem de pixels ativos com Zoneamento: O método consiste na divisão da imagem em zonas de mesmo tamanho e posterior aplicação da técnica de contagem de pixels ativos em cada região. Os valores resultantes são agregados em um vetor único que será utilizado na classificação. Neste trabalho foram adotados cinco zoneamentos, mostrados na Figura 1, de modo a favorecer os diferentes formatos de caracteres. São eles: 1×3 , 3×1 , 3×6 , 6×3 e 6×6 .

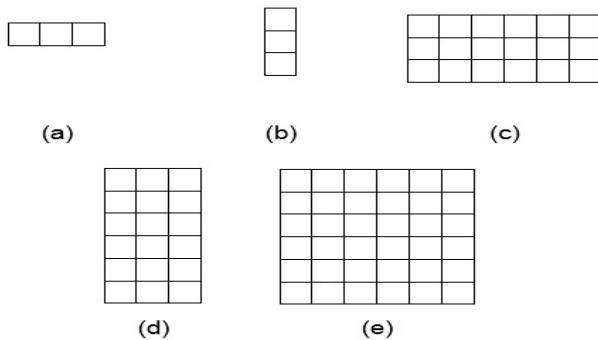


Figura 1: (a) Zoneamento 1×3 . (b) Zoneamento 3×1 . (c) Zoneamento 3×6 . (d) Zoneamento 6×3 . (e) Zoneamento 6×6 .

Os zoneamentos 1×3 e 3×6 têm o intuito de discriminar classes de caracteres que são assimétricos horizontalmente, tais como as letras I e T. Da mesma forma, os zoneamentos 3×1 e 6×3 pretendem diferenciar classes de caracteres assimétricos verticalmente, como as letras B e O. Por fim o zoneamento 6×6 foca na identificação de pequenas características discriminantes.

B. Distância Borda-Caractere

A técnica denominada distância borda-caractere consiste em, a partir da borda, realizar a contagem dos pixels do

fundo que antecedem os pixels do caractere nas oito direções indicadas na Figura 2 [10]. Assim sendo, esta técnica resulta em oito vetores de distância, compostos pela quantidade de pixels de fundo que antecedem o caractere em todos os caminhos possíveis seguindo as oito posições.

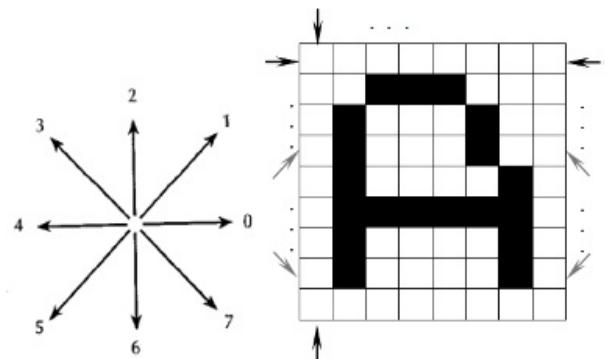


Figura 2: Oito posições para mensurar a distância borda-caractere. Adaptado de [10]

O conceito geral da técnica em questão deu origem a cinco métodos de extração de características propostos para o fim de análise. Sua abordagem original é apresentada a seguir.

IV. ABORDAGEM DE EXTRAÇÃO DE CARACTERÍSTICA PROPOSTA

Este trabalho inclui a implementação de todas as técnicas de extração de características apresentadas na Seção III. Além destas, foram propostas e implementadas outras sete abordagens descritas em seguida. É importante salientar que esses extractores são comumente empregados no reconhecimento de caracteres, mas neste trabalho eles serão empregados no reconhecimento de frutas.

A. Contagem de pixels ativos com diversos Zoneamentos

Neste método, o algoritmo descrito na seção anterior, correspondente à contagem de pixels ativos, é aplicado aos cinco zoneamentos descritos na mesma seção. Os vetores resultantes, ao final, são agrupados para então formar o vetor final de características. Assim sendo, o vetor será composto de 78 ($(1 \times 3) + (3 \times 1) + (3 \times 6) + (6 \times 3) + (6 \times 6)$) descriptores.

B. Contagem de pixels ativos para diversos Zoneamentos com Transformada de Fourier

Este método consiste na aplicação da Transformada Rápida de Fourier no vetor resultante da técnica de contagem de pixels ativos com diversos zoneamentos. Este procedimento resulta em um vetor de mesma dimensão que o vetor de aplicação, cujos elementos sintetizadores da imagem estão localizados em seus extremos. Deste modo, são tomadas três características de cada extremo, totalizando seis características que irão compor o vetor de característica final.

C. Distância Borda-Caractere com Transformada de Fourier no vetor de distância único

Este método aplica a técnica Distância Borda-Caractere agregando os vetores de distância resultantes da técnica em um vetor único. O vetor sequente é submetido a aplicação do algoritmo de Transformada Rápida de Fourier, que resulta em um vetor com as menores frequências localizadas em suas extremidades. Feito isto, são tomadas três características de cada um dos extremos, totalizando seis descritores que irão compor o vetor de características finais.

D. Distância Borda-Caractere com Transformada de Fourier em cada vetor de distância

Neste método os vetores de distância resultante da aplicação da técnica Distância Borda-Caractere são submetidos ao algoritmo de Transformada Rápida de Fourier, para o qual são tomadas três características dos extremos de cada vetor, totalizando seis descritores para cada posição, que comporão o vetor de características. Assim sendo o vetor final conterá 48 (6 x 8 posições) características.

E. Distância Borda-Caractere Pura com Zoneamento

O método em questão consiste na divisão da imagem em zonas de tamanhos iguais, e posterior aplicação da técnica Distância Borda-Caractere Pura em cada zona. Os vetores resultantes da aplicação da técnica são agrupados em um vetor único que será usado na classificação.

F. Distância Borda-Caractere com Transformada de Fourier no vetor de distância único de cada zona

Este método, como o anterior, exige prévia divisão da imagem em zonas de tamanhos iguais. Feito isto, é aplicada a técnica Distância Borda-Caractere em cada zona, para a qual os vetores de distância resultantes são agrupados em um vetor único. Logo, os vetores únicos de cada zona são submetidos à Transformada Rápida de Fourier, tomando-se duas características dos extremos do vetor resultante. Esta técnica é aplicada para os cinco zoneamentos descritos anteriormente.

G. Distância Borda-Caractere com Transformada de Fourier em cada vetor de distância de cada zona

Após a divisão da imagem em zonas de tamanhos iguais, é aplicada a técnica Distância Borda-Caractere em cada zona, onde cada vetor de distância, correspondente às oito direções, é submetido à Transformada Rápida de Fourier, tomando-se três características de cada extremo do vetor resultado da aplicação da Transformada de Fourier, totalizando seis descritores. Os valores selecionados são armazenados em um vetor que será usado na classificação. Esta técnica, como a anterior, é aplicada para os cinco zoneamentos descritos anteriormente.

A próxima seção apresenta os experimentos realizados com os vetores de características computados com as abordagens apresentadas nesta seção.

V. EXPERIMENTOS E RESULTADOS

Esta seção apresenta os experimentos realizados com os espaços de características computados durante as etapas de extração de características que seguiu as abordagens apresentadas na seção anterior. Finalmente, são apresentados alguns dos resultados obtidos com o sistema de Realidade Aumentada que utiliza o melhor espaço de características deste experimento.

Neste trabalho foram aplicados os espaços de características diretamente ao sistema Weka¹, utilizando os classificadores: Perceptron Multicamadas (MLP) e k-NN; ambos com as configurações padrão do WEKA.

Para compor os conjuntos de dados utilizados nos experimentos demonstrados a seguir, foram selecionadas 163 imagens de frutas, divididas em 15 classes diferentes: abacaxi, abacate, banana, cereja, damasco, kiwi, laranja, limão, maçã, manga, melancia, melão, morango, pera e uva. Essas imagens foram adquiridas na Internet, por meio da ferramenta *Google Imagens* e do site ViCoS [11]. Foram realizadas buscas por imagens de frutas individuais, com exceção da banana que tiveram imagens de pencas misturadas e, todas as imagens possuíam o fundo branco.

Nos experimentos, cada conjunto de dados foi dividido em duas partes, sendo 66% das 163 imagens (108 imagens) destinadas à fase de treinamento e 34% das 163 imagens (55 imagens) destinadas à fase de teste. A Figura 3 exemplifica uma parte do conjunto de imagens utilizado.

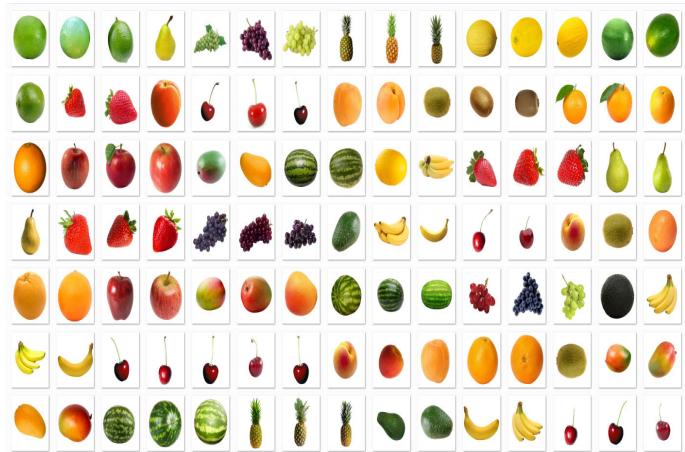


Figura 3: Exemplos das imagens de frutas utilizadas nos experimentos.

A Tabela I apresenta a porcentagem de taxa correta de classificação dos espaços de características comparados neste trabalho. É possível notar que o melhor espaço de características foi aquele computado com a abordagem proposta neste trabalho com a combinação das técnicas de zoneamento, de distância borda caractere e da aplicação da Transformada de Fourier (Distância Borda Caractere com Zoneamento 3x3 com Transformada de Fourier no vetor concatenado²).

¹O Weka é uma sistema composto por vários algoritmos de aprendizado de máquina para tarefas de mineração de dados. Ele está disponível em <http://www.cs.waikato.ac.nz/ml/weka/>.

Tabela I: Comparação dos diferentes espaços de características. As taxas de classificação foram computadas com os classificadores Perceptron Multicamadas (MLP) e k-NN.

Descrição	MLP	k-NN
Zoneamento 3x3	73,62%	63,80%
Zoneamento 3x3 e 6x6 com Transformada de Fourier no vetor concatenado	81,59%	63,80%
Zoneamento 3x3 e 6x6 com Transformada de Fourier em cada vetor	49,08%	47,85%
Distância Borda Caractere	16,56%	28,22%
Distância Borda Caractere com Transformada de Fourier no vetor de concatenado	92,63%	63,19%
Distância Borda Caractere com Transformada de Fourier em cada vetor de direção	59,50%	44,17%
Distância Borda Caractere com Zoneamento 3x3	22,69%	36,81%
Distância Borda Caractere com Zoneamento 3x3 com Transformada de Fourier no vetor concatenado	97,54%	70,55%
Distância Borda Caractere com Zoneamento 3x3 com Transformada de Fourier em cada vetor de direção	77,91%	58,28%

VI. APLICAÇÃO

Na seção anterior foram realizados experimentos para verificar qual abordagem de extração de características é a melhor na tarefa de reconhecer frutas. Com esse conhecimento adquirido, e também identificado o classificador que obteve a melhor taxa de acerto, foi desenvolvido um sistema de Realidade Aumentada para realizar o reconhecimento de frutas em imagens capturadas pela câmera de um computador em tempo real.

A Figura 4 apresenta o *pipeline* de execução do sistema de Realidade Aumentada proposto neste trabalho. Primeiramente é realizada a captura da imagem por meio de uma câmera, no caso deste trabalho, foi utilizada a câmera do computador. A próxima etapa consiste na binarização desta imagem, de modo que o fundo seja representado pela cor preta e o objeto de interesse pela cor branca. Neste trabalho, os testes foram realizados em um ambiente controlado, por isso, os fundos onde os objetos foram colocados eram brancos para facilitar a etapa de binarização. Em seguida, são extraídas as características de cada imagem capturada pela câmera; no caso, é utilizada a abordagem que obteve melhor taxa correta de classificação, ou seja, a abordagem ‘Distância Borda Caractere com Zoneamento 3x3 com Transformada de Fourier no vetor concatenado’. Uma vez que as características foram extraídas da imagem, elas servem como entrada para o processo de classificação, no caso, elas são passadas para o classificador Perceptron Multicamadas (MLP) para rotular a imagem capturada pela câmera. Finalmente, as informações sobre a fruta são exibidas na imagem capturada. Nessa etapa o sistema poderia ser adaptado para combinar a imagem real com a virtual. Como o foco deste trabalho foi propor uma abordagem para extração de características de frutas, o sistema de Realidade Aumentada exibe somente o nome da fruta na imagem capturada.

A Figura 5 apresenta três exemplos de reconhecimento de frutas realizados pelo sistema de Realidade Virtual que utiliza a abordagem de extração de características apresentada neste trabalho. No caso, foi realizada a classificação correta de um abacaxi (imagem do topo), de uma banana (imagem central) e de um limão (imagem de baixo).

Como qualquer sistema de visão computacional que seja

executado em um ambiente real, o sistema de Realidade Aumentada proposto neste artigo sofre com qualidade da imagem capturada e de outros fatores que podem atrapalhar o reconhecimento e a identificação da fruta, por exemplo, a segmentação do objeto de interesse ou sua separação do fundo da imagem. Para exemplificar um desses fatores, a Figura 6 apresenta a captura de uma mesma fruta, mas com condições de iluminação diferentes, o que modificou a sombra do objeto de interesse. Na imagem do topo a sombra fez com que as características extraídas não se comportassem de maneira correta, levando à classificação incorreta da imagem, no caso, a laranja foi classificada como um melão. Por outro lado, com outra iluminação (imagem de baixo) e menor influência da sombra, as características extraídas conseguiram discriminá-lo do objeto de interesse e levou o classificador a uma classificação correta, no caso, o objeto foi corretamente identificado como uma laranja.

O sistema captura as imagens e faz o seu reconhecimento em tempo real, entretanto, os experimentos foram realizados em um ambiente controlado, com fundo branco, sem oclusão, sem sombras e sem a rotação das frutas. Portanto, outras técnicas de visão computacional necessitam ser empregadas para que o rastreamento e reconhecimento dos objetos em um cena complexa sejam executados com precisão.

VII. CONCLUSÕES E TRABALHOS FUTUROS

A Realidade Aumentada tem se mostrado uma importante tecnologia para apresentar informações relevantes de modo a complementar o entendimento e exploração de usuários em uma cena, em algum contexto exploratório ou instrutivo. Este trabalho empregou a Realidade Aumentada para apresentar novas informações sobre objetos de interesse em um contexto de reconhecimento de frutas. Para tanto, foi proposto um sistema que não utilizasse marcadores fiduciais para a identificação dos objetos de interesse. Os objetos foram identificados na cena, ainda que em um ambiente controlado, e a partir de sua identificação as informações foram apresentadas para os usuários. Portanto, foram pesquisadas e comparadas abordagens de extração de características de modo a identificar qual seria a mais eficaz no seu reconhecimento. Como resultado



Figura 4: Pipeline do sistema de Realidade Aumentada.

dessas comparações, foram propostas novas abordagens de combinação de extratores já existentes na literatura. Além disso, foi identificada uma combinação de extratores que resultou em uma taxa de classificação correta de 97,5%.

Também foram apresentadas algumas das limitações do sistema, podendo comprometer o reconhecimento das frutas. Em trabalhos futuros, deseja-se explorar maneiras de evitar os problemas com iluminação e sombra dos objetos. Além disso, também é foco de trabalhos futuros o aprimoramento do sistema de Realidade Aumentada para que informações relevantes sejam recuperadas e apresentadas para os usuários,

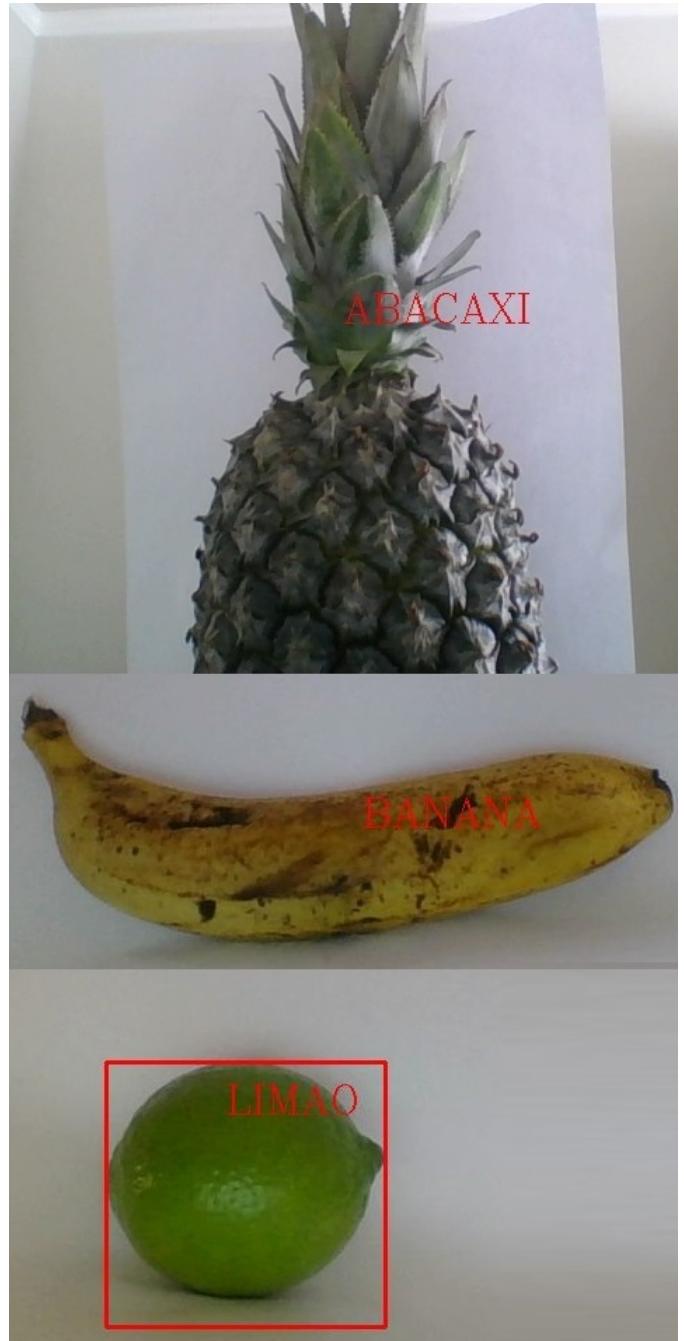


Figura 5: Sistema de Realidade Aumentada para o Reconhecimento de Frutas: exemplos de execução do sistema no reconhecimento de três frutas distintas.



Figura 6: Comparação de reconhecimento de uma mesma fruta variando a iluminação. Na imagem superior a laranja é identificada como melão, em razão da sombra projetada na imagem. Na imagem inferior a laranja é corretamente identificada, não havendo interferência da sombra.

por exemplo, informações nutricionais das frutas ou preços nos supermercados de uma determinada cidade. Nesse segundo exemplo, o usuário poderia capturar imagens de frutas e essas informações seriam apresentadas para ele sob demanda em um dispositivo móvel ou computador. Além de informações textuais, também será incorporado ao sistema a exibição de modelos 3D sobre a cena.

REFERÊNCIAS

- [1] H. Slay, M. Phillips, R. Vernik, and B. Thomas, “Interaction modes for augmented reality visualization,” in *Proceedings of the 2001 Asia-Pacific Symposium on Information Visualisation - Volume 9*, ser. APVis ’01. Darlinghurst, Australia, Australia: Australian Computer Society, Inc., 2001, pp. 71–75. [Online]. Available: <http://dl.acm.org/citation.cfm?id=564040.564049>
- [2] M. P. Neto, D. M. Eler, A. C. de Moarais, and J. R. F. Brega, “An immersive and interactive visualization system by integrating distinct platforms,” in *Proceedings of XIX International Conference on Information Visualization*. Washington, DC, USA: IEEE Computer Society, 2015, pp. 403–410.
- [3] M. C. d. Macedo, A. L. Apolinario Junior, and A. C. d. Souza, “A robust real-time face tracking using head pose estimation for a markerless ar system,” in *Proceedings of the 2013 XV Symposium on Virtual and Augmented Reality*, ser. SVR ’13. Washington, DC, USA: IEEE Computer Society, 2013, pp. 224–227.
- [4] M. C. d. F. Macedo and A. L. Apolinário Júnior, “A markerless augmented reality environment for on-patient medical data visualization,” in *Proceedings Conference on Graphics, Patterns and Images (SIBGRAPI 2015)*. Porto Alegre: Sociedade Brasileira de Computação, 2015.
- [5] W. C. SENG and S. H. MIRSAEE, “A new method for fruits recognition system,” *ICEEI*, 2009.
- [6] S. Arivazhagan, R. N. Shebiah, S. S. Nidhyan, and L. Ganesan, “Fruit recognition using color and texture features,” *Journal of Emerging Trends in Computing and Information Sciences*, vol. 1, no. 2, pp. 90–94, 2010.
- [7] H. Zawbaa, M. Abbass, M. Hazman, and A. Hassenian, “Automatic fruit image recognition system based on shape and color features,” in *Advanced Machine Learning Technologies and Applications*, ser. Communications in Computer and Information Science, A. Hassanieh, M. Tolba, and A. Taher Azar, Eds. Springer International Publishing, 2014, vol. 488, pp. 278–290.
- [8] S. Sengupta and W. S. Lee, “Identification and determination of the number of immature green citrus fruit in a canopy under different ambient light conditions,” *Biosystems Engineering*, 2014, image Analysis in Agriculture.
- [9] C. O. A. Freitas, L. S. Oliveira, S. B. K. Aires, and F. Bortolozzi, “Metaclasses and zoning mechanism applied to handwriting recognition,” *Journal of Universal Computer Science*, vol. 14, no. 2, pp. 211–223, 2008.
- [10] R. A. R. MIRANDA, F. A. D. SILVA, M. A. PAZOTI, A. O. ARTERO, and M. A. PITERI, “Algoritmo para o reconhecimento de caracteres manuscritos,” *Colloquium Exactarum*, vol. 5, pp. 109–127, 2013.
- [11] Visual cognitive systems lab. [Acesso em 22 de fevereiro de 2015]. [Online]. Available: <http://www.vicos.si/Downloads/FIDS30>

Construction of a 3D virtual university with recommendation of learning objects to support distance education

Rodrigo Valença Cavalcante Frade, Francisco Milton Mendes Neto & Rafael Castro de Souza

Universidade Federal Rural do Semi-Árido - UFERSA

Programa de Pós-graduação em Ciência da Computação

Mossoró, Brasil

rodrigo.rvcf@gmail.com, {miltonmendes, rafaelcastro}@ufersa.edu.br

Abstract—In last years the Distance Education (DE) has grown with the support of computational resources that provides new possibilities on teaching and learning process. However, the evasion of students has also been a major problem for this mode of teaching. Some causes of this problem are related to Virtual Learning Environments (VLEs) that usually treat similarly the students. Thus, this paper presents a three-dimensional virtual university that provides the Virtual Reality (VR) through a realistic and immersive scenario. In addition, this environment has resources that can monitor the actions of the students to, based on this, select learning objects (LOs) and offer it spontaneously.

Keywords-Multi-Agent System; Virtual Learning Environments; 3D Paradigm.

I. INTRODUÇÃO

A tecnologia da informação tem se mostrado ao longo dos últimos anos uma opção bastante promissora para a área de educação. Isso pode ser comprovado, por exemplo, através da notória influência que a mesma promove sobre a modalidade do Ensino a Distância (EaD). Atualmente é possível identificar a utilização constante de recursos didáticos ricos em tecnologias por professores e estudantes.

Muitos destes recursos são ferramentas digitais com um nível de interação bastante promissor, que auxiliam as tarefas inclusas nos processos de ensino e aprendizagem, aumentando a qualidade e a produtividade destas atividades. Os Ambientes Virtuais de Aprendizagem (AVAs) são um exemplo destas ferramentas e têm sido destaque por serem bastante utilizados atualmente em soluções para EaD.

Entretanto, muitos desses ambientes ainda necessitam incorporar características que contribuam nas tarefas voltadas à EaD. Ainda que exista uma grande quantidade de AVAs disponíveis no mercado que proporcionam algumas vantagens, muitos deles não conseguem reter a atenção dos jovens estudantes para atividades voltadas à aprendizagem, geralmente ocasionado pela ausência de entretenimento e devido tratarem de forma semelhante os seus usuários [1].

Uma forma de suprir essas necessidades pode ser alcançada pela implantação de técnicas computacionais que proporcionem uma maior atratividade aos estudantes. Em destaque pode ser citado o uso de cenários virtuais em terceira dimensão, que promovem a Realidade Virtual

(RV) em ambientes de aprendizagem. Segundo Amaral et al. [2], esses cenários proporcionam experiências de imersão fortemente realísticas, que envolvem e despertam a sensação em estudantes de estarem dentro do ambiente virtual, permitindo o desenvolvimento de atividades semelhantes às realizadas em âmbito real. Além disso, vale ressaltar que os ambientes 3D podem ainda proporcionar outras formas de visualização de conteúdos educacionais, além de que, com o auxílio de técnicas de inteligência computacional , surge ainda a possibilidade de que o ambiente virtual possa tratar o estudante de forma personalizada, seja por meio de suas preferências de conteúdos, horários de estudo, conexão de amigos, entre outros.

Com base nisto, o presente trabalho propõe uma opção para auxiliar nos processos de ensino e aprendizagem de cursos em EaD por meio da construção de um sistema que permite simular uma universidade virtual 3D em navegadores de Internet ou nativamente em dispositivos móveis com sistemas Android, e que ainda dispõe de um sistema multiagente, responsável pela oferta de Objetos de Aprendizagem (OAs).

Este artigo encontra-se estruturado em sete seções. A Seção II apresenta trabalhos existentes com propósitos semelhantes. A Seção III discorre sobre tecnologias e conceitos abordados neste trabalho. A Seção IV descreve a arquitetura do ambiente virtual 3D apresentado. A Seção V expõe especificações sobre o desenvolvimento do ambiente 3D e dos demais componentes. A Seção VI apresenta um cenário de uso. Por fim, a Seção VII contempla as considerações finais e trabalhos futuros.

II. TRABALHOS RELACIONADOS

Voss et. al. [3] apresentaram o TCN5 (TCN FIVE), acrônimo para *Teaching Computer Networks in a Free Immersive Virtual Environment*, que é um ambiente virtual tridimensional imersivo no apoio ao ensino da disciplina de Redes de Computadores. O ambiente virtual foi construído utilizando o OpenSim, que é um servidor gratuito de mundos virtuais, e o mesmo está integrado ao Moodle de modo que algumas extensões e módulos desse possam ser utilizados no ambiente virtual. Essa integração se dá pela utilização do Sloode (*Simulation-Linked Object Oriented Dynamic Learning Environment*) que permite integrar recursos do Moodle com ambientes virtuais.

Seok e Kim [4] propuseram um ambiente virtual Web 3D aplicado no ensino do aquecimento global. O ambiente foi implementado utilizando o O3D-WebGL. O WebGL é um novo padrão de desenho 3D na Web. Este permite a renderização 3D para Canvas em HTML5. Já o O3D é uma API Web *open-source* para a implementação de gráfico 3D em navegadores da Web.

O fator que difere este trabalho dos demais é que o ambiente virtual foi construído utilizando o Unity 3D, que, além de permitir a execução em sistemas Web, ainda possibilita a exportação de projetos para serem executados como aplicativos de forma nativa em dispositivos móveis, como o iOS e o Android.

Vale ainda ressaltar que o presente trabalho dispõe de um sistema multiagente com a finalidade de recomendar, de forma personalizada, OAs para o estudante de acordo com suas preferências, tais como: cursos, tipos de arquivo, localização virtual no ambiente e geográfica via GPS, etc. De modo que a recomendação seja capaz de se adaptar ao contexto do estudante, sem a necessidade de intervenções humanas, fator este que os trabalhos supracitados não contemplam.

III. FUNDAMENTAÇÃO TEÓRICA

A. Ambientes Virtuais de Aprendizagem

O progresso constante da EaD tem demarcado um novo horizonte no âmbito educacional. A desconfiança que existia no passado quanto à qualidade nos processos de ensino e aprendizagem dessa modalidade tem se extinguido com o auxílio de novas técnicas que envolvem muitos processos e ferramentas. Como destaque a esta contribuição, é possível citar os AVAs.

Na opinião de Litto e Formiga [6], os AVAs permitem novas e potencialmente diferentes experiências de aprendizagem que não devem ser desprezadas pelo professor na busca por estratégias para que os estudantes atinjam os seus objetivos de aprendizagem.

Os pesquisadores Li, Yue e Jáuregui [8] alegam que o mais importante de uma ferramenta para EaD é a forma como a mesma irá atrair a atenção do seu público, em especial, para o processo de aprendizagem. Então, é de fundamental importância levar em consideração os 3Is ilustrados na Figura 1.

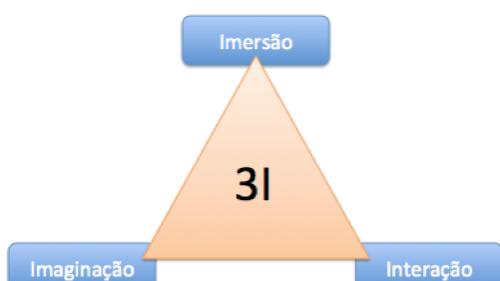


Figura 1: Os 3Is - Imersão, Imaginação e Interação.

Fonte: Adaptado de Li, Yue e Jáuregui [8].

Incluem-se, como exemplos, utilizando essas características providas pelos 3Is: i) Entretenimento: jogos, viagens virtuais, etc; ii) Saúde: cirurgias virtuais, tratamento de pacientes em UTI, reabilitação, etc; iii) Negócios: maquetes virtuais, edificações, projeto de interiores, etc; iv) Educação e Treinamento: simuladores de voo, jogos educativos, etc.

B. Aplicando o paradigma 3D em AVAs

Para prover a RV atualmente é comum o uso de interfaces 3D nos AVAs. Osório et al. [9] alegam que o paradigma 3D permite a representação da informação de um modo realístico, organizando a mesma de uma forma espacial e tornando sua visualização mais intuitiva por ser mais natural aos olhos humanos.

Baseado nestas afirmações, Bos et al. [10] concluem que ambientes virtuais 3D contribuem significativamente nos processos de ensino e aprendizagem mediados por tecnologias, uma vez que estes podem trazer uma gama de diferentes recursos.

Reforçando, Litto e Formiga [6] asseguram a intensa interação que um ambiente virtual 3D pode proporcionar, alegando que a possibilidade de criar locais de aprendizagem mais lúdicos e ricos, em várias dimensões, provoca nos estudantes uma interação mais intensa e prazerosa com os colegas, o professor, o conteúdo e, principalmente, os objetos e o próprio ambiente, em seu percurso de aprendizagem.

C. Sistemas Multiagente

De acordo com Luger [12], um Sistema Multiagente (SMA) consiste em um programa de computador com “resolvedores” de problemas, denominados de agentes inteligentes, localizados em ambientes interativos, onde os mesmos são capazes de executar ações flexíveis, autônomas e, ainda, socialmente organizadas que podem ser dirigidas para metas ou objetivos predeterminados.

Em outras palavras, ainda é possível afirmar que um SMA consiste em um conjunto de agentes colaborativos a fim de resolver um problema em comum, e que pode resolver um problema maior dividindo o mesmo em vários subproblemas, a fim de otimizar o tempo de solução.

Luger [12] também apresenta quatro critérios para um SMA. Ele alega que o mesmo deve incluir agentes que são: i) Situados: recebe entradas do ambiente no qual está ativo e pode, também, efetuar modificações dentro desse ambiente; ii) Autônomos: Interagi com o seu ambiente sem a intervenção direta de outros agentes; iii) Flexíveis: Intelligentemente *responsivo* e *proativo* conforme a situação atual; iv) Sociais: Interagi, de modo apropriado, com outro software ou com agentes humanos.

A utilização de um SMA pode auxiliar em tarefas mais complexas dentro de um ambiente de aprendizagem, como, por exemplo, monitoramento de atividades de estudantes, captura de informações sobre o seu contexto dinâmico, recomendação de conteúdo baseada em seu perfil de usuário, dentre outras.

IV. ARQUITETURA DA UNIVERSIDADE VIRTUAL 3D DE APOIO A EAD

Esta seção tem como finalidade descrever a universidade virtual em contexto técnico por meio da sua arquitetura apresentada na Subseção a seguir.

A. Descrição do Contexto Técnico

O desenvolvimento da universidade virtual teve como objetivo a reprodução de um ambiente imersivo proporcionado através do uso de algumas tendências tecnológicas aqui abordadas. Deste modo, a grande contribuição considerada neste trabalho foi a tarefa de integrar essas tecnologias, avaliada como um desafio devido à complexidade de efetuar a comunicação entre todos os componentes que contemplam a arquitetura do projeto (Figura 2).

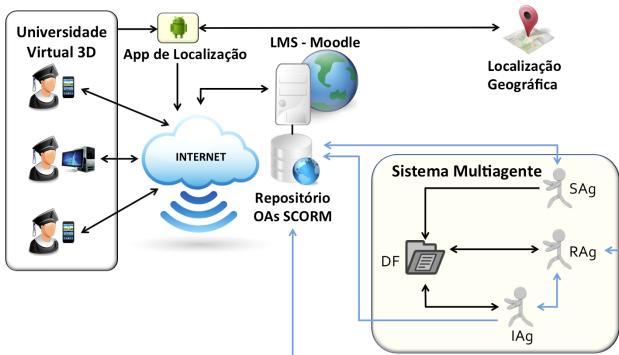


Figura 2: Arquitetura do projeto. Fonte: Autoria própria.

Os principais componentes que compõem essa arquitetura são:

- 1) SMA: Sistema multiagente criado para proporcionar uma maior autonomia ao sistema e prover a recomendação de forma automática e personalizada de Objetos de Aprendizagem (OAs), que são entidades digitais com fins educacionais [13].
- 2) AVA - Moodle: O Moodle foi imprescindível para o funcionamento da ferramenta, pois ambos se comunicam. A autenticação do estudante na universidade virtual, por exemplo, é feita a partir do cadastro prévio em um AVA. Neste caso foi utilizado o Moodle, assim como o repositório de OAs que se encontra na base de dados desse AVA. Sua relevância neste projeto também está atrelada a oferta dos cursos a distância, deste modo, sem esse componente seria inviável.
- 3) Repositório de OAs SCORM: Este é o repositório utilizado para armazenamento dos OAs. Ele fica localizado na base de dados do Moodle. Seus OAs devem ter sido desenvolvidos seguindo o padrão SCORM [14]. Através das informações fornecidas no SCORM e dos metadados do padrão LOM, que é parte integrante do SCORM, os agentes serão capazes de comparar as informações dos OAs com os perfis dos estudantes.
- 4) Universidade Virtual: Consiste em um ambiente 3D interativo com NPCs, que representam os agentes

de software. É nele que ocorre a oferta de OAs. Este sistema deve estar inteiramente ligado ao banco de dados do AVA utilizado para prover as suas funcionalidades.

- 5) App de Localização via GPS: Este software tem a responsabilidade de cadastrar a localização de ambientes reais, como laboratórios ou departamentos, onde são especificadas as áreas de estudo. No ato da autenticação, é possível capturar a localização do estudante, caso ele esteja em um ambiente já cadastrado. Essa informação é guardada para auxiliar na recomendação personalizada.

V. DESENVOLVIMENTO DA UNIVERSIDADE VIRTUAL 3D E SEUS COMPONENTES

Esta seção contempla a descrição das implementações de cada um dos componentes que compõem a arquitetura do sistema, bem como cita as ferramentas utilizadas para tal tarefa.

A. Desenvolvimento do Ambiente Virtual 3D

Para desenvolver o ambiente virtual 3D, foi de suma importância a escolha da ferramenta de implementação. Contudo, não foi uma tarefa simples, pois era necessário primeiramente analisar e testar as ferramentas existentes para então tomar a decisão baseada na necessidade do projeto.

Através de pesquisas, foi possível identificar uma ferramenta para o desenvolvimento de jogos 2D e 3D que apresenta uma popularidade muito grande atualmente, o Unity 3D [16]. Segundo Creighton [17], o Unity consiste em um motor de jogo (*game engine*) ou em uma ferramenta para criação de jogos que permite pessoas criativas desenvolverem projetos ricos, e a curto prazo, devido o montante de recursos que a mesma dispõe.

A Figura 3 exibe o interior de um dos laboratórios.

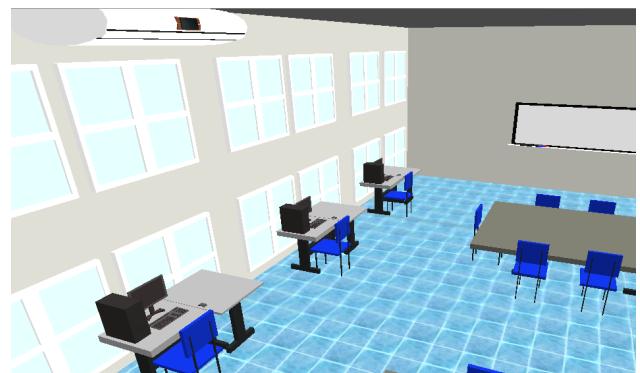


Figura 3: Interior de um dos laboratórios. Fonte: Autoria própria.

B. Criação dos Personagens 3D

Para o desenvolvimento dessa etapa foram utilizados o Fuse [18], o sistema online da Mixamo [21] e o Unity.

O Fuse é um software proprietário para criação de personagens 3D orientado a dados que vem com partes do corpo, roupas, malhas e texturas dinâmicas, alimentados

pela tecnologia *Substance Designer* (criação de texturas em alto nível) da empresa Allegorithmic [20]. A Mixamo é a empresa proprietária do Fuse que dispõe de um recurso online de *autorigger*, que consiste na aplicação automática de animações a personagens 3D carregados neste sistema web [21].

Primeiramente foi criado o avatar que consiste em um estudante. Em seguida, foram criados os NPCs que representam os professores/tutores em cada laboratório. Esta tarefa exige três passos básicos, que são: i) Criação do personagem em 3D (Fuse); ii) Aplicação da animação ao personagem (Sistema *Auto-Rigger* da Mixamo); iii) Aplicação dos *scripts* para utilizar as animações (Unity).

Esta ferramenta permite uma montagem do objeto 3D por partes do corpo, através de um banco que disponibiliza essas partes pré-prontas. Primeiramente o usuário escolhe a cabeça, em seguida o corpo, depois os braços, pernas e roupas. No final é permitido efetuar algumas customizações em texturas, nas expressões faciais etc. A Figura 4 exibe a etapa referente à confecção do NPC.



Figura 4: Etapa - Confecção do NPC. Fonte: Autoria própria.

Com a concretização do personagem 3D, se faz necessária a aplicação de animações para proporcionar Realidade Virtual. Neste processo entra a utilização do sistema on-line da empresa Mixamo.

Vale ressaltar que na modelagem de um personagem 3D é imprescindível a criação do seu esqueleto (Figura 5) para auxiliar nos movimentos de todas as partes do corpo. Após isso, aplicam-se as animações.

Agora o personagem está apto a ser equipado com animações. O site da Mixamo contém um número significativo de animações gratuitas para serem aplicadas. Comumente utilizam-se as seguintes expressões relacionadas às animações básicas: *idle*, *walking*, *right turn* e *left turn*; respectivamente: parado, andando, vira à direita e vira à esquerda. Todos os personagens da universidade virtual receberam essas animações básicas. Por fim, foram exportados em formato FBX [22] reconhecido pelo Unity.

Uma vez concluída a criação das partes fundamentais dos personagens, chega-se à necessidade da implementação da lógica de funcionamento destes avatares agora dentro do cenário em 3D. Então foram importados

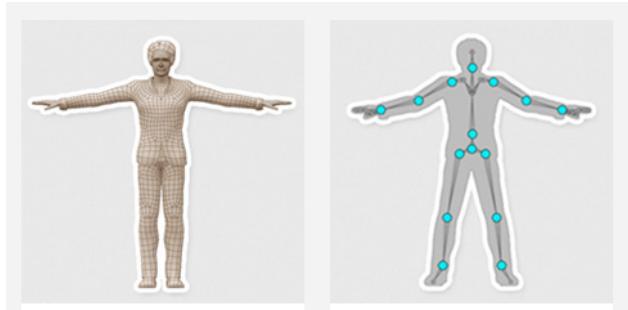


Figura 5: O antes e o depois da aplicação do esqueleto no Mixamo. Fonte: Mixamo [21].

todos para o Unity e, em seguida, desenvolvidos *scripts* específicos para prover algumas funcionalidades.

A ideia foi empregar comandos capazes de proporcionar ações conforme o ambiente. Como principais ações, podem ser citadas: mover o personagem estudante de acordo com determinados botões do teclado, mover aleatoriamente os NPCs, fazer com que os NPCs parem ao detectar o personagem estudante dentro de um determinado raio de distância e, a partir de uma aproximação, siga o avatar do estudante até que ele ultrapasse essa aproximação.

Todos os scripts mencionados neste trabalho foram omitidos devido o foco deste artigo ser o ambiente virtual 3D, entretanto, são apresentadas algumas figuras que consistem na semântica das implementações, facilitando o entendimento do leitor sobre as funcionalidades desses scripts.

C. Desenvolvimento do Projeto de Captura da Localização Real

Para o desenvolvimento deste projeto foi utilizada a plataforma Android SDK que apresenta características importantes, como, por exemplo, ser um software livre, aceitar uso de várias APIs, uso da linguagem de programação Java, etc [23], [24].

O processo de identificação da localização real do estudante é efetuada com o auxílio de classes Java, responsáveis por fornecerem acesso aos serviços de localização geográfica, e que podem representar essa localização através das seguintes informações: latitude, longitude, data e hora, entre outras. A Figura 6, representa esse processo de captura da localização real do estudante, a partir da versão para Android da Universidade 3D.

É importante lembrar que o sistema necessita do cadastramento dos laboratórios ou departamentos para poder ser feita a comparação sobre a localização real do estudante e a virtual. A localização virtual do estudante é baseada em que prédio, da universidade 3D ele se encontra.

D. Desenvolvimento do Sistema Multiagente

No desenvolvimento dos agentes foi utilizado o JADE (*Java Agent Development Framework*), uma plataforma específica para o desenvolvimento e execução de sistemas multiagente [25].

Os agentes criados monitoram os estudantes, selecionam os OAs mais adequados e ofertam estes OAs aos

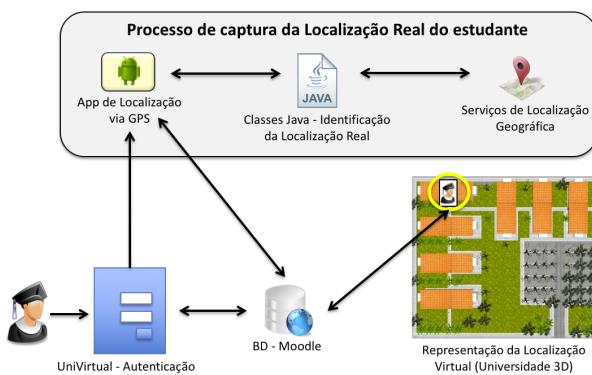


Figura 6: Processo de captura da Localização Real.
Fonte: Autoria própria.

estudantes de acordo com seus contextos. Para isto foi necessária a implementação de três tipos de agentes que comporão o SMA. A organização deste SMA é do tipo comunidade de especialistas, com base em Artero [15], visto que cada um dos tipos de agentes criados possuem o mesmo nível. Entretanto, cada um deles é especialista em determinada tarefa. Foram utilizadas as bibliotecas do JADE [25] para a criação dos seguintes agentes: i) Agente Estudante (SAg – *Student Agent*); ii) Agente de Interface (IAg - *Interface Agent*); iii) Agente Recomendador (RAg – *Recommender Agent*).

A responsabilidade dos SAGs consiste no monitoramento das atividades dos estudantes e recuperação das suas preferências de aprendizagem, obtidas através dos seus perfis de usuários. O seu comportamento é realizado a partir da autenticação do estudante na universidade virtual. Estes agentes também são capazes de capturar as informações do contexto dinâmico do estudante, além de sua localização geográfica. Por fim, todas essas informações são enviadas para o RAg.

Por sua vez, o RAg, de acordo com as informações fornecidas pelos SAGs e as obtidas dos OAs disponíveis no repositório, detecta os OAs que correspondem ao contexto atual do estudante. Inicialmente ele encontra estes OAs mais adequados e, depois de selecioná-los, os repassa para o IAg que, por sua vez, irá executar a oferta dos OAs ao estudante.

O papel principal do IAg é de receber a lista de OAs selecionados pelo RAg para, em seguida, disponibilizá-los aos estudantes. Após esse procedimento, os NPCs poderão consultar estes OAs para fazer o repasse ao estudante logado no ambiente tridimensional.

VI. CENÁRIO DE USO

Na Figura 7 estão ilustrados dois cenários referentes a um teste de usabilidade. Primeiramente a tela de autenticação [1], da qual o estudante fornece o mesmo usuário do Moodle e a senha que ele cadastrou para a universidade virtual. Em seguida, o cenário inicial do ambiente [2] já carregado logo após a autenticação do usuário.

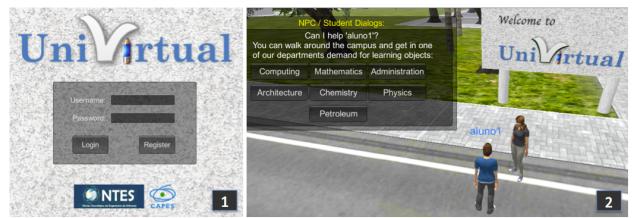


Figura 7: Tela de autenticação [1] e cenário inicial do ambiente 3D [2]. Fonte: Autoria própria.

Ao carregar a cena inicial, o estudante obtém a posse do controle de locomoção do avatar, que o representará neste universo digital. Este controle atende aos comandos tradicionais do teclado: W = frente, A = esquerda, D = direita, Q = girar à esquerda, e E = girar à direita, no caso do uso em computadores. Caso o usuário esteja em um dispositivo móvel, irá aparecer o joystick virtual que atende ao toque conforme a Figura 8.



Figura 8: Comandos da UniVirtual no Android. Fonte: Autoria própria.

Todo o ambiente virtual contém ao menos um NPC em cada cena. Nesta cena inicial (Figura 8), por exemplo, o estudante pode se deparar com um personagem virtual que orienta-o quanto aos prédios e seus respectivos cursos. Já os demais NPCs encontrados no interior de cada prédio são responsáveis pela recomendação dos OAs.

Na Figura 9, pode-se ver o personagem que representa o estudante autenticado prestes a entrar no departamento de computação e, em seguida, o momento em que ocorre a recomendação de OAs.

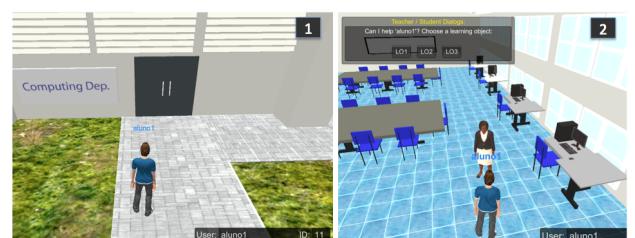


Figura 9: Cenário inicial [1] e sugestão de OAs [2]. Fonte: Autoria própria.

No momento em que ocorre a identificação da proximidade do estudante, o NPC aciona o SMA que já se

encontra proativamente em execução, por sua vez o SMA entra em ação e faz a seleção do OA mais indicado para o perfil do estudante autenticado.

VII. CONCLUSÕES E TRABALHOS FUTUROS

Neste artigo foi apresentado todo o processo de desenvolvimento de um ambiente virtual tridimensional com recomendação de OAs. O sistema abordado objetiva tornar a aprendizagem mais atrativa com o emprego de imersão, proporcionando assim, a Realidade Virtual. Esse ambiente pode ser aplicado em qualquer ferramenta de gestão de aprendizagem, devendo-se apenas fazer algumas adaptações.

Para trabalhos futuros, pretende-se aprimorar o ambiente 3D inserindo mais inteligência em seus objetos de cenas, incluindo a implementação de técnicas que permitam a abertura dos OAs de um modo bastante interativo, na expectativa de aumentar a imersão já contida na universidade virtual.

Também almeja-se efetuar testes em um contexto real, onde os próprios estudantes avaliariam a ferramenta auxiliando na identificação de possíveis falhas, visando o seu aprimoramento.

REFERÊNCIAS

- [1] E. A. Oliveira, P. i. Tedesco, *collaboration*: “Um modelo de Colaboração Inteligente Personalizada para Ambientes de EAD”. Revista Brasileira de Informática na Educação, 18(1), p.17-31. 2010.
- [2] É. Amaral, B. G. Avila, L. M. Tarouco, “Aspectos Teóricos e Práticos da Implantação de um Laboratório Virtual no OpenSim”. Anais do Simpósio Brasileiro de Informática na Educação, Vol. 23, No. 1. 2012.
- [3] G. B. Voss, V. Oliveira, F. B. Nunes, F. Herpich, R. D. Medina, M. Bercht, “Construção e Análise de um Mundo Virtual 3D para o Ensino e Aprendizagem de Redes de Computadores”. Anais do Simpósio Brasileiro de Informática na Educação. Vol. 25. No. 1. 2014.
- [4] K. Seok,Y. S. Kim, “*Implementation of Interactive Web 3D Based Virtual Educational System using O3D-WebGL*”. *International Journal of Multimedia and Ubiquitous Engineering*. Vol.9, No.4 2014.
- [5] F. A. P. Fialho, “Gestão do conhecimento e aprendizagem: as estratégias competitivas da sociedade pós-industrial”. Visual Books. 2006.
- [6] F. M. Litto, M. Formiga, “Educação a distância: o estado da arte”. Pearson. 2012.
- [7] R. Marcelino, “Ambiente virtual de aprendizagem integrado a mundo virtual 3D e a experimento remoto aplicados ao tema resistência dos materiais”. Tese de Doutorado em Engenharia de Minas, Metalúrgica e de Materiais – Universidade Federal do Rio Grande do Sul. 2010.
- [8] Z. Li, J. Yue, D. A. G. Jáuregui, “A new virtual reality environment used for e-learning”. In *IT in Medicine & Education. ITIME'09. IEEE International Symposium*. Vol. 1, pp. 445-449. 2009.
- [9] F. S. Osório, S. R. Musse, C.T.D. Santos, F. Heinen, A. Braun, A. T. S Silva, “Ambientes Virtuais Interativos e Inteligentes: Fundamentos, Implementação e Aplicações Práticas”. XXIV Congresso da SBC-JAI. 2004.
- [10] A. S. Bos, P. Dalssasso, L. H. C. Rosa, G. Bernardi, F. M. Müller, A. Z. Cordenonsi, “A História da Computação através de um Museu Virtual Interativo utilizando Mundos Virtuais 3D”. In Anais dos Workshops do Congresso Brasileiro de Informática na Educação. Vol. 2, No. 1. 2013.
- [11] Moodle. (2015, 2 February) Site oficial do Moodle [Online]. Available: <http://moodle.org/>.
- [12] G. F. Luger, “Inteligência Artificial: Estruturas e estratégias para a solução de problemas complexos”. 6º Edição. [S.I]: Bookman. 2014.
- [13] Objetos de Aprendizagem (2015, 5 February) [Online]. Available: <http://www.nuted.ufrgs.br/arquead/objetos.html>.
- [14] ADL. “*ADL Guidelines for Creating Reusable Content with SCORM 2004*”. Advanced Distributed Learning. [S.I]: 4ed. 2008.
- [15] A. O. Artero, “Inteligência Artificial - Teoria e Prática”. Livraria da Física. 2009.
- [16] Unity (2015, 5 February) Site oficial do Unity [Online]. Available: <http://unity3d.com/unity>.
- [17] R. H. Creighton, “*Unity 3D Game Development by Example: A Seat-of-Your-Pants Manual for Building Fun, Groovy Little Games Quickly*”. Packt Publishing Ltd. 2010.
- [18] Fuse (2014, 10 December) Fuse character creator high-quality 3D characters in minutes, not days [Online]. Available: <http://www.mixamo.com/fuse>.
- [19] CSharp (2015, 7 January) Introdução à linguagem C# e ao .NET Framework [Online]. Available: <https://msdn.microsoft.com/pt-br/library/z1zx9t92.aspx>.
- [20] Allegorithmic (2014, 10 December) *Substance: simply the most powerful texturing toolset out there* [Online]. Available: <http://www.allegorithmic.com/fuse>.
- [21] Mixamo (2014, 10 December) Site oficial da mixamo [Online]. Available: <http://www.mixamo.com/>.
- [22] FBX (2015, 15 January) Interoperabilidade: Autodesk FBX [Online]. Available: <http://www.autodesk.com.br/adsk/servlet/index?siteID=103425&id=11150639>.
- [23] F. T. Franciscato, R. D. M Medina, “Learning e Android: um novo paradigma?”. RENOTE. 2008.
- [24] Android (2014, 10 December) Site oficial do sistema android [Online]. Available: <http://www.android.com/>.
- [25] Jade (2014, 10 December) *Java development framework: an open-source platform for peer-to-peer agent based applications* [Online]. Available: <http://jade.tilab.com/>.
- [26] J. Carvalho, “Ensino-aprendizagem do PLE em modalidade híbrida”. Revista Animação e Educação. 2010.

Um Estudo sobre a Aceitação e Uso da Realidade Aumentada em Cenários de Negócio sob a ótica da Teoria Unificada de Aceitação e Uso da Tecnologia

A Study about the Acceptance and Use of Augmented Reality applied to Business Scenarios according to Unified Theory of Acceptance and Use of Technology

Joao Vitor de Matos Silva

Faculdade de Economia, Administração e Contabilidade
de Ribeirão Preto (FEA-RP)
Universidade de São Paulo (USP)
Ribeirão Preto, SP, Brasil
jvmsilva@fearn.usp.br

Ildeberto A. Rodello

Faculdade de Economia, Administração e Contabilidade
de Ribeirão Preto (FEA-RP)
Universidade de São Paulo (USP)
Ribeirão Preto, SP, Brasil
rodello@fearn.usp.br

Abstract—The augmented reality usage applied to business scenarios still requires studies to confirm its benefits and difficulties. This paper presents a study based on the Unified Theory of Acceptance and Use of Technology (UTAUT) to identify the main topics related to augmented reality acceptance and use, according to the Theory. The research involved 88 participants. Using the factorial analysis, the answers pointed to behavioral intention and effort expectancy as the fundamental items for augmented reality acceptance and use. Finally, this study indicates strategies to the companies developing applications based on augmented reality.

Keywords—UTAUT; augmented reality; user acceptance

I. INTRODUÇÃO

A ideia fundamental de um sistema de Realidade Aumentada (RA) é aumentar a percepção e a interação dos usuários com o mundo real, completando o mundo real com objetos virtuais tridimensionais (3D) que parecem coexistir no mesmo espaço como o mundo real [1] [6]. De maneira geral, a RA oferece uma interface atraente e intuitiva entre usuário e computador.

Segundo Van Kleef et al. [2], “Apesar da grande quantidade de possibilidades com aplicações de Realidade Aumentada, apenas alguns modelos de negócios exploram seu potencial. Isso pode ser causado por não ser claro para as empresas como elas podem agregar valor com a RA”

De maneira complementar, Rodello e Brega [3] destacam que, “utilizando a RA a empresa é capaz de utilizar os quatro

elementos do paradigma de novas comunicações: entreter, informar, persuadir e interagir”. Algumas empresas como Ray Ban, Lacta, Stella Artois, Volvo, dentre outras, já utilizam a RA como ferramenta de Marketing, fornecendo experiências únicas aos seus consumidores. Porém, ainda segundo Rodello e Brega [3], seu desenvolvimento ainda apresenta custo elevado, e o seu funcionamento é melhor observado com produtos que requerem uma grande interação do usuário.

Com a intenção de indicar estratégias para utilização mais assertiva da RA em cenários de negócio, nessa pesquisa realizou-se um estudo sobre a aceitação do usuário (consumidor) em situações de uso da RA no marketing de produtos. O modelo utilizado foi o proposto por Venkatesh et al. [4], denominada Teoria Unificada de Aceitação e Uso da Tecnologia – UTAUT (*Unified Theory of Acceptance and Use of Technology*), que atualmente é considerado o modelo mais completo para essa finalidade, uma vez que abrange outros oito estudos sobre aceitação de tecnologia e os unifica em um único modelo.

Com a análise destacaram-se fatores que, de acordo com o UTAUT, influenciam os usuários e são determinantes na sua aceitação, bem como o seu uso, auxiliando gestores e empresas na elaboração e desenvolvimento de estratégias e de ferramentas que utilizam a RA, de forma a agregar valor para os seus negócios. Além disso, no âmbito das aplicações de RA, testes e análises utilizando o modelo UTAUT são pouco difundidos.

Para melhor exposição do tema, este artigo apresenta na seção II, o embasamento teórico, envolvendo os seguintes

temas: Realidade Aumentada aplicada a cenários de negócio e detalhes do modelo UTAUT. Os procedimentos metodológicos são apresentados na seção III, enquanto a seção IV apresenta os resultados e análise que suportam as conclusões (seção V).

II. EMBASAMENTO TEÓRICO

A. Realidade Aumentada aplicada a cenários de Negócios

A área de Negócios envolve, principalmente, empresas e suas ações para alcançar determinadas metas estratégicas e competitivas. Por uma empresa entender-se que seja “um sistema administrado projetado e operado para atingir determinado conjunto de objetivos” [5].

Uma Empresa com fins lucrativos tem os objetivos de [5]: i) realizar lucro para seus donos; ii) fornecer bens e serviços aos clientes; iii) prover renda para seus funcionários; e iv) aumentar o nível de satisfação de cada um dos seus envolvidos.

Segundo Van Kleef et al. [2], apesar de uma grande quantidade de aplicações para a RA, poucos modelos de negócios exploram essa tecnologia, isso pode ser causado por as empresas não saberem como utilizar a RA, o que está relacionado diretamente com a falta de estudos na área.

A utilização da RA poderia trazer diversos benefícios para as organizações, como por exemplo: alavancar vendas de produtos por meio de uma interação do cliente durante o processo de compra do produto, com o uso de propagandas para *smartphones*, computadores e televisões utilizando a RA, modelagem de roupas para os clientes a distância. Tudo isso poderia funcionar como uma forma de inovação das suas estratégias de marketing, de relacionamento com o cliente, entre outros. Porém, as empresas atualmente pouco exploram essas possibilidades.

B. Adoção e Aceitação da Tecnologia de Informação

Existem vários fatores que influenciam as decisões individuais no contexto da adoção e difusão da Tecnologia da Informação (TI). Depois de analisar a literatura sobre aceitação do usuário e discutir e comparar oito modelos de destaque e suas extensões, Venkatesh et al. [4] formularam e validaram empiricamente uma teoria unificada que integra elementos desses oito modelos. O UTAUT foi modelo desenvolvido para ser um ponto de partida para investigar a adoção de TI e pode ser usado para identificar os fatores que influenciam a intenção de usá-la, bem como para ser adotada por uma organização.

A teoria sobre UTAUT se sustenta em quatro constructos principais: expectativa de desempenho, expectativa de esforço, influência social e condições facilitadoras. Os três primeiros são determinantes diretos da intenção de uso e comportamento e o quarto é determinante direto do comportamento de uso (Fig. 1). A intenção comportamental refere-se ao grau em que uma pessoa (usuário) formulou planos conscientes para realizar ou não algum comportamento futuro uso especificado.

A expectativa de desempenho é definida como o grau em que um indivíduo acredita que a utilização do sistema irá ajudá-lo a conseguir ganhos de desempenho no trabalho. É composto por cinco constructos: utilidade percebida,

motivação extrínseca, ajuste a tarefa, vantagem relativa e expectativas de resultados.

Expectativa de esforço, por sua vez, é definida como o grau de facilidade associado com o uso do sistema. Compreende três constructos: facilidade de utilização percebida, complexidade e a facilidade de utilização. É possível considerar que expectativa de esforço seja definida como o grau de facilidade associada com a utilização do sistema, sendo referente à quantidade de esforço percebido que o usuário precisa colocar para aprender e operar o mesmo.

A influência social refere-se ao grau em que um indivíduo percebe que outras pessoas importantes acreditam que ele deveria usar o sistema. Os principais constructos de influência social são: norma subjetiva, fatores sociais e imagem.

Finalmente, condições facilitadoras é o grau em que um indivíduo acredita que existe uma infraestrutura organizacional e técnica suficiente para suportar a utilização do sistema. Esta definição é apoiada por três constructos diferentes: controle comportamental percebido, condições e compatibilidade.

O modelo UTAUT ainda prevê ainda quatro constructos moderadores da intenção e uso da TI (Fig. 1): o gênero, a idade, a experiência do indivíduo e a voluntariedade. A experiência refere-se ao grau de familiaridade do usuário com o sistema em questão, familiaridade esta adquirida por meio do uso prévio da tecnologia; e a voluntariedade contempla o grau em que o usuário acredita ser obrigatório ou não o uso do sistema em seu trabalho. Gênero, idade, experiência e voluntariedade de uso são fatores que podem afetar os determinantes diretos da intenção comportamental e do comportamento de uso (expectativa de desempenho, expectativa de esforço, influência social e condições facilitadoras).

A validação do UTAUT é apresentada e discutida em [4]. Do ponto de vista teórico, o UTAUT fornece uma visão refinada de como os determinantes da evolução da intenção e do comportamento ao longo do tempo.

III. PROCEDIMENTOS METODOLOGICOS

A pesquisa levantou informações sobre a aceitação dos possíveis usuários (clientes) frente ao uso da RA por meio da ferramenta *Rayban Virtual Mirror* (<http://www.rayban.com/international/virtual-mirror>). O método de obtenção dos dados da pesquisa foi pela aplicação de questionários a um grupo de pessoas participantes, selecionadas por conveniência, a fim de verificar a aceitação da tecnologia, após o uso da ferramenta citada anteriormente, ou seja, inicialmente as pessoas foram convidadas a utilizar o *Rayban Virtual Mirror* e, em seguida, respondiam um questionário.

Os questionários foram aplicados de dois modos. O primeiro virtualmente, por meio da distribuição de *links* via *email* ou mídias sociais para pessoas específicas. Neste modo, os respondentes utilizavam o seu próprio computador para explorar o software e responder o questionário. O segundo método foi realizado presencialmente, em uma sala equipada com computador e *webcam* disponibilizada para a utilização do software pelos voluntários e aplicação do questionário nos

mesmos. O questionário aplicado foi baseado no modelo UTAUT, proposto por Venkatesh et al. [4] (Quadro 1) e adaptado para aplicação com indivíduos que tiveram contato com alguma ferramenta RA que, para o nosso estudo, foi o *Rayban Virtual Mirror*.

Esta ferramenta foi escolhida para o estudo pois é facilmente acessada via internet, é parte de ação de marketing de um empresa cuja marca tem presença mundial e é considerada uma boa experiência do usuário frente a RA. Como mencionado, o questionário foi estruturado com base no modelo formulado por Venkatesh et al. [4] que é composto por 19 itens, apresentados no Quadro 1. Foi utilizada uma escala de avaliação Likert de 5 pontos, onde 1 representava “Discordo totalmente” e 5 “Concordo totalmente”.

Antes da sua aplicação, o questionário foi testado em uma amostra de seis voluntários com o objetivo de identificar possíveis dúvidas de compreensão e problemas de formatação das questões do questionário. Os voluntários não indicaram nenhum problema no questionário.

IV. APRESENTAÇÃO DOS RESULTADOS

A. Caracterização do Aplicativo

O *Rayban Virtual Mirror* é uma aplicação web que permite ao usuário experimentar virtualmente os produtos oferecidos pela empresa em um espelho virtual, por meio do vídeo captado pela *webcam* do dispositivo do cliente. A ferramenta contém uma das caracterizações feitas por [1] [5] de que “A RA permite que o usuário visualize os objetos virtuais sobrepostos ou compõe uma cena com o mundo real. Dessa forma, objetos reais e virtuais coexistirão num mesmo espaço”. Portanto, o aplicativo permite ao usuário ver a imagem virtual dos óculos, sobreposta a sua imagem ou vídeo. A aplicação também conta com a opção de modelos virtuais oferecidos no site.

A aplicação apresenta várias opções de modelos dos óculos fabricados pela empresa, e permite também que o usuário teste simultaneamente até quatro modelos diferentes, podendo assim fazer comparações entre eles. Na ferramenta também é possível tirar fotos do usuário experimentando os modelos e também permite o compartilhamento da imagem em mídias sociais, assim como o próprio armazenamento desta no computador.

B. Caracterização da Amostra

No total, foram coletadas 95 respostas. Porém, sete não puderam ser processadas devido a estarem incompletas. Sendo assim, a amostra desta pesquisa é composta por 88 participantes, cuja média de idade é de 22 anos. Quanto ao gênero, 55% são homens e 45% são mulheres.

A maioria dos respondentes são graduandos do curso de bacharelado em Administração de Empresas, 54,5%, seguido do curso de Economia, com 13,6%, do curso de Economia Empresarial e Controladoria, 7,9%, do curso de Nutrição e Metabolismo, 5,7%, e de 4,5% do curso de Contabilidade, ainda, 13,6% não informaram a situação de sua ocupação. Todos pertencentes a mesma Instituição de Ensino.

C. Resultados do Questionário UTAUT

Os dados coletados foram analisados estatisticamente com o uso do software SPSS (*Statistical Package for the Social Sciences*) versão número 22, utilizando-se as técnicas multivariadas de análise factorial, componentes principais e alfa de Cronbach, conforme indica o modelo UTAUT.

Primeiramente na análise factorial, agruparam-se as perguntas, reduzindo a dimensão dos dados multivariados, com o objetivo de identificar fontes subjacentes de variação comuns a duas ou mais variáveis com mesmo padrão de resposta e com isso compreender padrões de intercorrelações [7]. A análise dos componentes principais mostra uma relação entre características extraídas dos dados, visando a redução de fatores e identificando o número destes que conseguem captar a maior parte da variação dos dados. Por fim, o teste alfa de Cronbach foi feito para confirmar a confiabilidade das respostas. Procurou-se assim, confirmar os construtos descobertos por Venkatesh et al. [4], assim como o destaque de alguns construtos quando se tratando de ferramentas que utilizam RA.

Na execução do software para análise factorial, foram extraídos os fatores pelo método das componentes principais, que permite ao pesquisador identificar o número de componentes e a quantidade de informação que estes captam da variação total dos dados. O objetivo da extração de componentes é encontrar um conjunto de fatores que formem uma combinação linear das variáveis originais. Assim, se as variáveis presentes na amostra são altamente correlacionadas, elas serão combinadas para formar um fator.

Cada fator encontrado é formado pela combinação de variáveis que são altamente correlacionadas. Porém, para interpretar esses fatores e identificar quais variáveis fazem parte da sua composição, foi utilizada uma rotação do tipo Varimax. A ideia básica da rotação é identificar as variáveis que tenham alta correlação com os fatores, e variáveis que tenham baixa correlação. Este modelo, portanto, visa correlacionar uma determinada variável a um único fator, reforçando sua interpretação [7]. Os resultados da rotação podem ser vistos na Tabela 1.

É possível perceber na Tabela 1 que o autovalor do primeiro fator explica 23,63% da variação, seguido de 18,93% explicados pelo segundo fator e de 10,26% no terceiro fator, seguindo assim, até o sexto fator é atingida uma variância acumulada de aproximadamente 72,57%, que é composta pela soma das variâncias individuais. Sendo assim, esses seis fatores juntos conseguem captar grande parte da variação total dos dados da pesquisa, sendo eles os relevantes para a realização de uma análise mais detalhada.

Para provar a confiabilidade do questionário, foi calculado o Alfa de Cronbach, que mede a correlação entre respostas em um questionário por meio da análise do perfil das respostas dadas pelos respondentes. Trata-se de uma correlação média entre as perguntas. Este coeficiente varia entre 0 e 1, quanto mais próximo do valor máximo, maior a confiabilidade das escalas [7][8]. Como é possível observar na Tabela 1, o alfa dos fatores variou entre 0,909 e 0,641. O limite mínimo aceito para o alfa é de 0,60, que aponta uma boa confiabilidade da

escala do questionário. Não foi possível calcular o alfa de Cronbach do fator seis por conter apenas um item. Porém este representa 5,46% da variância explicada, não influenciando na confiabilidade da pesquisa.

Cada fator encontrado na análise factorial (Tabela 1), foi classificado e interpretado com base nos estudos da UTAUT, onde cada item do questionário estava incluído dentro de um dos cinco construtos descritos por Venkatesh et al. [4]. Dessa forma, foi possível distribuir seis fatores dentre os cinco construtos, sendo eles: Expectativa de Performance, Expectativa de Esforço, Influencia Social, Condições Facilitadoras e Intenção de Uso. Na análise factorial foram descobertos seis fatores, porém não será analisado o fator seis separadamente por explicar apenas 5,46% da variância dos dados e por conter apenas um item.

O primeiro fator que tem a maior variância explicada de 23,63%, é composto pelos itens 3A, 3B, 3C e 5E (Quadro 1). Assim, em sua maioria é formado por itens advindos do construto “Intenção de Uso”, que indica que este construto contém os principais determinantes da aceitação do uso da RA na ferramenta *Rayban Virtual Mirror*. Se tratando da ferramenta utilizada no presente estudo, a média dos itens que compõem este fator se manteve um pouco acima de três, sugerindo uma leve tendência ao uso da aplicação futuramente.

O segundo fator pode ser classificado no contexto da “Expectativa de Esforço”, pois é formado apenas pelos itens deste construto, 1A, 1B e 1C. Este fator apresenta o quanto foi fácil para os usuários explorarem a ferramenta de RA. Como mencionado por Hair et al. [9], haverá maior aceitação de certa tecnologia quanto maior for a facilidade de uso percebida. Nesse fator foi possível identificar 18,93% da variação do uso na pesquisa, apresentando fatores importantes para a aceitação do uso da tecnologia pesquisada.

O terceiro fator é composto por itens de três dos construtos identificados por: os itens 1D, 2C e 2D e 5C, sendo ele “Expectativa de Esforço” (item 1D), “Expectativa de Performance” (itens 2C e 2D) e “Condições Facilitadoras” (item 5C), que explica 10,26% da intenção de uso dos participantes. Como dois dos quatro fatores são do construto “Expectativa de Performance”, que reflete o quanto a ferramenta aumentará a qualidade do serviço oferecido, é possível concluir que este construto é importante para a aceitação do uso da RA. É possível reforçar a importância deste construto pois o sexto fator encontrado é formado apenas pelo item 2B, que sozinho explica 5,46% da variância da aceitação do uso.

O quarto fator representa o construto “Influencia Social”, pois é formado em sua totalidade pelos itens 4A, 4B e 4C. No modelo proposto por Venkatesh et al. [4], a “Influência Social” trata da percepção do usuário sobre a influencia que demais membros podem ter em sua convivência social na utilização do sistema. Este fator explica 7,9% da aceitação e uso da RA. Se tratando do *Rayban Virtual Mirror*, a média dos itens deste fator ficaram abaixo de dois, mostrando que os usuários não percebem uma influencia social na utilização do sistema.

Por último, o quinto fator explica 6,38% da variância dos dados, e é formado pelos itens 5A e 5B, ambos do construto

“Condições Facilitadoras”, que reflete o quanto os respondentes acham que as condições da ferramenta facilitam o uso da RA.

Assim, com a análise factorial foi possível associar os fatores que mais explicam a variação dos dados da pesquisa aos construtos destacados no modelo UTAUT [4], que explicam a aceitação e uso de tecnologias. Dentre estes construtos, a “Intenção de Uso” tem o maior autovalor, e por consequência, explica a maior variação da aceitação da RA como ferramenta de marketing pelos usuários, seguidamente, a “Expectativa de Esforço” também explica uma grande parcela da variação na aceitação da tecnologia, o que oferece diretrizes para a elaboração e desenvolvimento de ferramentas de marketing que utilizem RA.

V. CONCLUSÕES.

O estudo fez uma análise de fatores importantes na aceitação e uso da tecnologia de RA como uma ferramenta de marketing, por potenciais usuários do software *Rayban Virtual Mirror*.

Com aplicação de uma adaptação do questionário do modelo UTAUT, foi possível confirmar os construtos descobertos por Venkatesh et al. [4] como determinantes na aceitação de uso. Dois se destacam na aceitação e uso da RA como estratégia de marketing: a “Intenção de Uso”, que reflete a intenção dos usuários de utilizar a ferramenta futuramente, assim como fazer proveito da RA, e a “Expectativa de Esforço”, que confirma que quanto maior a facilidade no uso da tecnologia, maior será a aceitação da mesma.

Ainda foi objetivo desta pesquisa, apontar estratégias que possam ser adotadas para a melhor utilização da RA no marketing. Recomenda-se que as empresas que pretendem utilizar este tipo de tecnologia, deem atenção no que se trata do desenvolvimento de uma ferramenta que facilite o processo de compra de produtos, que vá fazer com que o usuário repita o processo futuramente e procure sempre utilizar da ferramenta, assim como atenção para o desenvolvimento de um software que seja fácil e intuitivo para o usuário. Não pode também desconsiderar a atenção para os outros três construtos descritos anteriormente, “Expectativa de Performance”, “Influencia Social” e “Condições Facilitadoras”.

Com os resultados, apesar de suas limitações, por se tratar de uma tecnologia nova e pouco difundida, e de existirem poucos estudos na área, espera-se auxiliar e orientar gestores no uso da RA como estratégia de marketing.

AGRADECIMENTOS

Os autores agradecem ao CNPq e à Universidade de São Paulo pelo apoio concedido por meio de bolsa PIBITI.

REFERENCIAS

- [1] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, B. MacIntyre. “Recent Advances in Augmented Reality”. IEEE Comput. Graph. Appl. 21, 6 (November 2001), 34-47. 2001.
- [2] N. Van Kleef, J. Noltes, S. Van Der Spoel, “Success factors for augmented reality business models.” Study tour Pixel, pp. 1-36, 2010.

Workshop de Realidade Virtual e Aumentada (WRVA 2015)

- [3] I. A. Rodello, J. R. F. Brega, "Realidade Virtual e Aumentada em Ações de Marketing." *Realidade Virtual e Aumentada: Aplicações e Tendências*, p. 45, 2011.
- [4] V. Venkatesh, M. G. Morris, G. B. Davis, F. D. Davis, "User Acceptance of Information Technology: Toward a Unified View", *MIS Quarterly*, 27(3): 425-478, 2003.
- [5] T. S. Bateman, S. A. Snell, S. A. "Administração: construindo vantagem competitiva". Editora Atlas. ISBN 85-224-1923-X., 1998.
- [6] I. A. Rodello, S. R. R. Sanches, A. C. Sementille, J. R. F. Brega "Realidade Misturada: Conceitos, Ferramentas e Aplicações." *Revista Brasileira de Computação Aplicada* 2(2), 2–16 . 2010.
- [7] J. Lattin, J. D. Carroll, P. E. Green, "Análise de dados multivariados." São Paulo: Cengage Learning, 2011.
- [8] N. K. Malhotra, "Pesquisa de marketing: uma orientação aplicada". Porto Alegre: Bookman. 2006.
- [9] J. F. Hair et al. "Análise Multivariada de Dados". Bookman, 2009.
- [10] F. D. Davis, "Perceived usefulness, Perceived ease of use, and user acceptance of Information Technology." *MIS quarterly*, p. 319-340, 1989.

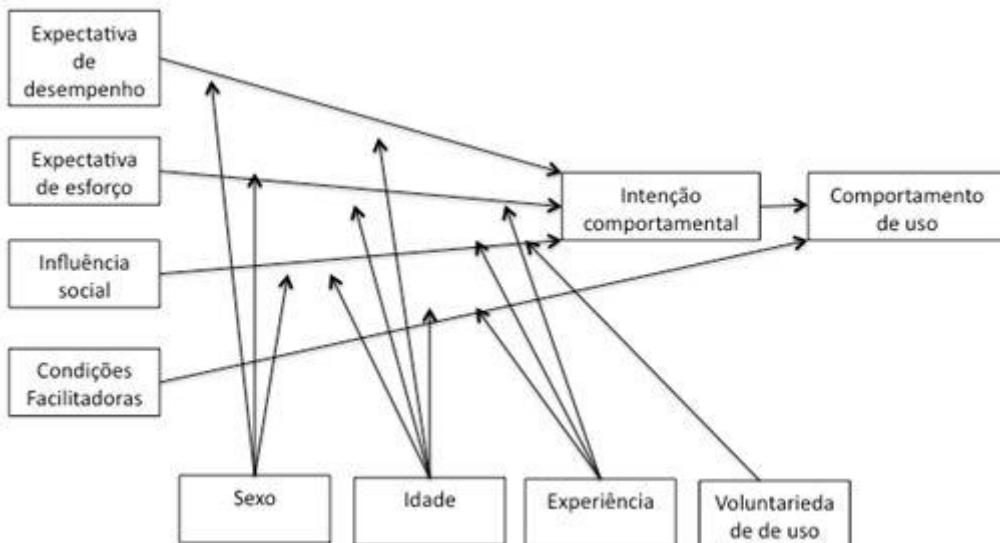


Fig. 1. Teoria Unificada de Aceitação e Uso da Tecnologia (UTAUT) [4]

QUADRO 1: QUESTIONÁRIO ADAPTADO DO MODELO UTAUT, ORIGINALMENTE FORMULADO POR [4]

Expectativa de Esforço.
1A: Minha interação com o Ray-Ban Virtual Mirror foi clara e compreensível.
1B: O Ray-Ban Virtual Mirror foi fácil de se usar.
1C: Aprender a operar o Ray-Ban Virtual Mirror foi fácil para mim.
1D: O uso do Ray-Ban Virtual Mirror é complicado, é difícil entender o que está acontecendo.
Expectativa de performance.
2A: O uso do Ray-Ban Virtual Mirror não terá nenhum efeito na minha escolha de compra.
2B: O uso do Ray-Ban Virtual Mirror pode diminuir o tempo necessário no processo de compra.
2C: Eu me senti apreensivo usando o Ray-Ban Virtual Mirror.
2D: O Ray-Ban Virtual Mirror me intimidou um pouco.
Intenção de uso.
3A: Eu pretendo usar o Ray-Ban Virtual Mirror nas próximas compras.
3B: Eu prevejo que irei usar o Ray-Ban Virtual Mirror nas próximas compras.
3C: Eu planejo usar o Ray-Ban Virtual Mirror nas próximas compras.
Influencia social.
4A: Pessoas que influenciam meu comportamento, pensam que eu deveria usar o Ray-Ban Virtual Mirror.
4B: Pessoas que são importantes para mim, pensam que eu deveria usar o Ray-Ban Virtual Mirror.
4C: Usar o Ray-Ban Virtual Mirror me proporciona um status.
Condições facilitadoras.
5A: Eu tenho o controle do uso do Ray-Ban Virtual Mirror.
5B: Eu tenho o conhecimento para usar o Ray-Ban Virtual Mirror.

Workshop de Realidade Virtual e Aumentada (WRVA 2015)

5C: O Ray-Ban Virtual Mirror não é compatível com outros sistemas que utilizo.
5D: Um guia foi disponibilizado para mim para uso do Ray-Ban Virtual Mirror.
5E: O uso do Ray-Ban Virtual Mirror é compatível com meu estilo de compra.

Tabela 1: Resultados da Análise Fatorial. Fonte: Os Autores

Tabela - Analise Fatorial (Rotação Varimax)

Itens	Fatores					
	F1	F2	F3	F4	F5	F6
3A: Pretendo usar o Ray-Ban Virtual Mirror nas próximas compras.	,913					
3B: Prevejo que irei usar o Ray-Ban Virtual Mirror nas próximas compras.	,900					
3C: Planejo usar o Ray-Ban Virtual Mirror nas próximas compras.	,890					
5E: O uso do Ray-Ban Virtual Mirror é compatível com meu estilo de compra.	,725					
1A: A interação com o Ray-Ban Virtual Mirror foi clara e compreensível.		,744				
1B: O Ray-Ban Virtual Mirror foi fácil de se usar.		,697				
1C: Aprender a operar o Ray-Ban Virtual Mirror foi fácil.		,616				
1D: O uso do Ray-Ban Virtual Mirror é complicado, é difícil entender o que está acontecendo.			,664			
2C: Me senti apreensivo usando o Ray-Ban Virtual Mirror.			,830			
2D: O Ray-Ban Virtual Mirror me intimidou um pouco.			,526			
5C: O Ray-Ban Virtual Mirror não é compatível com outros sistemas que utilizo.			,554			
4A: Pessoas que influenciam meu comportamento, pensam que eu deveria usar o Ray-Ban Virtual Mirror.				,878		
4B: Pessoas que são importantes para mim, pensam que eu deveria usar o Ray-Ban Virtual Mirror.				,688		
4C: Usar o Ray-Ban Virtual Mirror me proporciona um status.				,524		
5A: Tenho o controle do uso do Ray-Ban Virtual Mirror.					,775	
5B: Tenho o conhecimento para usar o Ray-Ban Virtual Mirror.					,795	
2B: O uso do Ray-Ban Virtual Mirror pode diminuir o tempo necessário no processo de compra.						,640
Autovalores	4,49	3,60	1,95	1,50	1,21	1,04
% Variância explicada.	23,63%	18,93%	10,26%	7,90%	6,38%	5,46%
% Variância acumulada.	23,63%	42,56%	52,82%	60,72%	67,11%	72,57%
Alfa de Cronbach	0,909	0,764	0,655	0,681	0,641	

Integrating Microsoft Kinect and Razer Hydra to Improve Natural Interaction in Virtual Environments

Vanessa Avelino Xavier de Camargo, Alexandre Cardoso, Igor Andrade Moraes, Edgard Lamounier

Virtual and Augmented Reality Group - GRVA

Federal University of Uberlândia - UFU

Uberlândia – MG, Brazil

vanessaxcamargo@gmail.com, alexandre@ufu.br, igorandradesystem@hotmail.com, elamounier@gmail.com

Abstract— Due to the increasing development of innovative and low-cost devices for games, a recent research initiative to explore this technology on Natural User Interface (NUI) and Virtual Reality (VR) applications has been identified. Thus, this paper presents an approach to combine two input devices (Microsoft™ Kinect and Razer™ Hydra) in order to enhance the interaction feedback within virtual environments. Microsoft™ Kinect offers non-touchable tracking with a wide range application in different fields of research. The controller Razer™ Hydra, on the other hand, is a relatively new sensor that has not have its full capabilities explored yet. The main characteristics related to NUI design were described to comprehend the state-of-art of interaction techniques and related works. A hybrid system composed by a full-body tracking device – Microsoft™ Kinect – and a handheld six degrees-of-freedom motion controller – Razer™ Hydra – was developed as a proof of concept to evaluate this approach and analyze its challenges and possible further applications. Preliminary results have shown that the proposed integration endorses the hypotheses of improvement.

Keywords— *Natural User Interface, input devices, hybrid tracking, interaction, Microsoft Kinect, Razer Hydra.*

I. INTRODUCTION

People naturally communicate through gestures, expressions, movements, and discover the world by looking around and manipulating physical objects. This way, Human-Computer Interaction (HCI) researchers have developed, over the last twenty years, a broad range of new interfaces that diverge from the "window, icon, menu, pointing device" (WIMP) or Direct Manipulation interaction style [1]. One of these series of interfaces is referred as Natural User Interfaces (NUI), which premise is to allow that users interact with technology the same way as they are used to interact with real world, according to what evolution and education taught them to do [2]. Effectively, this definition includes any activity or human behavior such as gestures, physical objects manipulation, virtual objects manipulation, face expressions, body movements, spoken language, sign language and real world metaphors [3] [4].

According to [5], there are seven typical characteristics of the Natural Interface, which are explained below. The design should be user-centered, i.e. in a Virtual Reality system, users

are active participants and the machine should respond to various human actions. The interface must be multi-channel, intending to make full use of one or more of the sensory and motor channels to capture the complementary characteristics of user's intention to enhance the naturalness of human-computer interaction. As opposed to precise input of keyboard and mouse, NUI might represent the inaccuracy of people's actions and thoughts, and so, the computer should understand user's requests, even to correct their mistakes. It should support high input bandwidth to fast import of large quantities of information. Finally, Natural User Interface must present at least one of these three kinds of interaction – voice-based, image-based and behavior-based. Voice-based interaction means computer technology to study how people interact through natural voice or machine synthesized voice, approaching computer and user's language. Image-based interaction can be divided into processing, recognition and image perception, to understand the image and react to it. Yet, in behavior-based interaction, the computer recognizes human behavior through the positioning, tracking, movement and expression characteristics of human body to understand human actions and behavior, and respond to it (feedback).

NUI is often confused with Reality-Based Interaction (RBI). However, RBI, also known as Reality User Interface (RUI), is a design strategy that increases the realism of interface objects and allow users to interact even more directly with them—using actions that correspond to daily practices within the non-digital world. Some examples of these post-WIMP interaction styles are: virtual, mixed and augmented reality, tangible interaction, ubiquitous and pervasive computing, context-aware computing, handheld, or mobile interaction, perceptual and affective computing as well as lightweight, tacit or passive interaction [1].

Hardware industries have developed a variety of input devices to make the use of this interface possible, and, consequently, new fields of application are emerging. Devices such as Nintendo™ Wii, PlayStation™ Move, Microsoft™ Kinect and Razer™ Hydra are examples of devices that use gesture recognition to provide more natural interaction than common ones (keyboard and mouse). Researchers in different areas, like education, simulation, rehabilitation, training, etc., are using them and stating their powerfulness and limitations.

Nintendo™ Wii Remote, PlayStation™ Move and Razer™ Hydra are motion gaming systems, which work through controller-based accelerometers and gyroscopes to sense tilting, rotation and acceleration [6]. Alternatively, Microsoft™ Kinect enables interactions between users and a game without the need to touch a controller [7]. Since its release, in 2010, it became very popular among gamers and researchers, who are extracting the most of its features.

Having the NUI characteristics in mind, especially the ones related to user-centered and multi-channel design, we hypothesized that the integration of two of these devices must enhance the natural interaction in three dimension (3D) environments, dealing with each other weaknesses. The devices chosen were the Microsoft™ Kinect, due to its non-invasive and wide spread technology, and the Razer™ Hydra, because of its precisely hand tracking and relatively unexplored know-how. The motion data from Kinect – even with its shortcomings – is well suited for establishing coarse-grained spatial relations in actions such as pointing, gesturing shape, or performing other continuous gestures. Hand-held controllers, on the other hand, are appealing, because most humans are naturally accustomed to using hand tools. They also provide greater tracking accuracy with less latency than inexpensive full-body tracking devices like Microsoft™ Kinect. Additionally, Razer™ Hydra is equipped with buttons that offer multiple benefits when compared to issuing commands via gesture recognition: buttons allow instantaneous and error-free signaling with haptic feedback [8].

Along with Bowman [9], a hybrid tracking approach put more than one tracking technology together to help increase accuracy, reduce latency, and provide a better overall 3D interaction experience. The major difficulty of this attempt is that more components produce more complexity. Still, the extra complexity is warranted if tracking accuracy is significantly improved.

As a proof of this concept, we develop a 3D training system prototype in which an inexperienced electrical maintenance technician has the opportunity to practice his or her job virtually, avoiding the risk of receiving the electric shock. Preliminary results confirm that the combination of both devices improves the natural interaction, opening a wide possible application area, as explained in the next sections of this paper. Section 2 contains a detailed description of the referred devices, while section 3 lists related works. Section 4 explains the methodology for the development of the prototype, since the reasons to choose the devices until the system architecture and implementation. Section 5 discusses the issues related to the approach. Finally, Section 6 presents the summary of the research, with its conclusions and future work.

II. DEVICES DESCRIPTION

Here we present an overview of each device, enumerating their main attributes, hardware characteristics and interaction constraints.

A. Razer™ Hydra

This device, developed by a partnership between Sixense and Razer, has a combined system, with a base station and two controllers.



Fig. 1. Razer™ Hydra overview.

Each controller is a thumb-ergonomic analog stick for fluid control with a non-slip satin grip surface and true six degree-of-freedom (DoF) magnetic motion tracking, and has four Hyperresponse action buttons, a rapid-fire trigger and a bumper, as shown in Figure 1. The base station has a low-power magnetic field with low power consumption, an ultra-precise sensor for one millimeter and one-degree tracking and low latency feedback.

Using magnetic motion sensing, the best available technology of its kind, the Razer™ Hydra's compact base station can compute the exact location and orientation of controllers in your hands, down to a millimeter and degree. Combined with Razer-grade ultra-low latency, the user can get a fluid and unhindered experience that is accurate and precise [10]. Table 1 shows a comparative among other sensors available and the magnetic sensor, which stands out for absolute position and orientation tracking in 3D space, in addition to six degrees of freedom.

The controllers and the base station have three magnetic coils each. These coils work together with the amplifier circuit, the digital sign processor and positioning algorithm to translate field data into orientation and position data. This allows 3D interaction with the virtual environment, enabling more spatial control than with a common mouse.

Besides that, the wire connection between the base station and the controllers, allied with the Hyperresponse technology, support ultra-low latency, providing virtually instantaneous feedback between user commands and display [10]. The device's precision can be useful in applications such as medical images, 3D modeling, virtual reality and training softwares, besides games. Sixense provides the free download of SDK to interested developers, in addition to forums and direct contact with the software developers. The platform is easy to implement through available libraries [10].

In general, magnetic tracking systems are accurate to within 0.1 inch in position and 0.1 degree in orientation. Nonetheless, their main disadvantage is that any ferromagnetic or metal objects present in the room may distort the magnetic field, reducing the precision [9].

TABLE I. SENSOR COMPARATIVE.

Characteristics	Type of sensor			
	Magnetic	Accelerometer	Ultrasonic	Camera
Degrees of freedom	6	3 (+3 with gyroscope)	3 (+3 with gyroscope)	3 (+3 with gyroscope)
Reach	3m-5m	10m	3m-5m	3m-5m
Absolute position	✓	✗	✓	✓
Absolute orientation	✓	✗	✗	✗
Reliability	✓	✗	✗	✗
Low latency	✓	✗	✗	✗

B. Microsoft™ Kinect

This device's main capabilities are using an RGB camera that allows face recognition; a depth sensor that enables it to scan the surroundings; built-in microphone to capture sound; and a microchip to track and recognize movements, as shown in Figure 2.

Its depth sensor captures 3D videos with an infrared projector and a CMOS (Complementary Metal-Oxide Semiconductor) image sensor, which measures the distance between its pixels and the objects. These two sensors work together to identify the room and the user in 3D. The depth sensor perceives the nearer objects as brighter and the further ones in a darker tone. It uses the generated image, by machine learning, to identify a human body, 48 joints per user. Then, a library of pre-recorded positions processes the image in case the Kinect does not capture all the joints, to recognize movements and positions. Microphones enable it to capture sound commands [11].

Albeit it is constituted by a variety of sensors, the optical tracking promoted by computer vision techniques is Kinect's worthiest feature. On the other hand, it is also responsible for some drawbacks. Setting up vision-based tracking systems can be difficult because many parameters must be set in order to track the user or physical objects properly. Yet, occlusion is stated as its major disadvantage [9].

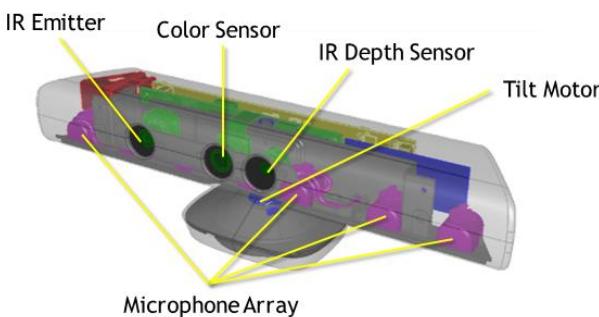


Fig. 2. Microsoft™ Kinect overview.

III. RELATED WORK

The use of Microsoft™ Kinect in NUI researches is extensively well known. Although Microsoft primarily designed it for games, its features turned it into a novel and cheap way to achieve good results in different types of application.

Many works focus on its algorithms to solve problems. [12] analyzes Kinect as a 3D measuring device, experimentally investigating depth measurement resolution and error properties and making a quantitative comparison of Microsoft™ Kinect accuracy with stereo reconstruction from SLR cameras and a 3D-TOF camera. [7] presents a comprehensive review of recent Kinect-based computer vision algorithms and applications, including preprocessing, object tracking and recognition, human activity analysis, hand gesture analysis, and indoor 3-D mapping. [13] proposes a robust detection scheme to detect and classify different types of shadows, caused by occlusion of the Microsoft™ Kinect's structured light, based on their distinct local shadow patterns as determined from geometric analysis, without assumption on object geometry.

On the other hand, NUI researchers are not enough exploring the Razer™ Hydra. Most of works use it as a more precise joystick, providing natural interaction and accuracy in their systems. [14] results give insight into the usability of 3D User Interfaces (3D UI) for adolescents with Autism Spectrum Disorder (ASD), using Razer™ Hydra as the interaction device. [15] proposes a low-cost simulation of minimally invasive surgery having Hydra as the most suitable, low-cost and currently available hardware to use as controllers. In [16] patients use the Razer™ Hydra game interface to play a series of custom games designed to improve focusing, decision making (executive function), short-term and long-term memory, progressing in difficulty over 6 weeks of therapy.

Yet, some works have integrated Microsoft™ Kinect and/or Razer™ Hydra with other devices. [17] describes the development of a system to support training of young drivers to make the learning of the traffic rules more friendly and exciting. To achieve this, they developed a game called KinardCar, where they use the Microsoft™ Kinect sensor integrated with an Arduino microcontroller. [8] provides 3D user interface building blocks for creating immersive VR applications with spatial interaction and stereo 3D graphics, while supporting affordable VR peripherals like Microsoft™ Kinect, PlayStation Move, Razer™ Hydra, and Oculus Rift. [18] shares the challenges encountered while setting up an open source game development platform called Cryengine and while integrating this engine with VR gadgets such as Kinect, audio system, Razer™ Hydra, and Head Mounted Display. [19] presents a VR demo called TurboTuscany, which combines a Kinect controlled full body avatar with Oculus Rift head-mounted-display. They also implemented three positional head tracking schemes that use Kinect, Razer™ Hydra, and PlayStation™ Move controllers.

As presented in this section, the union of two or more devices amplify the range of possible applications in NUI researches. This approach extends the definition of RBI, increasing the awareness of body and environment [1]. Although valid and well grounded, in the given works, the choice of which device would be most suitable seems based only on the purpose of the application, without considering 3D interfaces design

orientations. The developed prototype flourished its choice of devices according to important factors found in the specialized literature.

IV. METHODOLOGY

This section describes the methodology used to choose the input devices, based on literature, and the elaboration of the prototype, specifying its architecture and implementation.

A. Choice of input devices

Many factors must be considered when choosing an input device for a particular 3D User Interface. Device ergonomics, the number and type of input modes, the available technique to device-mapping strategies, and the type of tasks the user will be performing play important role in this assignment [9].

According to Bowman [9], devices should be lightweight, require little training and provide a significant transfer of information to the computer with minimal effort. Both Microsoft™ Kinect and Razer™ Hydra present these characteristics. Microsoft™ Kinect is non-touchable and non-wearable, resulting in no weight attached to user's body. These qualities are also the reason why it requires minor training and effort regardless user-computer information transference and interaction. Even though Razer™ Hydra is wired and touchable, its light and ergonomic format adapts naturally to user's hands, being smooth and fine to hold. Besides this, for the planned application, handheld devices are right and proper, since in a real situation, the technician will probably work with tools in hands.

Another crucial factor in choosing the correct device is the number and type of input modes. As exposed in Table 1, the 6 DoF provided by Razer™ Hydra is the best way to interact with a 3D environment, making possible rotation and translation in the three axis (x, y, z). Microsoft™ Kinect, in turn, provides 3 DoF plus more three with the aid of a gyroscope.

For immersive interaction systems, it is important for the UI to deliver information about the user or physical object's location in 3D space. Having accurate tracking is a fundamental of making interaction techniques usable within virtual environments [9]. By way of it is defined in Section 2, the magnetic tracking of Razer™ Hydra and the optical tracking of Microsoft™ Kinect were chosen down to their potentialities and vulnerabilities.

B. Prototype architecture and implementation

The proposed prototype has been developed using Unity 3D, which is a powerful engine with a variety of tools that can be utilized to meet specific needs. It has been chosen because it is robust, multiplatform, widely used, and its editor is intuitive. Unity 3D is available as a free personal edition and a 75-dollar-per-month professional edition. Both versions come with all features, with the possibility to purchase the charged-version facilities apart. For Razer™ Hydra interaction, the Sixense plugin, developed by Sixense Studios and available for free download at the Asset Store, was used. Its scripts give parameters to implement each button action, such as point, hold, fist and grip. For Microsoft™ Kinect interaction, OpenNI framework was used, since the Unity support for Microsoft™ Kinect SDK has only been released in October, 2014. The system architecture is exemplified in Figure 3.

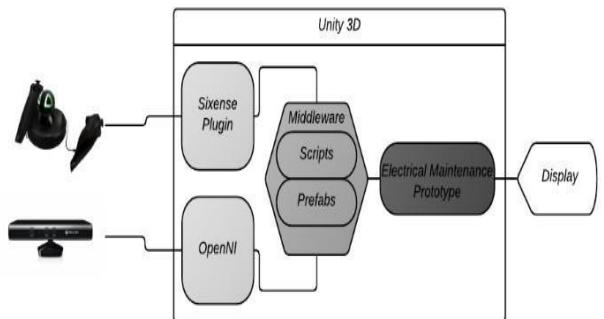


Fig. 3. Prototype architecture.

The proof of concept system developed was based on an electrical maintenance technician who has to deal with a problem in an electricity system. Figure 4 shows the first scene of the system, in which the user, represented by an avatar, has to walk up to his toolbox to choose the appropriate tools and Personal Protective Equipment (PPE). Microsoft™ Kinect tracks user's limbs as he/she walks in place, i.e. lifts his/her feet up and down, but not physically translates. Figure 5 shows the PPE available for the user, who has the option to pick them up or not.

As soon as the user chooses his equipment, he/she must pick up the controllers of Razer™ Hydra, which represents his/her tools, and go to the ladder to climb up it. To calibrate the sensor, user should position both controllers on his /her shoulders, and then press left and right triggers. It is a specification from Sixense plugin, which will correctly track user's hand in its proper real position. As soon as the action of climbing the ladder could not naturally be represented by gestures, it is automatically done once the avatar gets sufficiently close to the ladder. It is now time to work on, taking care of what to do and how to do the service. At this moment, the interaction switches from Kinect-based to Hydra-based, transparently for the user.

Figure 6 presents this interaction, as the operator works on the electricity system.



Fig. 4. Unity scene overview.



Fig. 5. Personal Protective Equipment options available for the user.

V. RESULTS

Seven researchers that work in our laboratory tested the prototype. In view of the fact that the hypotheses elicited as goal was to improve interaction in virtual environments by combining two different input devices, their background in interaction design and implementation was suffice to the prototype first evaluation.

After finishing the given task, they answered a questionnaire with questions regarding NUI design and interaction. These questions are part of the design guidelines proposed by Widgor and Wixon [20]. According to them, to build a user experience that is natural to the user, one must:

- 1) *Create an experience that, for expert users, can feel like an extension of their body.*
- 2) *Create an experience that feels just as natural to a novice as it does to an expert user.*
- 3) *Create an experience that is authentic to the medium—do not start by trying to mimic the real world or anything else.*
- 4) *Build a user interface that considers context, including the right metaphors, visual indications, feedback, and input/output methods for the context.*
- 5) *Avoid falling into the trap of copying existing user interface paradigms.*



Fig. 6. Third person view of the scene.

The searched results were related to rule number 4. All the researchers answered affirmatively about the input methods for the context approach, highlighting that this combined system has potentialities to be used into training and simulation systems to reproduce tasks that are either dangerous, expensive, or impracticable to recreate.

VI. DISCUSSION

As far as integrating devices is concerned, it is well known that challenges will be faced. Some issues faced during the development of the prototype are drawn out in this section.

Microsoft™ Kinect's hand detection is unstable if the hand is far from the sensor, due to its low resolution. Besides this, holding a wand seems more natural for the users, since in the real world, there will be something on their hands. This way, a hand-based controller, such as the Hydra, is better suited for precision tasks as the one chosen for the proof of concept.

The full body tracking that Microsoft™ Kinect provides, although, is efficient and accurate under determined circumstances, especially in applications in which users stand upright near the device, so it reduces the camera view-range limitation, occlusions, and unpredictable human poses [7].

Similarly, Hydra's limitations rely primarily on the interference of any metal object placed near it. Yet, as it is not wireless, the range of user's space must be controlled.

VII. CONCLUSION

In this paper we have discussed the prominent possibilities emerged from integration of two tracking devices aiming to enhance interaction in virtual environments. The elected NUI devices were Microsoft™ Kinect and Razer™ Hydra.

Thus, the integration of them seemed to be a reliable and effective approach to provide natural interaction in virtual environments at a low cost value. The gained results may be applied to different areas of study, including education, training and simulation.

Future work is intended to integrate other devices, such as Oculus Rift and Leap Motion to evaluate their contributions to improve user's level of immersion. Yet, it is envisioned to have experts (maintenance technicians) evaluate the system, as to carry out advantages and disadvantages of it in a real field of application.

REFERENCES

- [1] Jacob, R.J., Girouard, A., Hirshfield, L.M., Horn, M.S., Shaer, O., Solovey, E.T. and Zigelbaum, J. 2008. Reality-based interaction: a framework for post-WIMP interfaces. Proceedings of the SIGCHI conference on Human factors in computing systems (2008), 201–210.
- [2] Valli, A. The design of natural interaction. *Multimedia Tools and Applications*, p. 295–305, 2008.
- [3] Iacolina, S. et al. Natural interaction and computer graphics applications. In: The Eurographics Association. *Eurographics Italian Chapter Conference 2010*. [S.I.], 2010. p. 141–146.
- [4] Jain, J.; Lund, A.; Wixon, D. The future of natural user interfaces. In: *Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems*. New York, NY, USA: ACM, 2011. (CHI EA '11), p. 211–214. <http://doi.acm.org/10.1145/1979482.1979527>
- [5] Liu, W. Natural user interface- next mainstream product user interface. In: *Computer-Aided Industrial Design Conceptual Design (CAIDCD)*,

Workshop de Realidade Virtual e Aumentada (WRVA 2015)

- 2010 IEEE 11th International Conference on. [S.l.: s.n.], 2010. v. 1, p. 203–205.
- [6] Rouse, M. WhatIs.com. 2011. <<http://whatis.techtarget.com/definition/natural-user-interface-NUI>>
 - [7] Jungong Han; Ling Shao; Dong Xu; Shotton, J. Enhanced Computer Vision with Microsoft Kinect Sensor: A Review. *Cybernetics, IEEE Transactions on*, vol.43, no.5, pp.1318, 1334, Oct. 2013, doi: 10.1109/TCYB.2013.2265378
 - [8] Tuukka M. Takala. 2014. RUIS: a toolkit for developing virtual reality applications with spatial interaction. In Proceedings of the 2nd ACM symposium on spatial user interaction (SUI '14). ACM, New York, NY, USA, 2014, 94-103, doi:10.1145/2659766.2659774
 - [9] Bowman, D., Kruijff, E., LaViola, J., Poupyrev, I. (2005). 3D User Interfaces: Theory and Practice. Boston: Addison-Wesley.
 - [10] RazerZone.Enter the Hydra: PC Gaming Motion Sensing Controllers. <<http://www.razerzone.com/minisite/hydra/what>>
 - [11] Boas, Y. A. G. V. Overview of virtual reality technologies. Interactive Multimedia Conference 2013. 2013.
 - [12] Smisek, J.; Jancosek, M.; Pajdla, T. 3D with Kinect. Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on, vol., no., pp.1154, 1160, 6-13 Nov. 2011, doi: 10.1109/ICCVW.2011.6130380
 - [13] Teng Deng; Hui Li; Jianfei Cai; Tat-Jen Cham; Fuchs, H. Kinect Shadow Detection and Classification. Computer Vision Workshops (ICCVW), 2013 IEEE International Conference on, vol., no., pp.708, 713, 2-8 Dec. 2013, doi: 10.1109/ICCVW.2013.97
 - [14] Chao Mei, Lee Mason, and John Quarles. 2014. Usability issues with 3D user interfaces for adolescents with high functioning autism. In Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility (ASSETS '14). ACM, New York, NY, USA, 99-106, doi:10.1145/2661334.2661379
 - [15] Kasper Grande, Rasmus Steen Jensen, Martin Kraus, and Martin Kibsgaard. 2013. Low-cost simulation of robotic surgery. In Proceedings of the Virtual Reality International Conference: Laval Virtual (VRIC '13). ACM, New York, NY, USA, Article 6, 4 pages, doi:10.1145/2466816.2466823
 - [16] Burdea, G.C.; Defais, C. Repetitive bimanual integrative therapy (RABIT) virtual rehabilitation system. *Virtual Reality (VR)*, 2013 IEEE, vol., no., pp.99, 100, 18-20 March 2013, doi: 10.1109/VR.2013.6549381
 - [17] Amorim Vaz, F.; de Souza Silva, J.L.; Sol Dos Santos, R. KinardCar: Auxiliary Game in Formation of Young Drivers, Utilizing Kinect and Arduino Integration. *Virtual and Augmented Reality (SVR)*, 2014 XVI Symposium on, vol., no., pp.139, 142, 12-15 May 2014, doi: 10.1109/SVR.2014.49
 - [18] Zainudin, A.R.R.; Yusof, A.M.; Rusli, M.E.; Yusof, M.Z.M.; Mahalil, I. Implementing immersive virtual reality: Lessons learned and experience using open source game engine. *Open Systems (ICOS)*, 2014 IEEE Conference on, vol., no., pp.105, 111, 26-28 Oct. 2014, doi: 10.1109/ICOS.2014.7042638
 - [19] Takala, T.M.; Matveinen, M. Full body interaction in virtual reality with affordable hardware. *Virtual Reality (VR)*, 2014 IEEE, vol., no., pp.157, 157, March 29 2014-April 2 2014, doi: 10.1109/VR.2014.6802099
 - [20] Daniel Wigdor and Dennis Wixon. 2011. *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture* (1st ed.). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA

Reuso de Objetos Virtuais a Partir de Grafos de Cena para Concepção de Ambientes Virtuais

Reuse of Virtual Objects from Scene Graphs for the Conception of Virtual Environments

Camilo Barreto, Alexandre Cardoso, Edgard Lamounier, Paulo Prado, Leandro Mattioli, Alexandre Carvalho.
 Universidade Federal de Uberlândia - Grupo de Realidade Virtual e Aumentada - Uberlândia, Minas Gerais
 E-mail: {barretojuniormail, alex.cardoso.ufu, elamounier,
 leandro.mattioli, acs.carvalho10}@gmail.com, prmprado@uol.com.br

Abstract—The use of Virtual Environments in system control operation, in particular power systems (power generation and propagation), provides new paradigms to daily activities of monitoring and controlling. The existence of a three-dimensional geometric model, true to the real structures, when used to develop Virtual Reality systems, allows the acquisition of mental models that improve communication between Control Room and field operators. Besides this, it enables the conception of simulators and training environments.

This paper proposes a methodology of creating scene graphs applied to electric power substations during model groups' creation phase, in order to automate the virtual scenes replication.

Index Terms—scene graph; window; virtual environment; VRS;

I. INTRODUÇÃO

A Realidade Virtual (RV) é, antes de tudo, uma “interface avançada do usuário” para acessar aplicações executadas no computador, tendo como características a visualização de, e movimentação em, ambientes tridimensionais em tempo real e a interação com elementos desse ambiente [1].

Para construir ambientes de Realidade Virtual, tais como simuladores, um dos principais desafios é a otimização no desenvolvimento de elementos que aparecem repetidamente e em larga escala. Tais elementos quando não são desenvolvidos de forma eficiente, atrasam a construção de uma cena devido ao tempo gasto com modelagem, posicionamento e conexão com elementos da cena [2].

A criação de Sistemas de Realidade Virtual (SRV) aplicado à subestações (SE) de energia elétrica (Figura 1), revelou uma gama de desafios importantes para modelagem de Ambientes Virtuais (AV). Um dos problemas na criação de SRV, é a geração em grande escala de AV tendo em vista a baixa quantidade de insumos para sua produção. Vale salientar outros problemas na criação de AV, como o posicionamento de equipamentos, complexidade em conectá-los e o tempo de produção.

O objetivo deste trabalho é apresentar um método utilizando grafos de cena para modelagem automática de Ambientes Virtuais construídos por uma ferramenta de diagramas, visando a padronização de arranjos, reuso de elementos e conexão entre modelos, reduzindo recursos e tempo de desenvolvimento.

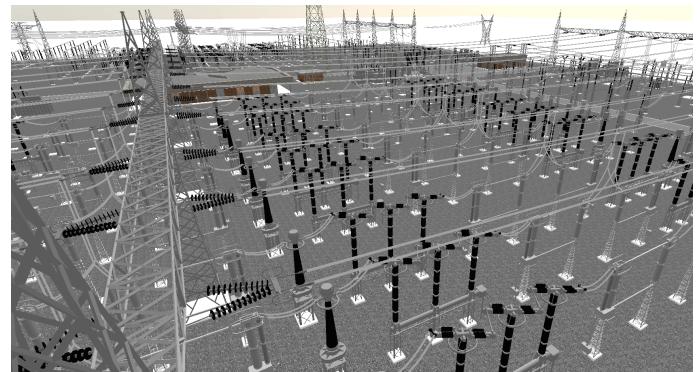


Fig. 1. Ambiente Virtual da Subestação de Ribeirão das Neves.

II. GRAFOS DE CENA E ARRANJOS FÍSICOS DE SUBESTAÇÕES

A. Grafos de Cena

Segundo Walsh [3], Grafos de cena são estruturas de dados usadas para organizar e gerenciar o conteúdo da cena hierárquicamente.

O grafo de cena (Figura 2) é formado, por nós conectados por arestas compondo um grafo acíclico direcionado. Cada nó possui um conjunto de atributos que podem, ou não, influenciar seus nós conectados. Os nós são organizados de uma maneira hierárquica correspondendo semântica e espacialmente com o mundo modelado [4].

Os nós podem ser divididos em três categorias: nó raiz, nós intermediários que são chamados de nós internos ou nós de agrupamento e os nós folha que estão localizados no final de um ramo. O nó raiz é o primeiro nó do grafo e todos os outros nós estão ligados a ele direta ou indiretamente. Os nós internos possuem várias propriedades, sendo o uso mais comum o de representar transformações 3D (rotação, translação e escala) [4].

Os grafos de cena utilizam um princípio chamado de herança, isso significa que os nós filhos armazenam o estado do nó pai, onde o estado significa a posição, orientação e

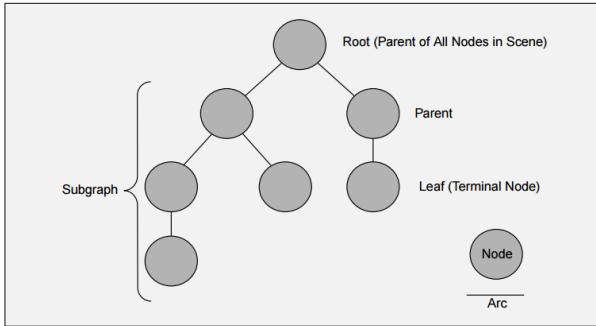


Fig. 2. Corpo de um grafo de cena [4].

escala dos elementos no ambiente virtual. A herança é uma propriedade dos grafos de cena que determina que cada nó deve herdar as propriedades de seus ancestrais.

B. Arranjos de Subestações

Subestação é um conjunto de equipamentos industriais interligados com objetivo de controlar o fluxo de potência, modificar tensões e alterar a natureza da corrente elétrica, assim como garantir a proteção do sistema elétrico [5].

O arranjo físico de uma subestação, é um conjunto de equipamentos que fazem parte do sistema de geração, transmissão e distribuição. Estes conjuntos de equipamentos são separados por nível de tensão, localização e a linha de transmissão. Segundo [6] existem 8 tipos de configurações de arranjos. São eles:

- Barra Simples;
- Barra Simples com By-Pass;
- Barra Simples Seccionado;
- Barra Principal e Transferência;
- Barra Dupla, um Disjuntor;
- Barra Dupla com Dois Disjuntores;
- Barra Dupla com Disjuntor e Meio;
- Barramento em Anel.

As configurações de arranjos não são aplicáveis somente a uma SE, outras possuem o mesmo tipo de configuração, podendo ser diferentes apenas no nível de tensão e modelo de equipamentos. Com a padronização dos arranjos, torna-se possível criar *templates* de arranjos que possam ser reutilizados em qualquer outra SE.

III. TRABALHOS CORRELATOS

Esta seção apresenta uma breve descrição de trabalhos relacionados ao desenvolvimento de ambientes virtuais de subestações de energia elétrica.

A. Scene Graph Organization and Rendering in 3D Substation Simulation System

A pesquisa de Guangwei [7] enfatiza a criação de modelos com níveis de detalhes diferentes para cada equipamento (Figura 3), abordando a técnica de grafos de cena para otimizar

a renderização dinâmica dos modelos, melhorando o processamento gráfico e gerando uma taxa de atualização de 30Hz para computadores em geral. A vantagem deste trabalho é a utilização de grafos de cena para a renderização do ambiente virtual, melhorando a performance do sistema, seja qual for o computador. Uma desvantagem é o sistema utilizar modelos com alguns níveis de qualidade diferentes para realizar o LOD (*Level of Detail*), que poderia ser automática utilizando apenas um modelo.

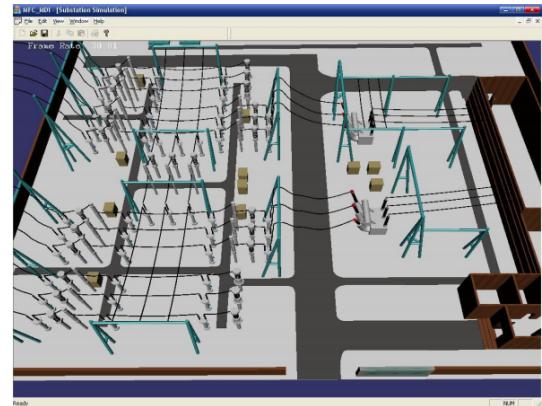


Fig. 3. Subestação com renderização dinâmica [7].

B. 3D Virtual Models Applied in Power Substation Projects

O trabalho de Quintana [8] apresenta uma metodologia para criação de modelos aplicados a subestações de energia elétrica do México (Figura 4). São apresentadas etapas para criação de um passeio virtual, coleta de informações, modelagem de equipamentos e renderização do passeio virtual. É abordado também, a complexidade de adequação de ambientes virtuais existentes a reformas nos pátios da SE. A vantagem deste trabalho é a realidade que a cena possui. Uma desvantagem é que o sistema não possui suporte para controle e monitoramento de sistemas críticos de energia em tempo real. Também não possui a produção extensiva de SE, se limitando apenas em uma.



Fig. 4. Visão aérea do modelo virtual da subestação de Xochimilco [3].

IV. DESENVOLVIMENTO

Nesta seção serão destacados os processos de criação de um método de modelagem de arranjos lógicos, baseados em grafos de cena para concepção de Ambientes Virtuais aplicados a subestações de energia elétrica.

O projeto do SRV é composto por várias etapas de desenvolvimento e segue um pipeline de produção. Na etapa de modelagem de elemento 3D, é necessário entrada de dados (fotos, CAD, dimensões) para a criação de modelos realísticos, aplicando as mesmas texturas e dimensões dos equipamentos reais, dispondo de um ambiente virtual fidedigno ao real.

Após a criação dos modelos 3D, é feita a preparação do cenário virtual. Nessa etapa os modelos criados são incluídos na cena, um dos métodos para posicionamento semiautomático de componentes foi descrito no trabalho de [9]. Trata-se de uma ferramenta de posicionamento de elementos a partir do reconhecimento de símbolos de plantas CAD.

Com a utilização desta ferramenta o tempo de produção foi reduzido consideravelmente, entretanto existem vários obstáculos na criação dos arranjos, como a grande quantidade de elementos e a complexidade de criação dos cabos condutores que os interligam.

O desenvolvimento da proposta foi dividida em módulos: criação de arranjos em uma ferramenta tipo *Windows* e utilização da ferramenta para posicionar os *Prefabs* no cenário.

A. Criação de Arranjos

Os arranjos possuem padrões que possibilitam sua reutilização em outros ambientes, alterando apenas os elementos que os compõem. Como foi descrito anteriormente, os arranjos possuem oito tipos de configuração, visando esse benefício foi desenvolvido uma ferramenta do tipo *Windows* (Figura 5a) para criar e utilizar *templates* de arranjos, possibilitando a rápida criação dos grafos de cena.

A ferramenta constitui de uma interface para criação de diagramas que define como o arranjo é criado, ou seja, uma interface bidimensional para criar grupos de objetos tridimensionais. Ela é composta por duas áreas de trabalho, o *Diagram Editor* e *Property Node*. O *Property Node* é o local que exibe todos os dados do bloco selecionado, possibilitando realizar configurações e inserção de modelos. O *Diagram Editor* é o local de criação e edição dos diagramas, nele é possível inserir blocos e criar qualquer configuração de arranjo. Possui os seguintes blocos:

- Bloco *Init*: inicia o diagrama;
- Bloco *Object*: insere ao fluxo do diagrama um modelo 3D;
- Bloco *End*: finaliza o diagrama.

Para criar um grafo de cena, primeiro é necessário inserir o bloco *Init* que define o começo do diagrama. Após a inicialização, é possível inserir os blocos do tipo *Object* que são correspondentes aos modelos 3D que serão incluídos no *Prefab*, cada bloco possui um campo para selecionar qual modelo ele representa. Os blocos são interligados por um fio que define o ordenamento e a ligação física, automatizando o processo de conexão entre eles. Para finalizar o grafo de cena, o bloco do tipo *End* deve ser inserido, isso define que o diagrama foi encerrado naquele ponto. Ao selecionar o bloco *End* um menu com configurações é exibido, é possível definir as distâncias entre elementos, sua posição, rotação e escala.

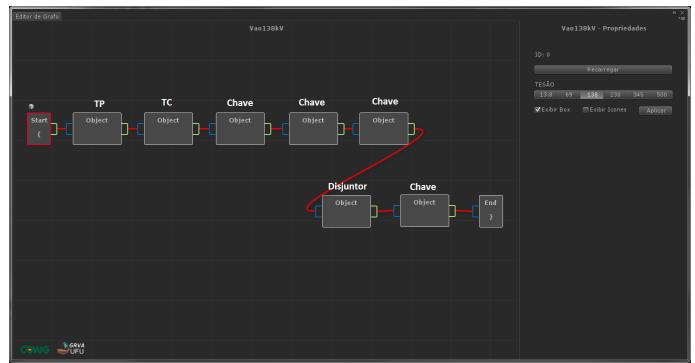


Fig. 5a. Ferramenta Windows para criação de grafos de cena.

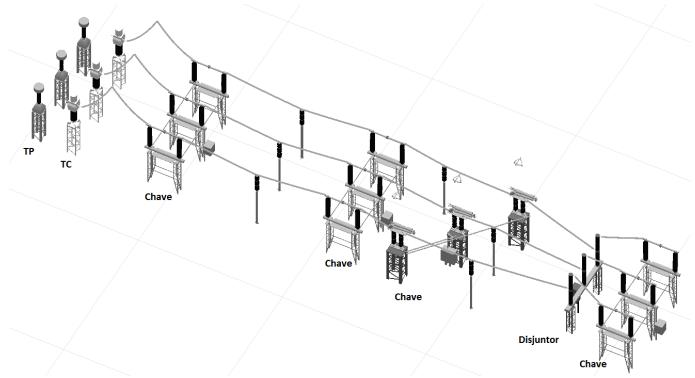


Fig. 5b. Grafo de cena criado pela ferramenta.

Finalizado, um grupo de objetos é exibido na tela, a Figura 5b mostra o resultado da criação.

O SRV em modo de edição possui vários *Scripts* Editores para a manipulação e configuração dos objetos 3D, os vãos criados a partir da ferramenta possuem um *Script* para definir a nomenclatura, tensão e algumas opções de visualização para o desenvolvedor da cena. Quando criado uma versão final esses *Scripts* podem prejudicar a performance da aplicação, para solucionar esse problema, foi adicionado uma função para retirar todos, restando apenas as estruturas estáticas dos modelos. A Figura 6 exibe as configurações possíveis para um vão do tipo barra dupla e disjuntor e meio.

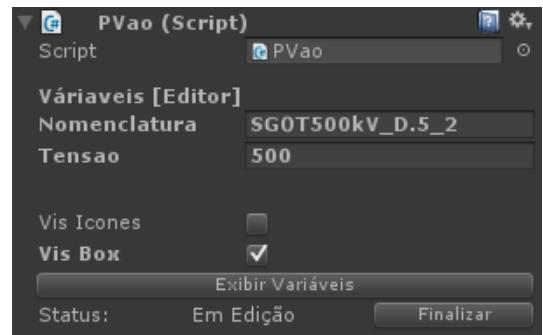


Fig. 6. Script para configuração de um arranjo.

A Figura 8 representa o grafo de cena da Figura 5b, trata-

se de uma hierarquia de modelos, onde o *Elemento Pai* é um *GameObject* que possui atributos de rotação, translação e escala. Após criação do grafo de cena, ele é armazenado em um tipo de arquivo chamado *Prefab*, podendo ser armazenado em bibliotecas e reutilizado posteriormente em outros AV.

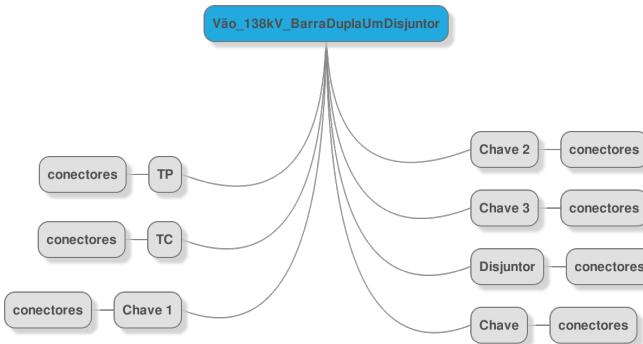


Fig. 7. Corpo do grafo resultante.

B. Posicionamento de Arranjos

Trata-se do processo semiautomático de posicionamento dos arranjos. A ferramenta utilizada neste processo é uma versão adaptada desenvolvida por [9]. O método utilizado para posicionamento foi feito pelo reconhecimento de símbolos de plantas CAD e transformando as informações em coordenadas. Os símbolos devem ser coerentes com os modelos 3D, respeitando o ponto de referência e o *pivot* do arranjo. Os nomes dos modelos devem corresponder aos nomes dos símbolos reconhecidos. A ferramenta instancia dinamicamente o arranjo, aplicando as coordenadas e matrizes de transformação. A Figura 7 apresenta o cenário virtual da subestação de São Gotardo após utilizar o processo de criação de grafos e posicionamento automático. A subestação possuí três níveis de tensão representadas em cores, para melhorar a visualização dos arranjos foram inseridas caixas de cores.

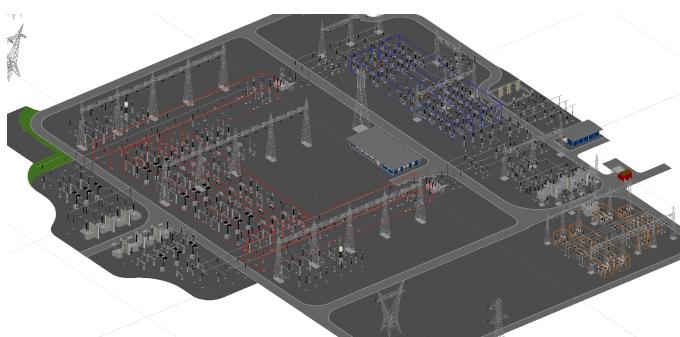


Fig. 8. Criação semiautomática da SE de São Gotardo.

V. CONCLUSÃO

A reprodução de cenários virtuais possui fatores que influenciam o tempo de entrega e a qualidade do produto, o

posicionamento manual de elementos é dificultado pela vista em perspectiva dos softwares de desenvolvimento, podendo enganar os desenvolvedores e gerar erros. Com o uso da ferramenta proposta neste trabalho torna-se possível a criação de arranjos lógicos utilizando grafos de cena em um ambiente bidimensional para criação ambientes tridimensionais, tornando o processo mais automático, eliminando trabalhos manuais e erros de desenvolvimento.

Foi realizada uma etapa inicial de validação da ferramenta para o desenvolvimento do sistema apresentado, pode-se relatar que a utilização de grafos de cena, apresentou resultados satisfatórios, quantitativos e qualitativos. Por se tratar de uma automatização na criação de cenários virtuais, os métodos manuais foram retirados e os prazos de reprodução foram reduzidos significativamente. A realidade e fidelidade do cenário reproduzido é prova de que um sistema automatizado e que segue padrões possui melhores resultados e qualidades.

A metodologia aqui descrita foi aplicada com sucesso na produção de ambientes de realidade virtual de subestações de energia elétrica. O processo provou trazer economia de tempo e custos durante todo o período em que foi usado pela equipe de desenvolvimento.

Futuramente é desejável que o sistema seja capaz de abranger outros tipos de arranjos mais complexos, aumentando a produção e qualidade de um SRV.

VI. AGRADECIMENTOS

Este trabalho tem custeio parcial da CAPES, do CNPq, da Fapemig e do projeto de P&D CEMIG GT411 – Aneel/Cemig em parceria com a UFU - convênio 4020000483, intitulado “Desenvolvimento de ambientes virtuais, para centro de operação de sistemas representativos das subestações e usinas da Cemig, associados a tecnologias de projeção 3D” a quem os autores agradecem.

REFERENCES

- [1] TORI, Romero; KIRNER, Claudio; SISCOUTO, Robson. *Fundamentos e Tecnologia de Realidade Virtual e Aumentada*. VIII Symposium on Virtual Reality, Belém-PA, 02 de Maio 2006.
- [2] ZHAO, Q. *A survey on virtual reality*. Science in China Series F: Information Sciences, 2009.
- [3] WALSH, A. E. *Understanding Scene Graphs*. Dr Dobb's Journal, 27:7, 17-26, 2002.
- [4] ROMANO, J. M. da Silva; ALBERTO B. Raposo; MARCELO Gattass Tecgraf. Depto. de Informática – PUC-Rio romano, abraposo, mgattass@tecgraf.puc-rio.br PUC-Rio.Inf.MCC11/04 Abril, 2004
- [5] FRONTIN, Sergio de Oliveira (Organizador). *Equipamentos de Alta tensão-Prospecção e hierarquização de inovações tecnológicas*. Brasília: Texeira, 2013.
- [6] MUZY, Gustavo L.. *Subestações Elétricas*. Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2012.
- [7] GUANGWEI, Y.; ZHITAO, G. *Scene graph organization and rendering in 3D substation simulation system*. In: Power and Energy Engineering Conference, 2009. APPEEC 2009. Asia-Pacific. [S.l.: s.n.], 2009. p. 1-4.
- [8] QUINTANA, J.; MENDOZA, E. *3D virtual models applied in power substation projects*. In: Intelligent System Applications to Power Systems, 2009. ISAP '09. 15th International Conference on. [S.l.: s.n.], 2009. p. 1-3.
- [9] MATTIOLI, Leandro; CARDOSO, Alexandre; LAMOUNIER, Edgard; SANTOS, Gesmar; PRADO, Paulo; FERREIRA, José Newton. *Automatização do processo de criação de Ambientes Virtuais para subestações de energia elétrica*. Encontro para Debates de Assuntos de Operação, novembro 2014.

Técnicas de Modelagem e Desenvolvimento para a Concepção de Usinas Hidrelétricas Virtuais

Modeling Techniques and Development for Virtual Hydroelectric Power Plants Design

Ígor Andrade Moraes*, Milton Miranda Neto, Alexandre Cardoso, Edgard Lamounier, Paulo Prado

Universidade Federal de Uberlândia

Grupo de Realidade Virtual e Aumentada - GRVA

Uberlândia – Minas Gerais

Email: {alex.cardoso.ufu, elamounier, voidmmn, igorandradesystem}@gmail.com, prmprado@uol.com.br

Abstract—The profits and benefits offered by Virtual Reality technology had drawn attention of professionals from several scientific fields, including the power systems', either for training or maintenance. For this purpose, 3D modeling is evidently pointed out as an imperative process for the conception of a Virtual Environment. Before the complexity of Hydroelectric Power Plants and Virtual Reality's contribution for the Industrial area, planning the tridimensional construction of the virtual environment becomes necessary. Thus, this paper presents modeling techniques applicable to several hydroelectric structures, aiming to optimize the 3D construction of the targets complexes.

Keywords—Virtual Reality; Hydroelectric; Power Plant; 3D Modeling; Power Systems

I. INTRODUÇÃO

No setor elétrico observa-se heterogeneidade entre os complexos responsáveis pela geração da Energia. Neste contexto, é relevante o desenvolvimento de soluções de software, notadamente, de Realidade Virtual que possam ser aplicadas a treinamento, manutenção, operação e controle de ativos presentes nas geradoras.

A realimentação positiva da aplicabilidade é constatada por soluções de Realidade Virtual nas mais diversas áreas científicas. Como exemplo, na Engenharia Elétrica observam-se diversas aplicações para o treinamento de profissionais e execução de procedimentos operacionais [1]. O uso de tal tecnologia também é fortemente justificado em contextos educacionais e de treinamento, assim como na visualização de informações.

Motivado por tal potencial, este trabalho apresenta o desenvolvimento de técnicas de modelagem capazes de otimizar o processo de geração de usinas hidrelétricas reproduzidas em ambientes virtuais e os resultados obtidos com tais técnicas.

II. PERFIL DA GERAÇÃO DE ENERGIA - SETOR ELÉTRICO

A geração de Energia no setor Elétrico é constituída por diferentes perfis de geradores: Hidrelétricos, Termoelétricos,

Termonucleares, Eólicos e Fotovoltaicos. Em média, segundo dados da ANEEL [2], 79% da Energia elétrica produzida no Brasil é gerada por Usinas Hidrelétricas. Estas geradoras de energia possuem instalações e características muito distintas, por serem dependentes do relevo onde foram instaladas. Este fato gera demandas específicas de treinamento, manutenção, operação e controle para cada uma delas.

No cenário atual, a geração é fortemente assistida por operação remota, distintamente do cenário de um passado próximo, quando um conjunto de operadores estava presente nas unidades locais. A operação remota deve levar em consideração as características individuais destas instalações, sendo necessário um treinamento local para que se estabeleça um modelo mental de cada arranjo elétrico existente. Combinando programas computacionais e computadores de alto desempenho, é possível a criação de um ambiente gráfico interativo capaz de promover experiências próximas à realidade [3], possibilitando que estes operadores não só operem as usinas remotamente, de forma adequada, como possam também ser treinados dispensando deslocamentos constantes para localidades da instalação de geradores.

A automação em diversos processos da geração de energia demanda novos sistemas de treinamentos que devem ser aplicados. No entanto, existe dificuldade de acesso a informações sobre as instalações e as novas técnicas de treinamento carecem no aspecto visual. Em resposta, a aplicação de tecnologias, que envolvam ambientes virtuais, no treinamento pode resolver os problemas mencionados anteriormente [4]. Evidentemente, a habilidade e o conhecimento do operador devem ser equivalentes à demanda da operação, mas para que o resultado esperado seja alcançado, a qualidade do treinamento também deve ser elevada. Ao mesmo tempo, a complexidade estrutural de uma usina hidroelétrica deve ser conhecida pelo operador para que uma tarefa seja realizada em tempo hábil [5].

III. TRABALHOS RELACIONADOS

Dado o impacto positivo causado pela modelagem tridimensional de sistemas complexos em diversas áreas, vários

são os trabalhos que fazem uso de objetos 3D em ambientes interativos no intuito de treinar e auxiliar profissionais. Em [6] a elaboração de modelos 3D auxiliou engenheiros no planejamento e na reconstrução da usina nuclear de Fukushima, a qual foi danificada pelo grande terremoto ocorrido em 11 de março de 2011. Apesar de apresentar uma obtenção de requisitos onerosa para uma reprodução fiel das estruturas, o trabalho ofereceu uma noção geográfica significativa para os envolvidos na reconstrução da usina e contribuiu com a recuperação da região afetada pelo terremoto.

Em [7], é apontada a baixa qualidade gráfica que simuladores voltados para treinamento em subestação apresentam e, como solução, é desenvolvida uma aplicação para treinamento flexível podendo ser aplicada em diversas subestações e usinas. [8] é um trabalho semelhante que atua em sistemas de energia e reconhece a superioridade do treinamento por meio de um ambiente virtual simulado, possibilitando interação e imersão, diante um ambiente de visualização convencional.

Finalmente, o trabalho [9] apresenta uma proposta de controle e gerenciamento de subestações de energia elétrica por meio de um ambiente virtual, podendo ainda ser utilizada no treinamento de novos funcionários. Por meio da imersão e de interações intuitivas, este correlato permite o controle de diversas subestações distribuídas pelo estado de Minas Gerais, refletindo o estado real dos equipamentos que compõem as subestações reais – Figura 1.



Figura 1 - VRCEMIG (2013)

IV. INSUMOS DE CONSTRUÇÃO DOS AMBIENTES VIRTUAIS DE USINAS HIDRELÉTRICAS

Como ponto de partida, visitas a usinas hidrelétricas foram promovidas no intuito de se obter requisitos de modelagem, os quais são armazenados na forma de fotos e vídeos. A partir dessas informações, adquiridas e arquivadas, padrões foram observados, o que possibilitou o desenvolvimento de técnicas aplicáveis a complexos hidrelétricos de forma mais ampla.

Com as visitas, foi observada a grande complexidade estrutural nas usinas. Assim, a modelagem de todos os ambientes é uma tarefa que apresenta inviabilidade, caso diversas usinas venham a ser modeladas. Apesar do presente trabalho apresentar técnicas desenvolvidas para serem aplicáveis a diversos espaços hidrelétricos, é de grande importância uma reprodução fiel do ambiente modelado, visto que cada complexo visitado apresenta características estruturais únicas.

Fora os ambientes visitados e fotografados, o ambiente virtual relacionado com uma dada Usina, deverá ser uma interface responsável pela seleção do ambiente a ser visitado e

prover condições da operação remota. Seus segmentos serão associados com ambientes que contêm informações topográficas do local que compreende a usina hidrelétrica, suas subestações. Para tal, também fundamenta concepção o conjunto de imagens de satélite da região e miniaturas representativas dos complexos apresentados no ambiente.

Observa-se, por demanda do processo de operação e controle, que três ambientes das Usinas são de maior frequência de funcionários e concentração de atividades e de elementos de controle: a Galeria Elétrica, a Galeria Mecânica e o Pátio de Transformadores.

A. Galeria Elétrica

Espaço que comprehende painéis de proteção e sincronização das máquinas geradoras. Estes painéis são, geralmente, dispostos de forma linear em torno de um corredor, adotando ordenamento e configuração diferente entre as diversas usinas visitadas – Figura 2.



Figura 2 - Galeria Elétrica da Usina Hidrelétrica de Amador Aguiar 1

B. Galeria Mecânica

Geralmente, é o espaço mais amplo da usina que abriga todas as máquinas geradoras, bem como o acesso para manutenção das mesmas. Também é um espaço que adota estruturas variadas entre as usinas visitadas, tanto no que se refere a disposição espacial quanto na configuração da máquinas geradoras.

C. Pátio de Transformadores

É uma área externa do complexo hidrelétrico que abriga os transformadores responsáveis pela elevação da tensão. Também existe um sistema contra incêndio que faz uso de água nebulizada voltado para os transformadores e para-raios que são distribuídos neste pátio.

V. ASPECTOS DA MODELAGEM E GERAÇÃO DE CENAS

Considerando a necessidade de se promover práticas que acelerem ou otimizem o desenvolvimento do trabalho de modelagem e concepção dos Ambientes Virtuais, foram adotadas convenções de modelagem e de Implementação. O software de modelagem tridimensional escolhido foi o 3DS Max, enquanto a Engine 3D para montagem do ambiente virtual, bem como as interações necessárias, é o Unity3D.

Em ambos os softwares, práticas de desenvolvimento foram documentadas e demandam fidelidade durante o trabalho para que o fluxo esperado de modelagem e montagem de cena seja alcançado. Um recurso notável da Engine escolhida para

montagem do ambiente é a possibilidade de se criar janelas customizáveis, um recurso denominado como *Editor Windows*, no qual funções e interações com a Engine podem ser definidas pelo usuário.

Assim, é possível adaptar o ambiente de desenvolvimento de forma a potencializar o trabalho que está sendo desenvolvido. Dos ambientes a serem modelados, a Galeria Elétrica e a Cena de Menu apresentaram maior complexidade de informações e semelhanças em suas construções tridimensionais. Por isso, técnicas de modelagem foram desenvolvidas para potencializar o processo de construção destes dois ambientes virtuais.

O ambiente que apresenta a interface responsável pelo direcionamento do usuário para os diversos ambientes dentro da usina hidrelétrica virtual, além de imagens de satélite capazes de abranger o espaço físico da usina e da subestação próxima, também concebe informações topográficas, as quais são obtidas por meio do Software Google SketchUp e diversos objetos tridimensionais. Assim, para a montagem da cena, uma ferramenta criada a partir do recurso *Editor Windows* da Engine 3D foi desenvolvida, denominada como “Interface Geradora de Mapa”, capaz de aprimorar o processo de produção do ambiente geográfico (Figura 3).

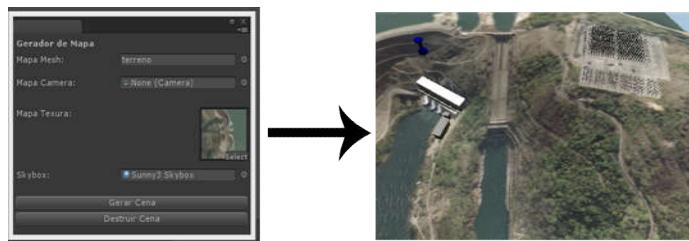


Figura 3 - Interface Geradora de Mapa

A interface apresentada demanda a seleção de quatro componentes para a geração da cena: a malha tridimensional do terreno, a câmera que deverá ser utilizada para a visualização do ambiente, a textura com imagens de satélite para a malha tridimensional e, por fim, o Material responsável pela representação visual do céu aberto.

Dessa forma, foi possível otimizar um processo de construção que poderá ser aplicado à diversas usinas hidrelétricas a serem reproduzidas virtualmente. Enfim, com a execução da interface desenvolvida, de forma geral, os seguintes processos serão executados automaticamente:

- Inserção e posicionamento do objeto que compreende o menu de seleção dos ambientes dentro da usina
- Posicionamento do terreno e câmera na cena.
- Aplicação da textura sobre o terreno
- Configuração da iluminação
- Aplicação da água sobre a área delimitada no mapa
- Inserção das partículas simuladoras das saídas de água.

Neste sentido, o recurso *Editor Windows* foi novamente utilizado, desta vez para a criação da Interface de Geração de

Painéis, no intuito de, novamente, aprimorar o processo de criação de cena. Esta última ferramenta é voltada para a Galeria Elétrica que, apesar de apresentar configurações estruturais distintas entre as diversas usinas, sempre dispõe de painéis de proteção distribuídos de forma enfileirada. Com a interface desenvolvida, é possível gerar, dentro da cena modelada, todo o conjunto de painéis já texturizados e distribuídos de acordo com a distância definida pelo usuário.

O conjunto de objetos gerados pela interface elaborada pode ser visualizado na Figura 4. Para tanto, um painel tridimensional foi modelado de modo que pudesse receber uma textura dotada de faces bem definidas, por meio de um processo de modelagem chamado *UVW Mapping*. Desta maneira, várias texturas, se configuradas da forma correta, podem ser perfeitamente aplicáveis ao mesmo painel. Para que esta etapa de produção funcione, cada textura deverá receber a imagem de um painel em uma área delimitada, criando assim um padrão aplicável em diversos cenários (Figura 5).



Figura 4 - Conjunto de Painéis Gerados pela Interface Desenvolvida

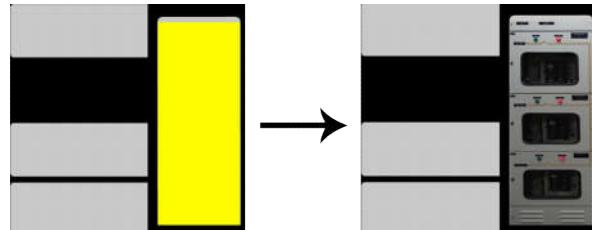


Figura 5 - Textura Preparada Para Aplicação

O processo de desenvolvimento não envolve somente convenções no processo de modelagem e montagem de cena. A organização do diretório que comprehende todos os componentes necessários para a criação do ambiente virtual também é de extrema importância para que partes do processo sejam automatizadas. Em ambas as interfaces desenvolvidas para auxiliar na criação do cenário, existem funções que buscam objetos reaproveitáveis em diretórios específicos. Como por exemplo, no cenário que apresenta o mapa onde se encontram a usina hidrelétrica e a subestação 3D, objetos tridimensionais e componentes devem estar no diretório correto de busca. As mesmas convenções se aplicam para a geração otimizada dos painéis presentes na galeria mecânica, visto que as texturas, corretamente nomeadas, também devem estar nos seus respectivos diretórios.

A partir da análise do material obtido por meio das visitas às usinas e dos processos de construção dos ambientes virtuais, uma estratégia inicial de desenvolvimento dos ambientes foi

elaborada, que corresponde ao diagrama apresentado na Figura 6. Com a obtenção de insumos, inicia-se desenvolvimento das cenas virtuais, bem como o armazenamento dos modelos desenvolvidos em uma biblioteca. Terminada a cena tridimensional, testes de validação devem ser realizados para garantir a fidelidade da representação com o ambiente real, assim como as interações básicas do usuário com o sistema. Para a entrega da cena produzida, uma integração com o sistema de operação remota deve ser realizada para que, desta forma, a aplicação seja capaz de apresentar informações reais sobre o estado das máquinas presentes nas usinas.

AGRADECIMENTOS

Este trabalho tem custeio parcial da CAPES, do CNPq, da Fapemig e do projeto de P&D CEMIG GT411 – Aneel/Cemig em parceria com a UFU - convênio 4020000483, intitulado “Desenvolvimento de ambientes virtuais, para centro de operação de sistemas representativos das subestações e usinas da Cemig, associados a tecnologias de projeção 3D” a quem os autores agradecem.

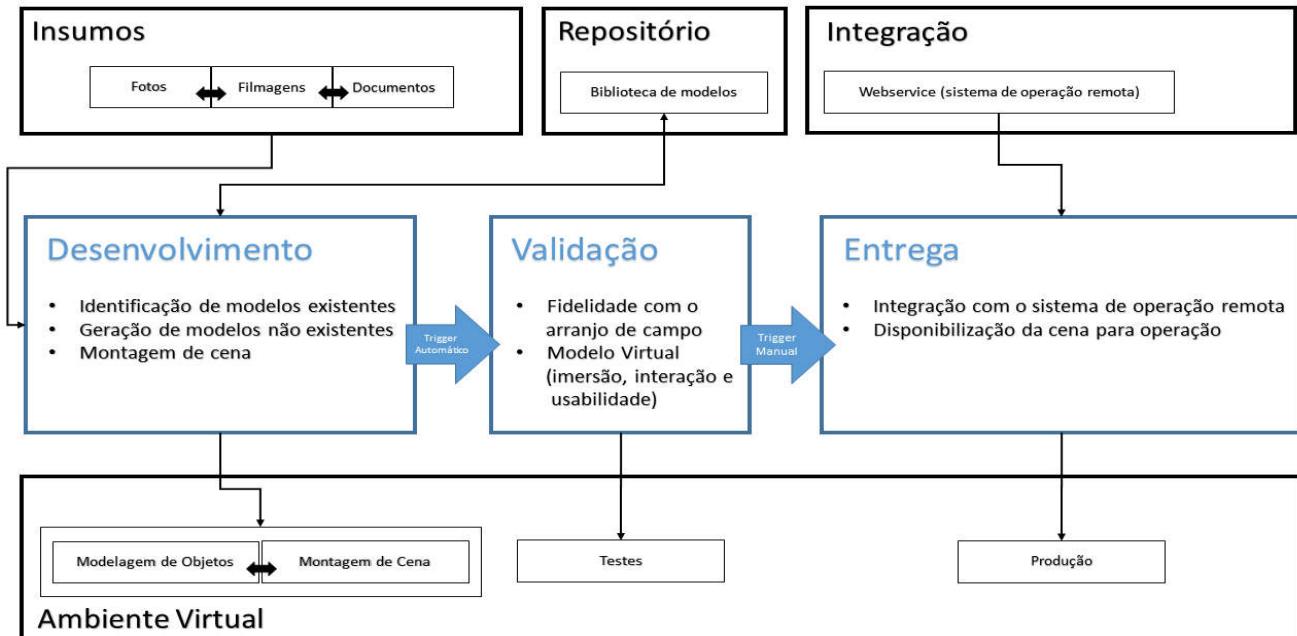


Figura 6 - Diagrama de Desenvolvimento das Usinas Virtuais.

VI. CONCLUSÕES E CONSIDERAÇÕES FINAIS

A diversidade operacional encontrada no setor elétrico nacional possibilita a aplicação das técnicas propostas neste trabalho, o qual deverá ser capaz de promover processos de treinamento mais eficientes e até mesmo a simplificação dos processos atuais de operação dos sistemas das Usinas Hidrelétricas.

A aplicação de treinamento nos ambientes virtuais criados permite a desoneração e dispensa a necessidade dos operadores se deslocarem até o local das unidades geradoras frequentemente, tornando possível a aquisição de conhecimento e experiência através do modelo mental adquirido por meio das cenas virtuais de todas as unidades criadas neste ambiente.

Por meio de entrevistas com funcionários, foi possível confirmar que as cenas, geradas por meio das técnicas de modelagem apresentadas neste trabalho, apresentaram qualidade gráfica alta, sendo que alguns dos ambientes desenvolvidos foram rapidamente reconhecidos por operadores que já haviam visitado o espaço real representado no sistema.

REFERÊNCIAS

- [1] GUANGWEI, Y.; ZHITAO, G. Scene graph organization and rendering in 3D substation simulation system. In: Power and Energy Engineering Conference, 2009. APPEEC 2009. Asia-Pacific. [S.I.: s.n], 2009. P 1-4.
- [2] AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA – ANEEL. Banco de Informações de Geração – BIG. 2003. Disponível em: <http://www.aneel.gov.br/aplicacoes/atlas/aspectos_institucionais/2_2.htm>
- [3] CARDOSO, Alexandre; LAMOUNIER, Edgard. A Realidade Virtual na Educação e Treinamento. In: TORI, Romero; KIRNER, Claudio; SISCOUTO, Robson. Fundamentos e Tecnologia de Realidade Virtual e Aumentada. Belém: SBC, 2006. p. 304-312.
- [4] HONGTAO, Zeng. A Remote Training System of Hydropower Plant Based on Interactive Virtual Environments. Sch. of Power & Mech. Eng. Wuhan, China, 2010.
- [5] GUO, Jiang. Virtual environment conception for CBM of hydro-electric generating units. Fac. of Hydropower & Digitalization Eng. Wuhan, China, 2002.
- [6] DOI. A. 3D topographic map generation of Fukushima Daiichi Power plant: Visualization of the reconstruction plan for effective information sharing. Iwate Prefectural Univ. Takizawa, Japan, 2014.
- [7] WANG, Wenju. Virtual reality in the substation training simulator. College of Electronics and Information Engineering. Shanghai, 2010.
- [8] GUO, Xiao Li. Application of the virtual reality technologies in power systems. Sch. of Inf. Eng., Northeast Dianli Univ., Jilin, China, 2010.
- [9] CARDOSO, Alexandre; LAMOUNIER, Edgard; LIMA, Gerson. et al. VRCEMIG: A Virtual Reality System for Real Time Control of Electric Substations. IEEE, Brazil, 2013.

Visualização Científica de Descargas Elétricas Atmosféricas

Scientific visualization of atmospheric electrical discharges

José Remo Ferreira Brega

Márcia A. Zanolí Meira e Silva

Gustavo Azarito Silva

Departamento de Computação
Universidade Estadual Paulista
(UNESP)

Bauru, SP, Brasil

remo@fc.unesp.br

marciazm@fc.unesp.br

gustavoazarito@hotmail.com

Mário Popolin Neto

Instituto Federal de São Paulo
(IFSP)

São Paulo, SP, Brasil

mariopopolin@gmail.com

Robson Siscoutto

Universidade do Oeste Paulista
(Unoeste)

Presidente Prudente, SP, Brasil

robson.siscoutto@unoeste.br

Resumo - Este trabalho tem como principal objetivo desenvolver uma aplicação web para visualização de Descargas Elétricas Atmosféricas, plotando marcas em um mapa, utilizando como base dados fornecidos pelo SIMEPAR. O sistema desenvolvido utiliza técnicas e ferramentas que possibilitam a criação do Sistema web sem que seja necessária a utilização de plug-ins ou softwares externos, funcionando em praticamente todos navegadores de internet, bem como várias plataformas. Esta ferramenta visa auxiliar os meteorologistas a monitorar as descargas elétricas de forma ágil e ajudar a sociedade em geral na prevenção de acidentes causados por esse fenômeno da natureza.

Palavras-chave: *descargas elétricas atmosféricas; visualização científica; meteorologia; sistema de monitoramento; web.*

Abstract - This paper aims primarily to develop a web application for Atmospheric Electrical Discharge visualization by plotting marks on a map based on the database provided by SIMEPAR. The developed system uses tools and techniques that allow the creation of the Web System in a way that plug-ins and external software aren't necessary, making it run virtually in all web browsers, as well as in several platforms. This tool was created in order to assist meteorologists to quickly monitor electrical discharges and consequently help the society in general in the prevention of accidents caused by this natural phenomenon.

Keywords: *atmospheric electrical discharges, scientific visualization, meteorology monitoring systems, web.*

I. INTRODUÇÃO

Desastres naturais são os mais variados eventos capazes de gerar grande impacto na sociedade e se distinguem de acordo com a natureza do fenômeno que os desencadeiam. Como exemplos destes desastres naturais, responsáveis por intensos acidentes, podemos citar tsunamis, furacões, tempestades e descargas elétricas.

Descargas Elétricas Atmosféricas (DEA), conhecidas como raios, são fenômenos que ocorrem pelo acúmulo de cargas elétricas em regiões da atmosfera. Quando o campo elétrico produzido pelas cargas excede a capacidade isolante do ar, em um local da atmosfera, dá-se inicio à descarga. Estima-se que, por ano, cerca de 100 milhões de Descargas Elétricas Atmosféricas da nuvem para o solo (Nuvem-Solo) atingam o Brasil, ou seja, três Descargas Elétricas Atmosféricas por segundo [1].

Por meio de monitoramentos de dados de Institutos Meteorológicos, é possível acompanhar as ocorrências de DEA em tempo real e, consequentemente, evitar alguns desses desastres.

Com esse objetivo, o acompanhamento de ocorrências de DEA em tempo real utilizando um sistema de monitoramento deste fenômeno que seja intuitivo aos seus usuários e que funcione de computadores simples aos mais robustos é de grande interesse. Além disso, de acordo com [2], as informações obtidas de descargas atmosféricas podem gerar produtos para a previsão do tempo, análise e

manutenção de sistemas elétricos, emissão de laudos de análise de eventos severos para seguradoras e empresas de engenharia em geral, entre tantas aplicações.

Uma ferramenta que pode complementar a eficiência desse monitoramento é uma aplicação *web*, definida como uma aplicação de software que utiliza a *web* como ambiente de execução. Uma das vantagens de uma aplicação *web* é não ter a necessidade de instalação prévia de bibliotecas e softwares para o funcionamento de tais aplicações permitindo o seu uso em uma grande variedade de PCs (Personal Computer), satisfazendo o interesse mencionado anteriormente.

Neste contexto, esta pesquisa desenvolveu uma aplicação que facilita a visualização de dados de Descargas Elétricas Atmosféricas (DEA) em ambiente *web*, no caso um mapa, utilizando dados fornecidos pelo SIMEPAR (Sistema Meteorológico do Paraná), tendo como meta melhorar a prevenção de acidentes causados por estes fenômenos no Brasil.

II. OBJETIVOS

O objetivo geral desta pesquisa foi desenvolver uma aplicação *web* para visualização de DEA, em um mapa, utilizando dados fornecidos pelo SIMEPAR, visando à identificação de regiões com maiores e menores incidências de raios a fim de alertar e prevenir acidentes causados por tais descargas.

Foram objetivos específicos:

- Investigar as técnicas de visualização científica;
- Entender a captação dos dados brutos, por meio da cooperação com o SIMEPAR, através dos sensores de detecção de DEA da rede RINDAT (Rede Integrada Nacional de Detecção de Descargas Atmosféricas);
- Analisar e organizar os dados recebidos de maneira com que possam ser representados por meio da *web*;
- Criar representações da localização de DEA em um mapa.

III. MATERIAIS E MÉTODOS

Este trabalho se desenvolveu de acordo com as seguintes etapas:

- Estudo das DEA e entendimento do método de captação de suas ocorrências pelos sensores da RINDAT ([1], [2], [3] e [4]);
- Solicitação dos dados ao SIMEPAR;
- Estudo das técnicas de visualização científicas de DEA ([5], [6] e [7]);
- Estudo da API *Google Maps V3* ([8]);

- Escolha da melhor técnica para implementação ([9], [10], [11] e [12]).
- Configuração *EasyPHP* ([9] e [13]);
- Desenvolvimento da aplicação *Web* ([13] e [15]).

Inicialmente realizou-se um estudo sobre as DEA destacando-se: suas definições; os tipos existentes (com ênfase nas DEA Nuvem-Solo, por serem as com maior poder de destruição); as DEA no Brasil e a sua detecção.

Na sequencia buscou-se o entendimento da captação dos dados brutos das DEA pelos sensores LPATS (*Lightning Position and Tracking System*) e IMPACT (*IMProved Accuracy from Combined Technology*) da RINDAT. A diferença entre esses sensores está relacionada à forma como a localização das descargas de raios é processada e a quantidade mínima de sensores para identificação da descarga atmosférica.

Em seguida foram solicitadas ao SIMEPAR amostras de dados, já processados pelas centrais de processamento e preparados para visualização e/ou armazenamento. O arquivo de dados fornecido pelo SIMEPAR estava em formato CSV e continha informações como localização, polaridade, intensidade da corrente, data, hora e coordenadas geográficas das descargas. Nesta etapa foi necessário o estudo de dois Sistemas de Coordenadas Geográficas, sendo o Sistema de Coordenadas UTM (*Universal Transversa de Mercator*) e o Sistema de Coordenadas Latitude e Longitude, que descrevem com absoluta exatidão qualquer localização na Terra [14].

Após receber os dados, fez-se uma pesquisa e estudo sobre técnicas de visualização desses dados em mapas da *web*. O processo de visualização de dados, basicamente, consiste em três etapas que, de acordo com [6] são: aquisição dos dados, transformação em forma apropriada para representação e renderização com a representação em monitor ou outro *display* (superfície de visualização).

A opção de desenvolver em plataforma *web* deve-se a não necessidade de instalações prévias de bibliotecas e softwares adicionais para se obter acesso à aplicação em diferentes plataformas, como *Windows*, *Linux*, *Android*, etc. Basta ao usuário ter conexão com internet e um *browser*. As páginas *web* são criadas de forma dinâmicas através de formulários *HTML* (*Hypertext Markup Language*), cuja principal função é processar as solicitações efetuadas pelo cliente, por meio do *browser*, e exibir os resultados de acordo com o código contido no servidor ([15]). Como exemplo de servidor *web*, pode ser citado o *Apache*. O *Apache HTTP Server* é um servidor *web* livre, líder de mercado de 1995 até os dias atuais, responsável por alimentar mais de 65% dos sites no mundo, de acordo com dados da empresa ([13]).

A tecnologia escolhida para realizar a representação dos dados nas páginas da *web* foi a API *Google Maps V3*. Uma API é um conjunto de padrões, funções, classes e métodos

para se utilizar em um software sem precisar conhecer detalhadamente a sua implementação, bastando apenas saber como utilizar os serviços ([16]). A *API* do *Google Maps* é um serviço gratuito, disponível para qualquer website. Nela está definido um conjunto de objetos e métodos *JavaScript* que permitem incorporar o *Google Maps* em páginas da web. Por meio da documentação da *API* fornecida pelo próprio desenvolvedor *Google*, foi possível estudar e entender grande parte da funcionalidade dessa *API*.

Conceitos de aplicações web e linguagens de programação também foram estudados. Destacou-se a linguagem *PHP* (acrônimo recursivo para *PHP: Hypertext Preprocessor*) por ser uma linguagem de programação livre, interpretada e de grande utilização em desenvolvimento para a web, além de atuar predominantemente no lado do servidor e combinar-se ao código *HTML*. Ela auxilia os desenvolvedores a escreverem páginas que serão geradas dinamicamente de forma rápida.

Outro destaque foi a linguagem *JavaScript*. É uma linguagem leve, interpretada e se incorporada a um código *HTML*, permite operar e manipular os objetos Modelo de Objetos de Documento (DOM). Atualmente é a principal linguagem *client-side*, ou seja, executada na máquina do usuário, utilizada em navegadores web ([10]).

Também se estudou o *CSS3 (Cascading Style Sheets)* utilizado para definir a forma como os objetos de um documento escrito em linguagem de marcação (*HTML*) serão apresentados. Seu principal benefício é separar o arquivo de formatação do documento, criando um *link* para uma página que contém os estilos. Uma de suas vantagens está relacionada ao seu uso em projetos com muitas páginas, pois neste caso, ao se desejar alterar a aparência basta modificar apenas um arquivo ([11]).

A seguir, instalou-se e configurou-se o *EasyPHPDEVServer 13.1 VC11* para a utilização do servidor *Apache* e linguagem *PHP* para, então desenvolver a aplicação web como um todo.

IV. RESULTADOS E DISCUSSÃO

O software desenvolvido é uma aplicação web clássica, com interface amigável e intuitiva, para visualização e monitoramento de descargas elétricas atmosféricas. Foi desenvolvido em parceria com o Instituto Tecnológico SIMEPAR, responsável por fornecer a base de dados, em arquivo de formato CSV, cujos parâmetros são Data, Hora, Coordenadas Geográficas, no formato UTM, e Geodésicas (Latitude e Longitude) e Valor da Corrente, em quilo ampere (kA).

Esta aplicação oferece recursos para monitoramento e previsões de DEA, auxiliando companhias de eletricidade e telecomunicações na análise de falhas, bem como prevenindo acidentes à população além de auxiliar especialistas em meteorologia a visualizarem, de forma

simples, as descargas captadas pelos sensores meteorológicos.

O sistema funciona, basicamente, plotando marcas em formato de raio de diferentes cores no *Google Maps*, de acordo com sua posição geográfica informada no arquivo CSV. As marcas de mesma cor representam as descargas que ocorreram exatamente no mesmo momento (hora, minuto e segundo), fazendo com que a identificação e visualização das descargas pelo usuário ocorram mais rapidamente.

A Fig. 1 ilustra a aplicação desenvolvida.

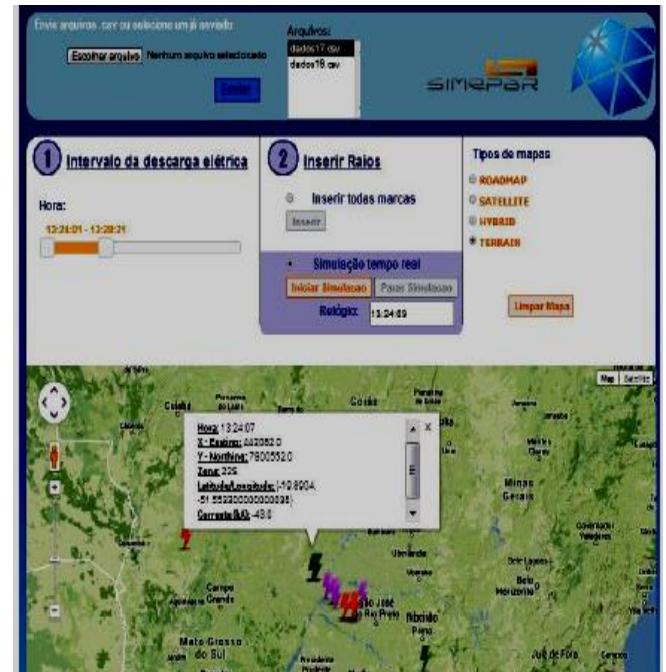


Fig. 1 – Aplicação Web desenvolvida.

Fonte: Foto da aplicação.

Inicialmente foi escolhida a tecnologia *WebGL* que é inovadora, muito utilizada em games e permite a utilização de gráficos 3D em navegadores de forma nativa. Porém encontrou-se algumas dificuldades com o servidor de mapas para essa tecnologia, tais como não possuir grandes recursos e não apresentar os nomes de alguns estados e cidades do Brasil, além de ser encontrado apenas em dimensão 3D. Este tipo de dimensão exige forte processamento da placa de vídeo acarretando lentidão no servidor de mapas e dificultando a execução da aplicação em dispositivos móveis e também em PCs de menor porte. Posteriormente percebeu-se a não necessidade de um mapa 3D para visualização das descargas e uma nova busca foi realizada encontrando-se a *API do Google V3* para incorporar o *Google Maps* ao browser. Em testes básicos verificou-se que essa *API* se comportava de forma rápida em qualquer dispositivo móvel ou PC e o mapa, mesmo sendo em 2D, atendia as necessidades do sistema em desenvolvimento. Por exemplo, o *Google Maps* possui todas as demarcações de cidades, estados e países; é possível efetuar zoom até que se encontrem as ruas das cidades e o usuário ainda pode

escolher entre quatro opções de tipos de mapas: *Roadmap*, *Satellite*, *Hybrid* e *Terrain*, sem que ocorra sobrecarga do sistema ou grande processamento da placa de vídeo do usuário.

Para a aplicação em questão, utilizou-se um servidor *Apache* instalado junto com o pacote *EasyPHPDEVServer 13.1 VC11* que realiza a comunicação cliente-servidor através do protocolo *HTTP*, e para retorno das informações processadas no servidor, beneficiou-se da linguagem *PHP*, que retorna ao cliente essas informações, a fim de se concretizar a visualização dos dados no *browser*.

No lado do cliente (*client-side*) aplicaram-se *scripts* escritos em *JavaScript* para implementação da *API Google V3* e para alguns seletores específicos utilizou-se de biblioteca *jQuery* e a estilização da página ficou por conta da linguagem *CSS3*.

Para utilizar o sistema o usuário deve carregar o arquivo CSV com os dados das DEA para o servidor através de um formulário *HTML* ilustrado na Fig. 2.



Fig. 2 – Envio do(s) arquivo(s) CSV para o servidor através de formulário *HTML*.

Fonte: Foto da aplicação.

Quando o botão “Enviar” é pressionado, o arquivo é enviado para o servidor e os arquivos enviados aparecem listados no *ListBox* “Arquivos”, Fig. 3.



Fig. 3 – Arquivo carregado no servidor e listado no *ListBox* “Arquivos”.

Fonte: Foto da aplicação.

O usuário, sempre que acessar a página, encontrará os arquivos que já foram enviados listados nesse *ListBox* e deverá então selecionar o arquivo desejado clicando sobre o nome dele. Nesse momento, abre-se um *frame*, abaixo dessa seção de envio de arquivo, contendo um menu com uma barra seletora para que selecione o intervalo de horário da descarga. Ao lado direito há um botão para inserir todas as marcas de uma só vez ou, logo abaixo deste, um botão para iniciar e parar a simulação em tempo real e, mais a direita, botões para seleção do tipo de mapa desejado e para limpar as marcas que já foram inseridas no mapa. Logo abaixo desse menu encontra-se o *Google Maps*, centralizado no Brasil, e ainda sem as marcações de raios (Fig. 4).

O seletor de intervalo das descargas elétricas (indicado com o número 1 na Fig. 4) carrega os horários de ocorrência das DEA de acordo com o arquivo CSV, tendo como limite inferior a primeira descarga do arquivo e limite superior a última. Há a possibilidade de selecionar o intervalo desejado

para que o sistema insira no mapa as marcações correspondentes a este intervalo selecionado.

Selecionado o intervalo de ocorrências das DEA, deve-se então escolher se deseja inserir todas as marcas, correspondentes a esse intervalo, de uma só vez ou se deseja simular que elas sejam inseridas no mapa como se estivessem sendo plotadas à medida que as descargas foram captadas pelos sensores.

As marcas inseridas possuem as informações de Data, Hora, Coordenadas Geográficas e Valor da Corrente e ao se clicar nelas abre-se uma caixa contendo essas informações (Fig. 5).

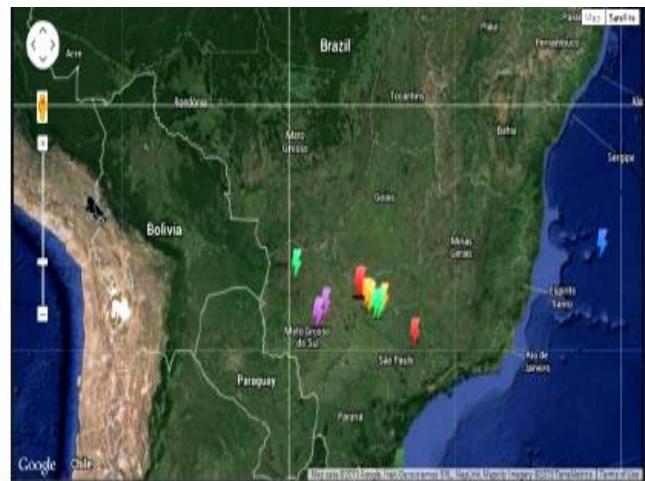
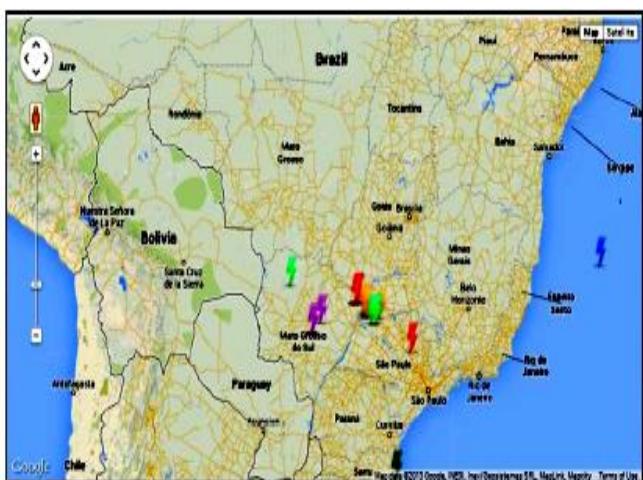


Fig.4 – Frame carregado ao selecionar um arquivo da lista.

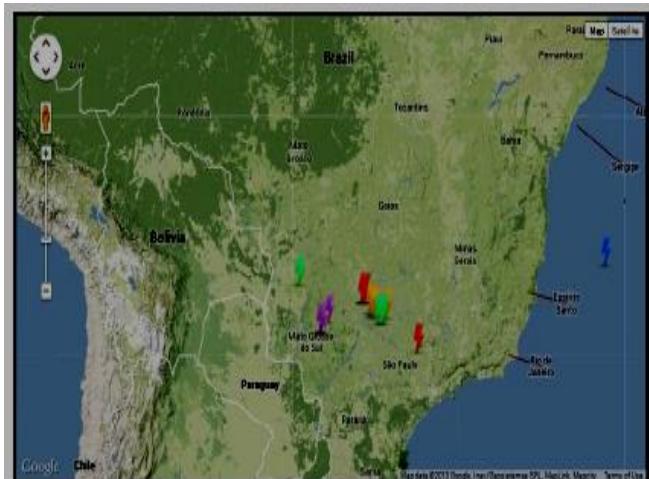
Fonte: Foto da aplicação



Finalmente, tem-se a opção de mudar o estilo do mapa para Roadmap (Fig. 6), Satellite (Fig. 7), Hybrid (Fig. 8), Terrain (Fig. 9).



Finalmente, tem-se a opção de mudar o estilo do mapa para Roadmap (Fig. 6), Satellite (Fig. 7), Hybrid (Fig. 8), Terrain (Fig. 9).



V. CONSIDERAÇÕES FINAIS

Esta pesquisa apresenta uma solução para visualização de dados sobre Descargas Elétricas Atmosféricas em ambiente *web* para monitoramento, alerta e possível prevenção de acidentes provocados por raios.

Os dados meteorológicos fornecidos pelo SIMEPAR são sigilosos e de fim estritamente científicos. Visando a preservação dos direitos de propriedade intelectual e industrial (Lei nº 9.279, de 14 de maio de 1996 que regula direitos e obrigações relativos à propriedade industrial, com as alterações introduzidas pela Lei nº 10.196, de 14 de fevereiro de 2001), foi disponibilizado apenas um arquivo de dados contendo 1.742 informações de descargas elétricas diferentes.

Por essa razão não foi possível desenvolver um sistema que plote os raios no mapa em tempo real, mas foi inserida uma simulação de tempo real de acordo com os horários das ocorrências dos raios contidos nos dados obtidos.

A simulação em tempo real permite que futuramente possa se estabelecer uma conexão em tempo real com o instituto e, consequentemente com suas centrais de processamento dos sensores meteorológicos, para que o sistema receba e plote no *Google Maps* os raios em tempo real, sem que sejam necessárias grandes adaptações no código já escrito. Tal comunicação poderá ser realizada através de um *Web Service*, que é uma forma de integrar sistemas desenvolvidos em plataforma diferentes, permitindo-se então comunicar a aplicação com as centrais de processamento, convertendo-se os dados em formato XML (*eXtensible Markup Language*) e transferindo-os por um protocolo SOAP (*Simple Object Access Protocol*), por exemplo.

Além disso, como continuidade deste trabalho pretende-se através do sensor de movimentos Kinect da Microsoft promover maior interação do usuário com as imagens plotadas. Será avaliada a viabilidade dessa iteração ser realizada com o ambiente gerado a partir das imagens de ocorrências de DEA apresentadas no *Google Maps* (2D) ou através do uso de mapas 3D.

- [1] LIMA, K. C. **Descargas elétricas atmosféricas em sistemas convectivos de mesoscala no sul da América do Sul**. Dissertação (Pós-Graduação em Meteorologia). Universidade Federal de Pelotas, Rio Grande do Sul, 2005.
- [2] BENETI, C.; VASCONCELLOS, C. Sisraios – um sistema interativo de análise e visualização de descargas elétricas atmosféricas. In **XII Congresso Brasileiro de Meteorologia. Foz do Iguaçu, 2002**.
- [3] INPE. **Descargas Nuvem-Solo**. Disponível em: <http://www.inpe.br/webelat/rindat/menu/desc.atm>. Acesso em: 26 out. 2013.
- [4] RINDAT. **Rede Integrada Nacional de Detecção de Descargas Atmosféricas**. Disponível em <http://www.rindat.com.br>. Acesso em: 05 jun. 2013.
- [5] CARD, S. K.; MACKINLAY, J. D. **Readings in Information Visualization: Using Vision to Think**. USA: Academic Press, 1999.

- [6] NETO, M. A.; BURIOL, T. M.; SCHEER, S. Técnicas de visualização científica aplicada à exploração de dados obtidos pelo método dos elementos finitos. In: **XXVIII Congresso Ibero Latino Americano de Métodos Computacionais em Engenharia, 2007, Porto. CMNE/CILAMCE 2007**. Porto: Faculdade de Engenharia - Universidade do Porto, p. 13–15, 2007.
- [7] BATTAIOLA, A. L.; SOARES, L. P. Estudo e uso exploratório de ferramentas de visualização científica. **Anais da VII Semana de Informática da UFBA**, p. 4, 1998.
- [8] GOOGLE. **V3: a solução para aplicativos do Google Maps para dispositivos desktop e móveis**. Disponível em: <https://developers.google.com/maps/documentation/javascript?hl=pt-BR&csw=1>. Acesso em 05 nov. 2013.
- [9] PHP. **Manual do PHP**. Disponível em: http://www.php.net/manual/pt_BR/preface.php. Acesso em: 05 jun. 2013.
- [10] FLANAGAN D. **Javascript: O GuiaDefinitivo**. 6 ed, Porto Alegre: Bookman, 2013.
- [11] JOBSTRAIBIZER, F. **Criação de sites com o CSS**. São Paulo: Digerati Books, 2009.
- [12] W3C. **Cascading Style Sheets (CSS) Snapshot 2010**. Disponível em: <http://www.w3.org/TR/CSS>. Acesso em: 05 nov. 2013.
- [13] APACHE. **Howthe ASF Works**. Disponível em: <http://www.apache.org/foundation/how-it-works.html#history>. Acesso em 17 mai. 2013.
- [14] FERREIRA, S. **Geografia – Ensino Fundamental – EJA: Coordenadas Geográficas**. Disponível em: <http://isoldaeja.blogspot.com.br>. Acesso em: 11 nov. 2013.
- [15] GLASS M. K. et al. **Beginning PHP, Apache, MySQL Web Development**. USA: Wiley Publishing, 2004.
- [16] ORENSTEIN, D. **QuickStudy: Application Programming Interface (API)**. Disponível em: http://www.computerworld.com/s/article/43487/Application_Programming_Interface. Acesso em: 10 out. 2013.

Desenvolvimento de Aplicações para o Oculus Rift: *Integração do Oculus Rift com o Street View*

Fernanda Mendes

Faculdade de Computação e Informática
Universidade Presbiteriana Mackenzie
São Paulo, SP, Brasil
nandayoko@hotmail.com

Valéria Farinazzo Martins

Faculdade de Computação e Informática
Universidade Presbiteriana Mackenzie
São Paulo, SP, Brasil
valfarinazzo@hotmail.com

André Santos Lima

Faculdade de Computação e Informática
Universidade Presbiteriana Mackenzie
São Paulo, SP, Brasil
andresantoslima@gmail.com

Marcelo de Paiva Guimarães

Universidade Federal de São Paulo (UAB/Unifesp)/
Programa de Mestrado da Faccamp, São Paulo, SP,
Brasil
marcelodepaiva@gmail.com

Resumo— Realidade Virtual e Realidade Aumentada trazem, entre outras características, a sensação do usuário de estar imerso num ambiente virtual. O surgimento do Oculus Rift trouxe a possibilidade de seu uso em sistemas computacionais compostos por imagens panorâmicas, tornando estes sistemas imersivos. Neste sentido, este trabalho apresenta a integração do Oculus Rift ao Google Street View, permitindo que o usuário consiga visualizar e navegar, através de comandos de voz, por localizações no mundo, tornando-o imerso neste ambiente. Neste trabalho, são apresentadas todas as fases de desenvolvimento desta aplicação, assim como os testes realizados com potenciais usuários.

Palavras-chave—Oculus Rift; Google Street View; Imersão; Interface natural

Abstract— Virtual Reality and Augmented Reality bring, among other features, the user's feeling of being immersed in a virtual environment. The emergence of the Oculus Rift brought the possibility of their use in computer systems composed of panoramic images, making these immersive systems. Thus, this paper presents the integration of the Oculus Rift to Google Street View, allowing the user can view and navigate through voice interface, for locations on the world, making it immersed in this environment. In this work are presented all development stages of this application, as well as tests carried out with potential users.

Keywords—Oculus Rift; Google Street View; Immersion; natural interface

I. INTRODUÇÃO

Verifica-se, no cenário atual, o surgimento de novas tecnologias destinadas a tornar a experiência do usuário com a interação mais natural. Isso pode ser verificado com o desenvolvimento e aprimoramento de tecnologias como Realidade Virtual (RV) [1], Realidade Aumentada (RA) [2], Interfaces por Voz [3] e Interfaces Gestuais [4], que possibilitam ao usuário interagir com o computador de forma semelhante ao mundo real, usando a voz, o tato, etc. [5].

Uma das características inerentes à Realidade Virtual é seu caráter imersivo [6]. Para prover a imersão, pode-se utilizar um capacete de visualização, como o Oculus Rift, que utiliza tecnologia de rastreamento de movimentos da cabeça em 360°, possibilitando que o usuário possa olhar ao seu redor de maneira semelhante ao qual faria na vida real, proporcionando uma experiência natural [7].

Vários trabalhos têm abordado o desenvolvimento de aplicações de RV imersiva, fazendo uso do dispositivo Oculus Rift como meio de interação. Reiners e seu grupo [8] utilizam o Oculus Rift em ambientes imersivos com cenários de indústrias e infraestrutura de logística com o objetivo de melhorar e incentivar o aprendizado do colaborador sobre sua saúde ocupacional e sua segurança. Os experimentos apresentam missões e desafios relevantes que devem ser realizados pelo funcionário, com foco no desenvolvimento da capacidade do usuário em identificar situações de alto risco e agir em conformidade ao risco apresentado.

Já, Hoffman e seu grupo [9], fornecem um tratamento de queimaduras com utilização de ambiente imersivo por meio do dispositivo Oculus Rift com a finalidade de diminuir a dor do paciente. Uma vez que os métodos convencionais (remédios e terapias) não funcionam nos procedimentos de limpeza das feridas, nos exercícios de alongamento dos membros queimados e no procedimento de esticar a pele, gerando dores insuportáveis ao paciente. Nesse contexto, intervenções psicológicas podem facilitar os efeitos de analgésicos farmacológicos padrão. Para tanto, o paciente utiliza o Oculus Rift para navegar num mundo confeccionado com gelo projetado enquanto o paciente passa pelos procedimentos médicos dolorosos necessários. Como o cérebro do paciente está ocupado em interagir com o ambiente de RV, o paciente se esquece da dor. Um teste foi realizado com um menino de 11 anos de idade que possuía várias queimaduras elétricas e de flash em sua cabeça, ombros, braços e pés. Ele passou por sessão de 20 minutos sem VR no primeiro dia, uma sessão com VR no segundo dia, e uma sessão final sem VR no terceiro dia.

No dia em que fez uso de RV como tratamento, a sensação da dor caiu de desagradável para moderadamente desagradável.

O objetivo geral dessa pesquisa é desenvolver uma aplicação que integre as possibilidades de movimento nas imagens panorâmicas com os comandos de direções do dispositivo Oculus Rift, obedecendo a comandos de voz do usuário para determinação da área a ser visitada. Como objetivo específico, tem-se a possibilidade de navegação no aplicativo Google Street View usando o dispositivo mencionado, para, por meio do rastreamento dos movimentos da cabeça, indicar as direções a serem tomadas, descartando, assim, mouse e teclado. A intenção é proporcionar ao usuário a sensação verdadeira de estar presente no local onde os olhos podem visualizar, sem que haja de fato a necessidade de deslocamento físico [10].

Este artigo está estruturado como segue. A seção II apresenta a metodologia do trabalho. A seção III versa sobre as fases de desenvolvimento da aplicação. Na seção IV são apresentados os resultados obtidos através dos testes com usuários. Finalmente, na seção V estão as considerações finais deste trabalho.

II. METODOLOGIA

Para alcançar os objetivos gerais e específicos deste trabalho foram executados os seguintes passos:

- Estudo do funcionamento e de desenvolvimento utilizando o equipamento;
- Configuração, instalação e execução do dispositivo Oculus Rift;
- Desenvolvimento propriamente dito da aplicação;
- Aplicação dos testes com o público-alvo; e
- Análise dos resultados obtidos na etapa de avaliação.

III. DESENVOLVIMENTO DA APLICAÇÃO

A aplicação desenvolvida utiliza o Street View. Nela o usuário visualiza o local onde irá trafegar e interagir com comandos de voz. Esses comandos estão associados a ações no ambiente, como, por exemplo, ir em frente.

Assim, o usuário pode colocar o Oculus Rift e visita qualquer lugar do mundo que deseja sem sair de casa. A aplicação oferece a oportunidade de visualizar paisagens fotografadas pela Google juntamente com a sensação de presença física no falso ambiente real que é apresentado.

A base para o desenvolvimento da pesquisa se deu por meio de análises de programas já existentes e exemplos que fazem uso do Oculus Rift e usam código aberto, como o Oculus Street View, por exemplo, possibilitando assim, um estudo específico e detalhado.

O desenvolvimento da aplicação foi realizado utilizando prototipação. Assim, as fases do desenvolvimento foram iterativas, e a cada ciclo de desenvolvimento, mais funcionalidades foram sendo implementadas. Estas fases

podem ser visualizadas na Figura 1 e explicadas nas próximas subseções.

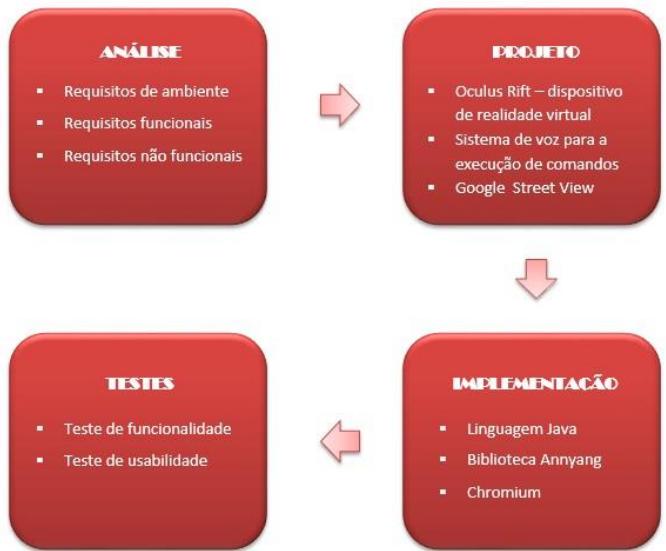


Fig. 1. Fases do Desenvolvimendo da Aplicação

A. Análise de Requisitos

Esta fase é composta por requisitos de ambiente, funcionais e não funcionais, conforme apresentado a seguir:

- **Requisitos de ambiente:** Para promover e usufruir de todas as funcionalidades da aplicação é necessário:
 - Estar num ambiente silencioso para melhor captação dos comandos pelo microfone.
 - Estar sentado de forma segura para evitar quedas, pois uma vez inserido no ambiente virtual, utilizando o Oculus Rift, perde-se a visão e a noção do espaço real.
 - Não fazer movimentos bruscos com a cabeça para reduzir a sensação de tontura e enjojo.
- **Requisitos funcionais:** O programa é executado online pela necessidade de acessar a aplicação Google Street View. Terá um campo para inserção de um endereço via comando de voz. A ação do usuário falar e o programa captar e interpretar as palavras ditas, tratam-se das principais funções que deverão ser executadas. Todos os comandos devem ser falados no idioma Inglês. Os requisitos funcionais podem ser descritos a seguir:
 - Entrar com o endereço;
 - Seguir em frente com o comando GO;
 - Retornar com o comando BACK;
 - Retornar à página inicial com o comando BACK TO SEARCH PAGE.
- **Requisitos não funcionais:** Corresponde ao tempo de resposta quando a aplicação é executada. O intervalo de segundos deve ser o menor possível entre a

recepção do comando por voz e a exibição da resposta, a fim de garantir uma busca rápida e permitir uma visualização do ambiente sem travar, sem atrasos e sem exibição de imagens borrasas. Atualmente, a aplicação demora, em média, 2 segundos a partir do momento do término da fala até a execução, o que é razoavelmente aceitável, de acordo com os testes com usuários. Se o local físico onde a pessoa está inserida tiver ruído de outras fontes, esse tempo aumenta até para mais de 5 segundos e, também, aumenta a probabilidade do reconhecimento errôneo da palavra pela aplicação.

B. Projeto

Este projeto foi desenvolvido em forma de aplicação web, escrito na linguagem Java devido a sua versatilidade em integrar com linguagens de programação distintas e o suporte do Java para diversas plataformas.

As ferramentas utilizadas para a implementação do programa foram o Oculus Software Development Kit – SDK v0.3.2 (ferramenta de configuração e bibliotecas do Oculus Rift fornecido pela própria Oculus VR, o Google Street View e o Netbeans IDE 7.3.1).

O Oculus Street View utilizado é uma versão modificada do Street View [11] para usar especialmente com o Oculus Rift. É possível fazer *download* dessa versão na comunidade de desenvolvedores de código aberto GitHub.

Foi utilizado também a biblioteca Annyang [12], que foi desenvolvida na linguagem Javascript para reconhecimento de voz em aplicações web. Ela não utiliza treinamento da voz para aumentar o grau de confiabilidade da voz. Esta biblioteca não possui suporte para o Português; por esta razão optou-se pelo idioma Inglês para o sistema de voz.

A aplicação desenvolvida funciona no navegador específico Google Chromium_webVR, que é uma versão alternativa do Google Chrome com modificações internas que lhe permite estabelecer comunicação direta com o Oculus. Esse navegador foi fabricado pela Google para ser usado para teste por programadores em aplicações web. O código dele é aberto.

A implementação do reconhecimento de voz na aplicação Street View utilizando o Oculus Rift DK, sendo necessário um servidor para a comunicação entre o navegador e o Oculus Rift. A versão DK2 não é exigida um servidor, mas precisa que o navegador tenha suporte ao WEBVR e ao WEBGL para funcionar.

Os navegadores que possuem suporte ao WebGL são o Chrome e o Firefox. O Chromium tem suporte ao WebGL e WebVR nativos, porém não disponíveis na versão de distribuição, somente na versão open source Chromium.

Existe uma limitação para o endereço requisitado que deve possuir coordenadas específicas para gerar apenas um resultado; por exemplo, se o comando for Go to Disney é informado e o usuário estiver longe desse local, um par de coordenadas desconhecidas é apresentada na tela do óculos. Isso ocorre porque o parque da Disney é uma área física muito grande.

C. Implementação

O desenvolvimento deste projeto teve uma duração de sete meses, a maior parte do tempo se passou em pesquisas de como desenvolver uma aplicação para o Oculus Rift.

A aplicação funciona de acordo com o diagrama apresentado na Figura 2, onde é possível verificar as relações entre cada parte.

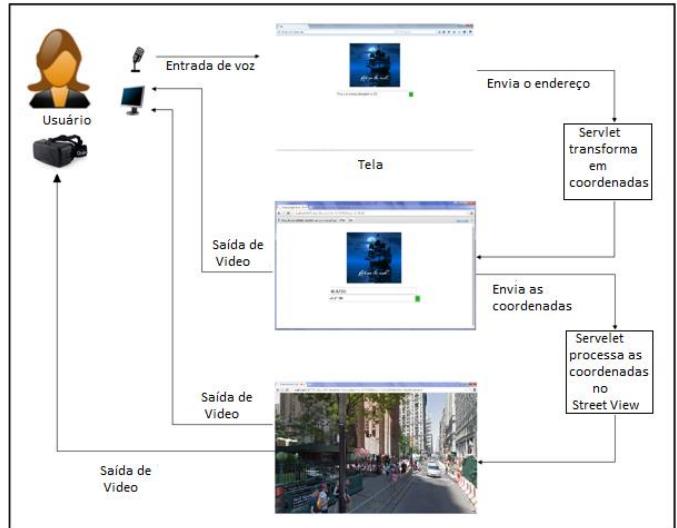


Fig. 2. Tela de Funcionamento do Programa

A tela principal da aplicação que se abre no navegador Chromium é parecida com uma página de busca comum, possui uma imagem atrativa no centro para chamar a atenção do usuário, um campo de texto que transforma as palavras faladas em texto do tipo String que será enviado para o processamento (Figura 3).

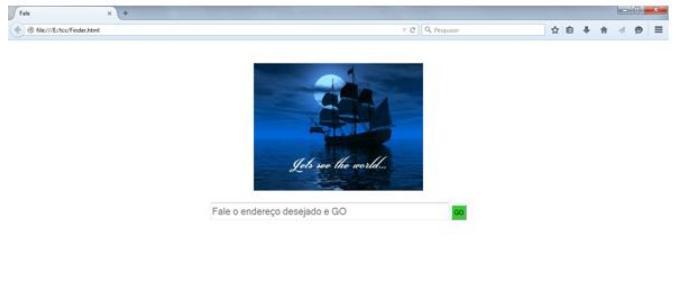


Fig. 3. Tela Inicial da Aplicação

Os comandos são recebidos por meio da voz que o microfone capta que simulam a digitação e o clique do mouse; não havendo a necessidade de usar as mãos para interagir com a aplicação.

Para visualizar o local desejado é necessário primeiramente dizer o comando seguido do endereço a ser interpretado; por exemplo, GO TO Wall Street New York. A mensagem via voz é captada pela biblioteca Annyang e enviada para uma servlet.

O Annyang processa a informação e mostra como saída o texto ditado no campo de texto onde estavam escritas as instruções de funcionamento em cinza “Fale o endereço desejado e GO” como uma variável do tipo String na tela inicial. Os comandos aceitos para buscar as coordenadas são GO e GO TO.

A servlet recebe a cadeia de caracteres que representa o endereço buscado e o envia para a API de Geocode do Google Maps, que transforma o endereço em coordenadas. A resposta dessa API é filtrada apenas para o nível de latitude e longitude, que são enviadas como variáveis para a página inicial. Essa medida intermediária foi definida para poder enviar as coordenadas na URL em forma de parâmetros para a aplicação do Oculus Rift como pode ser visualizada na Figura 4.

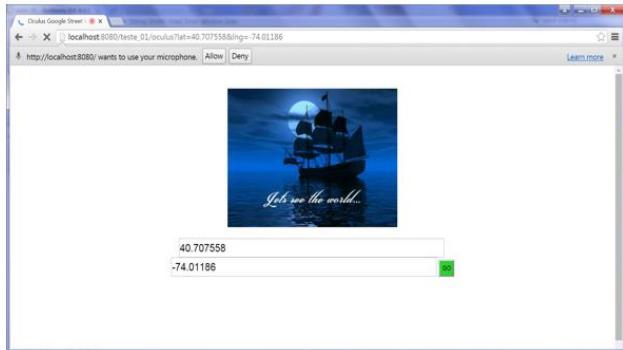


Fig. 4. Tela inicial da aplicação reconstruída contendo as coordenadas

As coordenadas que aparecem na imagem são enviadas para o Street View que exibirá a imagem do local solicitado no Oculus Rift como apresentado na Figura 5.



Fig. 5. Imagem do local localizado pelas coordenadas do endereço solicitado

Ao abrir essa tela, o usuário está livre para seguir o caminho que desejar por meio dos comandos de voz.

D. Testes

Para a melhoria progressiva do sistema e a integração efetiva com o usuário, foram criados dois questionários (pré-teste e pós-teste) com o objetivo de detectar falhas e pontos relevantes que devem ser modificados ou considerados.

O questionário de pré-teste tem como objetivo identificar o perfil dos usuários. Assim, versa sobre idade, sexo, grau de escolaridade, frequência de uso em tecnologia e internet e experiência com o uso de sistemas de mapa.

Os testes foram realizados com 15 pessoas, em três locais diferentes: na casa do grupo que desenvolveu a aplicação, na empresa em que um dos autores trabalha e na universidade em que os autores estudam/trabalham. Cada pessoa demorou de sete a dez minutos para testar a aplicação. Os equipamentos podem ser visualizados na Figura 6.



Fig. 6. Equipamentos utilizados

O segundo questionário (Tabela 1) é sobre a opinião dos usuários sobre usabilidade da aplicação. Esse foi preenchido após o uso da aplicação, com questões pertinentes ao mesmo e tem como finalidade a avaliação da experiência de uso, dinâmica, possíveis erros na execução. A tarefa era ditar um endereço; o programa vai exibir o local solicitado e o usuário deve andar entre as ruas através dos comandos de voz.

Tabela 1. Avaliação de usabilidade da aplicação

Experiência de uso do sistema				
Avalie o uso do sistema com os critérios descritos abaixo marcando com um X				
	Concordo fortemente	Concordo	Indeciso	Discordo
Esta aplicação tem uma apresentação gráfica agradável e legível				
É fácil entender o que é preciso fazer na aplicação				
A interface é bastante intuitiva				
Foi fácil aprender a utilizar a aplicação				
A velocidade de resposta é suficiente (tempo entre o reconhecimento do endereço e exibição do local)				
A aplicação funciona de maneira adequada quanto ao reconhecimento de minha pronúncia				
Quando faço hesitações (como, "mmmm", "aaah"), o sistema ignora captando somente palavras que definem o endereço				
Quando estou no Street View há congelamento da tela				
Quando viro a cabeça a imagem fica borrada				
Há distorções de cores quando se abre a imagem 3D				
Não consigo visualizar a imagem 100%				
É fácil se movimentar na aplicação utilizando o óculos Rift				
Usaria a aplicação para visitar lugares antes de ir fisicamente				
Senti tontura ao usar o óculos Rift				
Considerações sobre a aplicação				
Pontos positivos da aplicação:				
Pontos negativos da aplicação:				
Sugestões sobre a aplicação:				

IV. RESULTADOS

O resultado da aplicação do questionário de pré-teste aponta o perfil dos entrevistados: pessoas com faixa etária variada entre 14 e 46 anos, onde o maior grupo se concentra na faixa 20 anos a 29 anos, representando a metade da amostragem. Tratam-se de pessoas com nível de escolaridade de graduação e pós-graduação em sua maioria, que utilizam o computador por mais de 15 horas semanais e que já utilizam jogos digitais. Esse perfil aponta que todos utilizam o Google Maps e o Street View para procurar por endereços e atestam que esse último funciona bem.

Em relação aos resultados obtidos da aplicação do questionário de usabilidade (pós-teste) é possível perceber que:

1. 54% das pessoas concordaram fortemente e 46% das pessoas concordaram que a aplicação tem uma apresentação gráfica agradável e legível.
2. 54% dos entrevistados concordaram fortemente, 33% dos entrevistados concordaram e 13% dos entrevistados ficaram indecisos sobre ser fácil entender o que é preciso fazer na aplicação.
3. 47% dos usuários concordaram fortemente, 47% dos usuários concordaram e 6% dos usuários ficaram indecisos sobre a interface gráfica ser bastante intuitiva.
4. 67% do público-alvo concordou fortemente, 20% do público alvo concordou, 6% do público alvo não concordaram e 7% do público não respondeu que foi fácil aprender a utilizar a aplicação.
5. 34% das pessoas concordaram fortemente, 54% das pessoas concordaram, 6% ficaram indecisas e 6% das pessoas discordaram fortemente que a velocidade de resposta é suficiente.
6. 47% dos usuários concordaram fortemente, 6% dos usuários concordaram e 34% dos usuários ficaram indecisos sobre a aplicação funcionar de maneira adequada quanto ao reconhecimento da pronúncia.
7. 27% dos entrevistados concordaram fortemente, 27% dos entrevistados concordaram, 34% ficaram indecisos e 12% dos entrevistados não responderam sobre ao fazer hesitações como “mmh”, “Aaah”, o sistema ignora, captando somente palavras importantes que definem o endereço.
8. 6% das pessoas concordaram, 6% ficaram indecisas, 34% das pessoas discordaram, 41% discordaram fortemente e 12% das pessoas não responderam que quando estão no Street View há o congelamento na tela.
9. 2% do público-alvo concordou, 12% ficaram indecisos, 40% do público-alvo discordou e 36% do público-alvo discordou fortemente que quando viram a cabeça a imagem fica borrada.
10. 6% das pessoas concordaram fortemente 20% delas ficaram indecisas, 37% das pessoas discordaram e 37%

das pessoas discordaram fortemente sobre haver distorções de cores quando se abre a imagem 3D.

11. 12% dos entrevistados concordaram, 12% deles ficaram indecisos, 20% dos entrevistados discordaram e 40% deles discordaram fortemente que não conseguem visualizar a imagem 100%.
12. 60% dos usuários concordaram fortemente, 34% dos usuários concordaram e 6% dos usuários ficaram indecisos sobre ser fácil se movimentar na aplicação utilizando o dispositivo Oculus Rift.
13. 60% das pessoas concordaram fortemente e 40% das pessoas concordaram que tiveram a sensação de estarem presente fisicamente no local solicitado.
14. 54% do público-alvo concordou fortemente, 20% do público-alvo concordou, 20% do público-alvo ficaram indecisos e 6% do público-alvo discordou sobre usar a aplicação para visitar lugares antes de ir fisicamente.
15. 12% dos entrevistados concordaram fortemente, 27% dos entrevistados concordaram, 6% dos entrevistados ficaram indecisos, 27% dos entrevistados discordaram e 28% discordaram fortemente que sentiram tontura ao usar o dispositivo.

Através das informações supracitadas, é possível concluir, para este grupo de usuários que:

- A aplicação tem uma apresentação gráfica agradável e legível;
- É fácil entender o que é preciso fazer na aplicação;
- A interface gráfica é bastante intuitiva;
- É fácil aprender a utilizar a aplicação;
- A velocidade de resposta é suficiente;
- A aplicação funciona de maneira adequada quanto ao reconhecimento da pronúncia ao fazer hesitações como “mmh”, “Aaah”, o sistema ignora, captando somente palavras importantes que definem o endereço;
- Não há congelamento na imagem no Street View;
- Quando se faz movimentos com a cabeça a imagem não fica borrada;
- Não há distorções de cores quando se abre a imagem 3D;
- É sempre possível visualizar a imagem 100%;
- É fácil se movimentar na aplicação utilizando o dispositivo Oculus Rift;
- O programa fornece a sensação de estar presente fisicamente no local desejado;
- A maioria das pessoas usaria esse programa para visitar lugares antes de ir fisicamente;
- A sensação de tontura é variável de pessoa para pessoa, algumas sentirão e outras não.

V. CONSIDERAÇÕES FINAIS

Este trabalho apresentou o desenvolvimento de uma aplicação que integra o Oculus Rift com o aplicativo Street View. A aplicação utiliza um sistema de voz para entender os locais anunciados que são localizados pelo programa, contribuindo com uma melhor interação humano-máquina, pois deixa as mãos e o corpo livres no espaço físico para navegar no mundo virtual. Testes com 15 usuários foram realizados e os resultados para este grupo são bastante satisfatórios, conforme apresentados na seção 5.

De acordo com as informações coletadas no grupo de usuários, seria interessante a inserção de áudio no programa como um trabalho futuro. Duas opções são analisadas, a primeira uma música ambiente e agradável para relaxar enquanto visualiza o lugar escolhido; e a segunda, um ruído de cidade grande, com barulho de carros passando e buzinando algumas vezes e pessoas falando para simular melhor a sensação de presença física.

Sabe-se que o uso de novas tecnologias representa um desafio de implementação para os desenvolvedores. Este projeto, por ser inovador, também apresentou diversos desafios e dificuldades, desde o atraso na entrega do equipamento até o não suporte de navegadores para a navegação, o que consumiu muitas horas de projeto.

REFERÊNCIAS

- [1] Biocca, F., & Levy, M. R. (Eds.). (2013). *Communication in the age of virtual reality*. Routledge.
- [2] Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., & MacIntyre, B. (2001). Recent advances in augmented reality. *Computer Graphics and Applications*, 21(6), 34-47.
- [3] Cohen, M. H., Cohen, M. H., Giangola, J. P., & Balogh, J. (2004). *Voice user interface design*. Addison-Wesley Professional.
- [4] Mistry, P., & Maes, P. (2009, December). SixthSense: a wearable gestural interface. In ACM SIGGRAPH ASIA 2009 Sketches (p. 11). ACM.
- [5] Ribeiro, M. W. D. S., & Zorral, E. R. (2011). Realidade Virtual e Aumentada: Aplicações e Tendências. XIII Simpósio de Realidade Virtual e Aumentada, Uberlândia-MG-Brasil.
- [6] Burdea, G., & Coiffet, P. (2003). Virtual reality technology. *Presence: Teleoperators and virtual environments*, 12(6), 663-664.
- [7] Oculus VR. Oculus Rift NEXT-GEN VIRTUAL REALITY. Oculus VR. Disponível em: <<http://www.oculusvr.com/rift/>>. Acesso em: 19 mar. 2015
- [8] Reiners, T., Teras, H., Chang, V., Wood, L. C., Gregory, S., Gibson, D., ... & Teras, M. (2014). Authentic, immersive, and emotional experience in virtual learning environments: The fear of dying as an important learning experience in a simulation.
- [9] Hoffman, H. G., Meyer III, W. J., Ramirez, M., Roberts, L., Seibel, E. J., Atzori, B., ... & Patterson, D. R. (2014). Feasibility of articulated arm mounted oculus rift virtual reality goggles for adjunctive pain control during occupational therapy in pediatric burn patients. *Cyberpsychology, Behavior, and Social Networking*, 17(6), 397-401.
- [10] Google Maps. Nos bastidores do Street View. Google. Disponível em: <<http://www.google.com/intl/pt-PT/maps/about/behind-the-scenes/streetview/#devices>>. Acesso em: 23 mar. 2015.
- [11] Oculus Street View. Disponível em: <https://github.com/troffmo5/OculusStreetView>. Acesso em: 02/10/2015
- [12] Annyang. Speech Recognition that just works. Tal Ater. Disponível em: <<https://www.talater.com/annyang/>> Acesso em: 11 fev. 2015.

Sincronização de ambientes virtuais interativos por meio do grafo de cena do padrão X3D

Alisson Fernando Coelho do Carmo

PPGCC – Programa de Pós-Graduação
em Ciências Cartográficas

FCT/UNESP – Faculdade de Ciência e Tecnologia
Presidente Prudente, São Paulo, Brasil
alisondocarmo@gmail.com

Milton Hirokazu Shimabukuro

DMC - Departamento de Matemática e Computação
FCT/UNESP – Faculdade de Ciência e Tecnologia
Presidente Prudente, São Paulo, Brasil
miltonhs@fct.unesp.br

Abstract—This article describes the development and results in a research focused on sharing and synchronization of interactive three dimensional virtual worlds, using the features of X3D standard and Java Xj3D library. The approach is based on implementing a middleware with modules positioned in an intermediate layer between the representation of the scene with the standard X3D and processes and interactions with Xj3D library. The middleware offers features for creating three-dimensional environment, dynamic manipulation of the scene graph, synchronization of environments via network and to export graphics and interactive elements events. The results show a possibility for expansion of the capabilities from the use of X3D standard related to synchronous and asynchronous exploitation of the three-dimensional environment.

Keywords—X3D; Interactive Virtual Environments; Distributed Synchronization.

Resumo—Este artigo descreve o desenvolvimento e os resultados alcançados na investigação focada no compartilhamento e sincronização de mundos virtuais tridimensionais interativos, utilizando os recursos do padrão X3D e da biblioteca Java Xj3D. A abordagem utilizada é baseada na implementação de um middleware composto por módulos posicionados em uma camada intermediária, entre a representação da cena com o padrão X3D e os processamentos e interações com a biblioteca Xj3D. O middleware construído oferece recursos para criação do ambiente tridimensional, manipulação dinâmica do grafo de cena, sincronização dos ambientes via rede e exportação dos elementos gráficos e eventos interativos. Os resultados obtidos demonstram uma possibilidade para ampliação dos recursos obtidos na utilização do padrão X3D relacionados à exploração síncrona e assíncrona do ambiente tridimensional.

Palavras Chaves—X3D; Ambiente Virtual Interativo; Sincronização Distribuída.

I. INTRODUÇÃO

A eficiência e níveis de detalhes de representações visuais se beneficiam com o crescente avanço das tecnologias que oferecem a possibilidade de manipulação gráfica. Este fato tem popularizado cada vez mais a utilização do efeito de tridimensionalidade em ambientes virtuais. A manipulação e visualização tridimensional podem oferecer ao usuário um espaço virtual que se aproxime da realidade, de modo que não exija uma adaptação rigorosa do usuário ao novo universo do sistema virtual.

A exploração dos cenários virtuais, aplicáveis em diferentes contextos, pode ter seus resultados amplificados com a possibilidade de compartilhamento do mundo virtual entre diferentes agentes. O compartilhamento do mundo virtual caracteriza um cenário de colaboração aplicável ao contexto de exploração que pode ser síncrona – exibição coordenada simultânea – ou assíncrona – *replay* posterior.

A integração de fatores referentes à visualização tridimensional, interatividade e colaboração, pode ser realizada com tecnologias voltadas para Realidade Virtual (RV). Existem diversas plataformas que oferecem tais recursos que possibilitam a integração e publicação de conteúdos em ambiente Web, tais como WebGL, Unit3D e Flash. No entanto, a utilização destas tecnologias são dependentes do aprendizado de técnicas de codificação em linguagens de programação específica, como a linguagem C ou Javascript para Unit3D e Actionscript para construção de conteúdos dinâmicos em Flash. Por outro lado, existem tecnologias que definem padrões de codificações de conteúdos gráficos, como SVG, VRML, X3D.

O padrão X3D (*eXtensible 3D*) oferece recursos de interatividade e representação gráfica que podem potencializar a visualização em um ambiente tridimensional. Este é capaz de oferecer os recursos relacionados à representação, no entanto, para a construção de ambientes dinâmicos e com recursos de interatividade mais complexos pode ser necessário manipular a cena dinamicamente. Para isto, pode ser utilizada uma biblioteca de linguagem de programação que implemente os recursos da *Scene Access Interface*(SAI)¹, como a biblioteca Java Xj3D.

A principal motivação que justifica o desenvolvimento deste trabalho está no fato de que tanto o padrão X3D como a biblioteca Xj3D, não oferecem nenhum recurso nativo voltado à colaboração síncrona ou assíncrona. Para a colaboração assíncrona, deve existir uma forma de exportar a cena construída, para posterior execução, porém a SAI não oferece nenhum recurso para tal. Da mesma forma a colaboração síncrona não é imediata pois não existe nenhum recurso no padrão X3D para a sincronização de cenas em ambientes remotos. Neste contexto, este trabalho apresenta as investigações realizadas para permitir a representação de

¹A SAI está descrita na especificação do padrão X3D e comprehende um conjunto de funcionalidades e definições abstratas que devem ser implementadas por bibliotecas para serem capazes de manipular este padrão

informações em um ambiente tridimensional interativo que permita o compartilhamento, síncrono ou assíncrono.

Para transpor algumas barreiras relativas aos recursos necessários para o ambiente, neste trabalho, uma abordagem com um *middleware* é utilizada para gerenciar o cenário tridimensional. Este componente concentra todas as funcionalidades de representação visual, interatividade e colaboração em uma camada intermediária. Com isto, as cenas virtuais geradas pelo *middleware* podem ser sincronizadas através da rede em instâncias alocadas em máquinas distintas – colaboração síncrona –, além de permitir a exportação da representação visual gerada – colaboração assíncrona.

Este trabalho apresenta alguns resultados obtidos em estudos iniciados em um programa de formação complementar, realizado no Departamento de Matemática e Computação (DMC/FCT/UNESP), e refinados em um projeto de Iniciação Científica apoiado pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). Alguns dos resultados parciais podem ser encontrados detalhados em outros dois artigos apresentados no VIII Workshop de Realidade Virtual e Aumentada (WRVA 2011) [1] e na XXV Conference on Graphics, Patterns and Images (SIBGRAPI 2012) [2].

As demais Seções descrevem os principais aspectos envolvidos no desenvolvimento do trabalho. Na Seção II é apresentada uma breve descrição sobre os materiais e tecnologias utilizados. A Seção III apresenta algumas capacidades aplicadas no *middleware* desenvolvido. Os resultados dos testes realizados com a execução do *middleware*, como a exportação e sincronização do grafo de cena, são apresentados na Seção IV para então subsidiar as conclusões apresentadas na Seção V.

II. MATERIAIS E TECNOLOGIAS

Neste trabalho, foram priorizadas ferramentas e tecnologias abertas e livres. Além destas características, a escolha das ferramentas teve como critérios a independência de ambiente computacional para execução (portabilidade), a facilidade para o uso integrado das soluções criadas e a disponibilidade de recursos para comunicação via internet.

X3D é um padrão aberto de formato de arquivo e arquitetura de execução, internacionalmente reconhecido e aprovado pela ISO (*International Organization for Standardization*), livre de royalties, para representação e comunicação de cenas e objetos 3D utilizando XML (*eXtensible Markup Language*) [3]. É o padrão sucessor do VRML (*Virtual Reality Modeling Language*), no qual foram agregadas diversas melhorias originadas de respostas obtidas com a utilização do VRML pela comunidade, oferecendo subsídios para a construção de um padrão mais maduro e avançado que seu antecessor.

É um padrão atualmente adotado pela W3C (*World Wide Web Consortium*) para compartilhamento e publicação de gráficos tridimensionais na WEB, o qual existe um esforço mútuo entre grupos de trabalhos para a integração completa entre HTML (*HyperText Markup Language*) e X3D, assim como ocorre com o SVG (*Scalable Vector Graphics*) no contexto de gráficos bidimensionais em páginas WEB. Os avanços recentemente obtidos nesta vertente estão associados à definição do X3DOM (composição das siglas X3D e DOM - *Document Object Model*) como um ambiente de execução

capaz de inserir os recursos X3D diretamente em páginas HTML5 em um contexto declarativo de elementos tridimensionais pertencentes à própria hierarquia de nós HTML [4]. Tal integração é possível em razão de outras bibliotecas 3D que podem ser integradas com o elemento Canvas do HTML5 sem a necessidade de nenhum plugin adicional, como o renderizador WebGL [5].

O padrão X3D define um ambiente tridimensional através da construção de um grafo de cena (*Scene Graph*), representado pela estrutura XML. Cada elemento do grafo é chamado de nó. Cada nó representa um objeto definido nas especificações do padrão X3D. O padrão de utilizar um grafo de cena para a representação hierárquica está presente em diversas ferramentas computacionais e é um recurso de alto nível porque permite ao desenvolvedor focar no que renderizar ao invés de como, e foi popularizado pelo *SGI Open Inventor* [6], [7].

A variedade de recursos disponíveis no padrão X3D, conforme pode ser observado na publicação de Brutzman e Daly [8] e nas próprias especificações do padrão [3], atendem as necessidades para a construção de objetos e ambientes tridimensionais complexos. Um dos motivos para tal é a concepção do padrão X3D baseada nas características de seu predecessor VRML agregadas à maior facilidade para a integração de *scripts*, modularização dos interpretadores, formatos de arquivos comprimidos e binários, entre outros. Um dos principais avanços do padrão X3D é a própria estrutura de codificação utilizando o formato XML, mantendo ainda a retrocompatibilidade com o padrão VRML.

Dentre outras definições, nas especificações [3] do padrão X3D está caracterizada a existência da SAI, que é uma interface para todas as funcionalidades necessárias para manipulação dinâmica do grafo de cena, possibilitando o desenvolvimento de APIs (*Application Programming Interface*) em diversas linguagens de programação.

Uma das APIs disponíveis para manipulação do padrão X3D é o Xj3D que é um projeto do Web3D Consortium que oferece um conjunto de ferramentas totalmente desenvolvido na plataforma Java para a manipulação do padrão X3D [9]. Xj3D é composto de um Browser X3D, capaz de interpretar também o formato VRML, e uma biblioteca Java para a manipulação dinâmica do grafo de cena. Essa biblioteca implementa a SAI, tornando possível a criação, acesso e manipulação dinâmica da cena X3D, aliando todos os recursos oferecidos pelo padrão X3D à capacidade da linguagem de programação Java.

A utilização da linguagem Java permite que a aplicação seja executada em sistemas operacionais locais ou na WEB, como um *applet*, por exemplo. Todas as funcionalidades oferecidas pela linguagem Java podem ser combinadas ao poder de representação e interatividade do padrão X3D por meio da biblioteca Xj3D.

III. CONSTRUÇÃO E SINCRONIZAÇÃO DE AMBIENTES VIRTUAIS

Na implementação de um ambiente colaborativo e interativo para Visualização de Informação, existem três objetivos gerais, que são necessários no processo de criação e distribuição da cena gráfica:

- *Construção dinâmica do ambiente:* O mesmo ambiente virtual deve ser conhecido e visualizado por todos os usuários, de forma que todos possam explorar o mesmo ambiente composto pelos mesmos elementos gráficos;
- *Compartilhamento dos elementos do ambiente:* Após a geração do ambiente virtual, toda manipulação dinâmica dos objetos gráficos devem ser distribuídas e replicadas para os outros usuários conectados;
- *Sincronização de eventos:* Uma vez que todos os usuários possuem a mesma visão, contendo os mesmos objetos gráficos, todos os eventos relevantes ocorridos no ambiente pela interação, devem ser sincronizados em todos as instâncias conectadas.

A. Abordagem com um Middleware

Apesar da possibilidade de construção dinâmica da cena X3D diretamente com a biblioteca Xj3D, o compartilhamento da cena X3D construída a partir da biblioteca Xj3D não é uma tarefa imediata, pois, conforme apresentado por Carmo e Shimabukuro [1], os objetos que definem a cena não podem ser transmitidos diretamente pela rede por conterem componentes não serializáveis. Diante deste cenário, Carmo e Shimabukuro [2] propõem a criação de um *middleware* para gerenciar o grafo de cena voltado para a análise e representação visual por meio de técnicas de Visualização de Informação, em um ambiente tridimensional, interativo e colaborativo.

O *middleware* está posicionado em uma camada intermediária, configurando-o como ponte de comunicação entre a cena X3D e as funções da biblioteca Xj3D que a manipulam. Com isto, as responsabilidades do *middleware* podem ser expandidas, de modo que todas as operações relacionadas com a manipulação da cena, sincronização e interatividade possam ser concentradas em uma única arquitetura modularizada. A Figura 1 representa a arquitetura dos componentes que compõem o *middleware*.

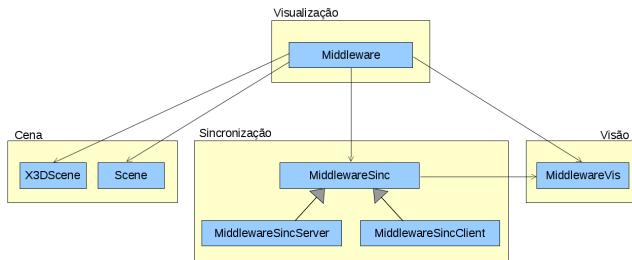


Figura 1. Estrutura modular do *middleware*

Como ilustrado no diagrama da Figura 1, o *middleware* pode ser segmentado em quatro módulos principais: Componentes da Cena; Componente de Sincronização; Componente de Visão e Componente de Visualização. A integração destes componentes define um ambiente com as funcionalidades de visualização, sincronização e interatividade. Os módulos da arquitetura definida por do Carmo e Shimabukuro [2] são descritos nas Seções seguintes, bem como os principais recursos relacionados.

1) *Componentes da Cena:* A construção do *middleware* surgiu da necessidade de se conhecer e registrar os atributos dos nós a serem adicionados à cena antes de sua inserção efetiva, em razão dos atributos com tipo de acesso restrito a inicialização, como descrito por Carmo e Shimabukuro [1].

Para superar esta dificuldade, este módulo é responsável por definir componentes que representam diretamente cada nó a ser utilizado para a construção da cena X3D. Com isto, estes componentes definem um grafo de cena abstrato e armazena todas as informações dos nós que a compõem, permitindo que estes dados sejam recuperados a qualquer tempo, bem como transmitidos, reconstruídos e manipulados. Estes componentes da cena foram construídos por meio de heranças de classes específicas que seguem a mesma disposição hierárquica dos nós definidos na especificação X3D.

Para que cada nó seja capaz de manipular diretamente seu correspondente na cena X3D real, existem dois métodos que devem estar presentes em cada um destes:

- *writeX3dScene:* função capaz de converter o referido nó abstrato em um nó X3D para ser adicionado à cena. Caso seja um nó responsável por agrupar outros nós, todos os outros são ativados recursivamente.
- *toXml:* a partir dos atributos definidos no nó abstrato, esta função o transforma na representação do nó de acordo com o formato XML do padrão X3D, permitindo que a cena seja construída e exportada para um arquivo externo.

2) *Sincronização:* O módulo de sincronização é responsável por conter todas as funções a serem executadas durante o processo de compartilhamento da cena. Neste cenário, existem dois componentes principais: cliente e servidor.

Os termos servidor e cliente, são utilizados para identificar, respectivamente, a instância ou usuário que possui o controle sobre a exploração do ambiente e aquele que apenas interage e acompanha a exploração do outro. Vale ressaltar que o *middleware* foi construído com a proposta de ser aplicado no contexto de exploração e análise de dados, cenário na qual a sincronização da visão do analista é uma informação importante para permitir a análise colaborativa. Esta arquitetura permite que apenas o servidor possua o conjunto de dados a serem analisados e todos as outras instâncias ainda possam interagir com as representações visuais.

Inicialmente, a cena X3D é gerada através do *middleware* e distribuída para os clientes conectados. Desta forma, todo processo de navegação e exploração do ambiente, executado pelo servidor será retransmitido aos clientes. No entanto, para aumentar a flexibilidade no processo de análise e exploração, o servidor pode transferir o controle para outro cliente, alterando a instância responsável pela execução da navegação.

Para que o processo de comunicação entre as instâncias conectadas seja realizado com sucesso, é necessária a definição de um protocolo de comunicação comum. Para isto, foi definido um segmento capaz de encapsular todas as mensagens de comunicação que trafegam entre as instâncias. Este segmento possui, basicamente, dois atributos: identificação da mensagem e conteúdo da mensagem. Exemplos de identificadores das mensagens são: criação de nova cena; adição de novo nó;

remoção de um nó existente; redefinição de servidor; alteração de atributo de um nó existente; solicitação de comando de servidor; atualização de câmera, entre outros.

Desta forma, ao detectar um evento, que deve ser transmitido aos outros clientes, o servidor cria uma nova mensagem, define o identificador, encapsula o conteúdo a ser transmitido e envia para todos os clientes. Dependendo do identificador da mensagem, pode ser necessário realizar mais de uma transmissão em várias iterações. Já o comportamento do cliente é definido após a sincronização inicial, onde é criado um ouvinte que aguarda o recebimento de novas mensagens com atualizações. Ao receber uma nova mensagem, o conteúdo é interpretado e convertido de acordo com o identificador registrado.

3) Visão: O componente de sincronização descrito na Seção III-A2 depende da geração e identificação de eventos durante a exploração, para efetuar a transmissão e sincronização da cena. Um dos eventos que pode ser transmitido, e que representa uma informação relevante no contexto de análise e exploração é a visão ou navegação no ambiente.

Com isto, todas as instâncias conectadas poderão ter a mesma perspectiva da visão do servidor da análise. Enquanto o servidor explora e interage com o ambiente, este componente captura a posição de cada movimento, que é encapsulado em uma mensagem pelo módulo de sincronização e transmitido aos clientes sincronizados. Todas as capturas de movimento são realizadas através de nós X3D sensores adicionados à cena.

4) Visualização: Os outros componentes do *middleware*, apresentados nas seções anteriores, tratam de assuntos relacionados com a construção do ambiente interativo e colaborativo para suportar o processo de visualização. Desta forma, este módulo é responsável pelas implementações das técnicas de visualizações e manipulação dos dados de entrada.

O módulo de Visualização é capaz de ler os dados de entrada, representá-los em uma estrutura intermediária, com as informações necessárias, para posterior conversão em uma representação na cena genérica, utilizando os componentes de cena descritos na Seção III-A1.

Portanto, para que a visualização aproveite todos os recursos de interatividade e colaboração disponíveis no *middleware*, basta que esta seja implementada neste componente de Visualização, respeitando a utilização dos componentes definidos na cena abstrata. Desta forma, a modularização do *middleware* proporciona um ganho relacionado à sua capacidade de extensão, no que diz respeito à implementação de novas técnicas de visualização.

B. Colaboração Síncrona: Sincronização da visão

A primeira etapa para sincronização dos ambientes é satisfeita a partir do sucesso no compartilhamento inicial e reconstrução da cena X3D através do *middleware*. No entanto, a partir deste momento, devem existir mecanismos que mantenham sincronizados todos os ambientes compartilhados. Para atender e manter a exploração do conteúdo gráfico sincronizada, é necessário, primeiramente, que todos os eventos envolvidos na exploração do ambiente sejam transmitidos a todos os usuários conectados.

Após investigação dos componentes X3D, notou-se a existência de recursos da própria SAI que poderiam ser aproveitados para a captura da posição de visão do usuário através de eventos da biblioteca Xj3D. No entanto, para privilegiar a utilização de recursos presentes no próprio padrão X3D, optou-se por estudar os comportamentos dos sensores de posição no ambiente X3D, mais especificamente o nó *ProximitySensor*, nativamente suportado pelo padrão.

O nó *ProximitySensor* possui todos os atributos necessários para se definir a posição visual do usuário, são eles: *position_changed*, *orientation_changed* e *centerOfRotation_changed*. Como a própria nomenclatura adotada pelos atributos dos nós X3D indica, todos são campos de saída de dados – geram dados –, que representam respectivamente a posição, orientação e o centro de rotação da visão do usuário. Este nó é adicionado ao módulo de visão, descrito na Seção III-A3, e utilizado para recuperação da perspectiva de visão do usuário.

A partir dos atributos recuperados do sensor, pode-se sincronizar a visão na cena aplicando-os diretamente sobre um ponto de visão, ou seja, um nó *ViewPoint* definido. Todo *ViewPoint* possui campos relacionados à posição, orientação e centro de rotação, respectivamente os campos de entrada e saída para: *position*, *orientation* e *centerOfRotation*.

C. Colaboração Assíncrona: Exportação da exploração

Uma das formas de colaboração assíncrona é a possibilidade de outros usuários replicarem as ações previamente realizadas, por meio de um mecanismo de *replay*. Como já abordado, o *middleware* utiliza uma representação intermediária do grafo da cena X3D. Nesta representação estão presentes todos os atributos que caracterizam cada nó, bem como suas relações hierárquicas que configuram o grafo. Tais garantias permitem que o *middleware* seja capaz de reconstruir a representação em arquivo no padrão X3D.

Para isto, a função de converter a representação em formato XML foi implementada em cada nó que compõe o *middleware*. A opção pela representação em formato XML, em oposição ao clássico VRML, é em razão da facilidade na interpretação do código gerado oferecida pelo padrão X3D.

A cena construída dinamicamente a partir da biblioteca Xj3D e representada através do *middleware*, além de ser transmitida pela rede e posteriormente reconstruída, pode ser exportada para a própria estrutura X3D, recurso previamente inexistente no conjunto de ferramentas Xj3D.

A exportação da cena X3D resolve o problema da replicação/construção posterior da cena, mas não mantém as ações de exploração desempenhadas. Observando a abordagem utilizada para a sincronização da cena X3D, apresentada na Seção III-B, é possível notar que a única diferença entre o processo síncrono e assíncrono é justamente o momento em que a sincronização ocorre: em tempo real ou posterior, respectivamente. Com esta abordagem, a captura dos movimentos está garantida, mas ainda necessita de um mecanismo para atualizar os movimentos no momento adequado, através do próprio *ViewPoint*.

Com o estudo direcionado sobre os aspectos envolvidos na animação de objetos no ambiente X3D, notou-se a possi-

bilidade de aliar os recursos da definição de rotas de eventos com os componentes interpoladores, capazes de representar o momento de acontecimento de um evento qualquer. Os componentes interpoladores utilizam dois conjuntos de dados, um para controlar o índice ou chave e o outro para relacionar cada informação de saída. Estes interpoladores geralmente operam em conjunto com algum tipo de sensor automático, neste caso, um sensor de tempo (*TimeSensor*), que dispara um evento a cada ciclo de tempo. A definição de uma rota de eventos entre estes dois nós – sensor e interpolador – permite a iteração automática e previamente definida dos valores armazenados.

O tipo de interpolador utilizado, depende única e exclusivamente do tipo de dado a ser armazenado. Neste caso, são utilizados dois componentes *PositionInterpolator* – para interpolação dos valores de posição e centro de rotação – e um *OrientationInterpolator* – para a orientação, além do próprio *TimeSensor* contendo os intervalos de atualização dos movimentos, bem como as rotas de eventos entre estes componentes.

Por fim, após a definição destes componentes, e inserção na cena abstrata do *middleware*, a exportação da cena permite que todos estes componentes estejam registrados no próprio arquivo X3D, e possam ser reexecutados em qualquer outro Brower X3D, sendo que a animação de exploração é automaticamente executada a partir da rota de eventos definida e os pontos de visão capturados e registrados durante a exportação.

IV. RESULTADOS E DISCUSSÕES

Para avaliar as capacidades de criação dinâmica de representações visuais, os recursos de colaboração síncrona e assíncrona, e a interatividade do ambiente, o *middleware* foi aplicado em um ambiente de teste utilizando dados aleatórios para a construção da cena tridimensional. Após a adição de algumas geometrias ao ambiente 3D, a cena foi compartilhada para outras três instâncias do *middleware* executadas na mesma máquina, como pode ser visto na Figura 2. Neste cenário, apenas uma das instâncias é o servidor da cena, na qual foi construída e gerada inicialmente e responsável pela exploração do ambiente, que é replicado aos outros clientes conectados. Apesar de neste caso o teste ter sido realizado em ambiente local, o *middleware* desenvolvido permite a sincronização a partir da rede, baseado na localização do endereço e porta dos processos que executam em máquinas remotas.

Para verificar a capacidade de replicação assíncrona, a mesma cena gerada na Figura 2 foi exportada e executada corretamente em outro Brower X3D, independente do *middleware*. Como o foco deste teste era avaliar a independência da cena gerada e exportada com o próprio *middleware*, para facilitar posterior execução, foi utilizado o Brower opensource FreeWRL. O resultado da execução no Brower FreeWRL, da cena gerada e exportada pelo *middleware* pode ser observado na Figura 3.

Utilizando esta mesma estrutura do *middleware*, ainda existe outra alternativa para o compartilhamento da cena X3D, baseada na transmissão do código X3D, representado em estrutura XML. Uma vez que é possível percorrer recursivamente o objeto da cena da camada intermediária para a reconstrução do ambiente X3D, a recursão pode ser utilizada para construir o código da cena X3D, e então transmiti-lo. Por ser representado

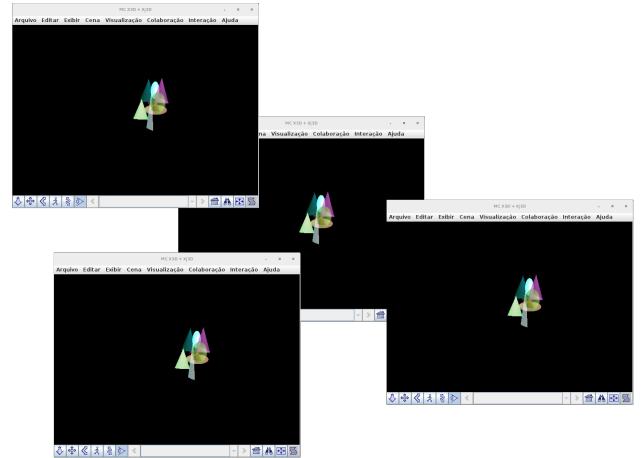


Figura 2. Representação de uma cena com geometrias aleatórias sincronizadas em outras instâncias do *middleware*

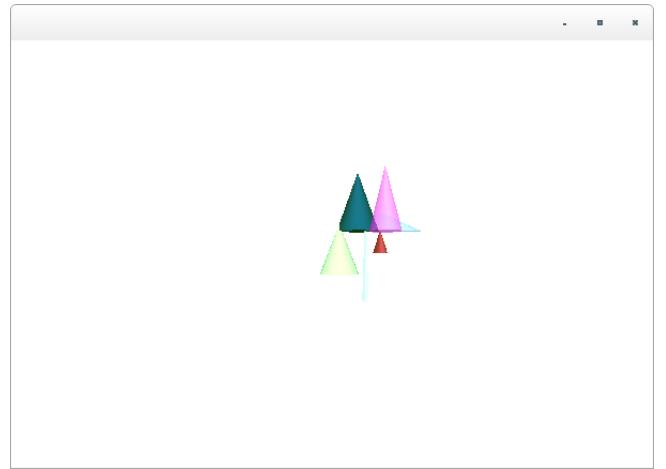


Figura 3. Cena com geometrias aleatórias gerada e exportada a partir do *middleware* e executada no Brower FreeWRL

em estrutura de texto puro, este pode ser facilmente transmitido pela rede.

A possibilidade de se exportar a cena diretamente para um arquivo no padrão X3D fomenta a análise relacionada ao tipo de representação serializável – arquivo XML e objeto cena do *middleware* – que se torna mais viável para o compartilhamento da cena via rede. Uma vez que o tráfego de dados, principalmente relacionados a ambientes complexos, pode se transformar em um fator de impacto relevante.

Para avaliar esta situação, foi observado o tamanho, em bytes, de cada representação de acordo com a quantidade de nós inseridos na cena. Como o objetivo é verificar a relação entre a quantidade de nós e o tamanho da representação, foi configurada uma situação de inserção incremental de nós contendo a representação geométrica de uma esfera, posicionada aleatoriamente no ambiente. Após cada inserção, foi capturado o tamanho dos objetos serializados, tanto da cena abstrata do *middleware*, quanto da cena exportada em arquivo XML.

O resultado deste teste pode ser observado na Figura 4. É possível notar o crescimento linear de ambas as representações,

no entanto, o ganho referente à representação a partir da cena abstrata do *middleware* torna-se relevante e significativo conforme o aumento da quantidade de nós da cena. Isto prova a eficiência em utilização do espaço da representação da cena X3D através do *middleware*, no que diz respeito ao compartilhamento da cena X3D via rede.

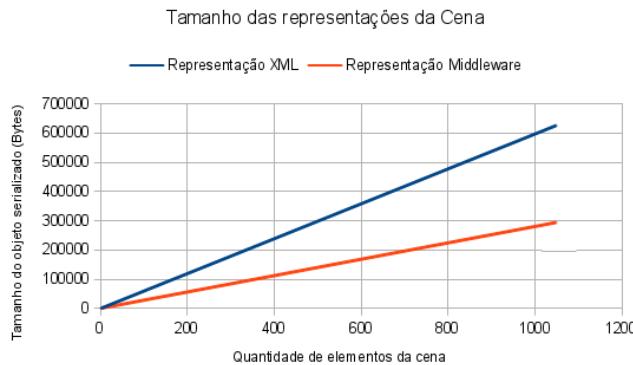


Figura 4. Gráfico comparativo entre a representação serializada por estrutura XML e cena abstrata do *middleware*.

Vale ressaltar que neste teste foram contabilizados apenas a quantidade de elementos gráficos presentes na cena. A complexidade de detalhamento de tais elementos não influenciam no resultado final do tamanho (bytes), pois a mesma cena, contendo a mesma quantidade de nós - consequentemente o mesmo nível de detalhe - foi utilizada em ambos as representações (*middleware* e estrutura XML).

V. CONCLUSÕES

Este trabalho aborda uma estratégia baseada em um *middleware* para a integração de recursos de sincronização em um ambiente de representação tridimensional. Neste cenário, a cena é construída dinamicamente, pode ser exportada para o formato X3D – tornando-se independente do *middleware* –, sincronizada com outras instâncias através da rede e ainda se beneficiar de todos os recursos de interatividade, representação e exploração tridimensional oferecida pelo padrão X3D.

O *middleware* proposto oferece algumas soluções para funcionalidades inexistentes tanto no padrão X3D, quanto na biblioteca Xj3D, como compartilhamento da cena X3D construída dinamicamente, sincronização da visão, exportação da cena X3D, além dos próprios recursos de interatividade do X3D.

O padrão X3D possui capacidades e responsabilidades de um padrão maduro e bem desenvolvido, em razão de ter herdado e melhorado vários aspectos da anterior tecnologia VRML. Por esta razão continua sendo um padrão W3C para comunicação de gráficos tridimensionais na WEB. Além disso, com a evolução e amadurecimento no uso de representações gráficas em ambientes interativos, são consideradas várias dimensões inter-relacionadas – sistema, usuário, colaboração e interação homem-máquina, para as quais a especificação de padrões abertos e a definição de novas tecnologias têm como alvos a interoperabilidade e a flexibilidade. O padrão X3D, a biblioteca Xj3D, a plataforma Java e a estruturação XML são ferramentas muito ricas que podem ser aplicadas

nas mais diversas áreas, beneficiando-se das capacidades de interatividade, representação e animação de objetos gráficos.

Sobretudo, com os resultados obtidos neste trabalho, pôde-se verificar a relevância e viabilidade da utilização do *middleware* proposto para suportar a manipulação da cena com recursos colaborativos e interativos. O ambiente construído oferece recursos de interatividade –através dos recursos exploratórios do X3D– e colaboração –por meio da sincronização ou *replay*– para enriquecer a interação com o mundo virtual.

AGRADECIMENTOS

Agradecimentos especiais ao Departamento de Matemática e Computação (DMC) da Faculdade de Ciências e Tecnologia/Unesp (FCT/UNESP) Campus de Presidente Prudente por ter propiciado o desenvolvimento inicial desta investigação através do Programa de Formação Complementar bem como pela estrutura oferecida para o desenvolvimento do projeto; à Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP) pelo financiamento e apoio à pesquisa em projeto de Iniciação Científica.

REFERÊNCIAS

- [1] A. F. C. do Carmo and M. H. Shimabukuro, "Estudo sobre o uso do x3d e java para a análise visual 3d interativa e colaborativa de dados via internet," in *[Anais do] WRVA'2011 8 Workshop de Realidade Virtual e Aumentada*, Uberaba, MG, Brasil, 2011.
- [2] ———, "Visualização de informação com recursos colaborativos e interativos utilizando o padrão x3d," *[Anais do] XXV Conference on Graphics, Patterns and Images (SIBGRAPI 2012)*, pp. 150–155, 2012.
- [3] Web3D, "X3d, h-anim, and vrml97 specifications," <http://www.web3d.org/x3d/specifications/>, outubro 2015.
- [4] J. Behr, P. Eschler, Y. Jung, and M. Zöllner, "X3dom: A domain-based html5/x3d integration model," in *Proceedings of the 14th International Conference on 3D Web Technology*, ser. Web3D '09. New York, NY, USA: ACM, 2009, pp. 127–135. [Online]. Available: <http://doi.acm.org/10.1145/1559764.1559784>
- [5] J. Behr, Y. Jung, J. Keil, T. Drevensek, M. Zoellner, P. Eschler, and D. Fellner, "A scalable architecture for the html5/x3d integration model x3dom," in *Proceedings of the 15th International Conference on Web 3D Technology*, ser. Web3D '10. New York, NY, USA: ACM, 2010, pp. 185–194. [Online]. Available: <http://doi.acm.org/10.1145/1836049.1836077>
- [6] A. E. Walsh, "Understanding scene graphs," *Dr Dobbs Journal of Software Tools*, vol. 27, no. 7, pp. 17, 20–26, 2002.
- [7] R. J. M. Silva, A. B. Raposo, and M. Gattass, "[monografias em ciência da computação]grafo de cena e realidade virtual," Master's thesis, Departamento de Informática, PUC, Rio de Janeiro, Abril 2004.
- [8] D. Brutzman and L. Daly, *X3D: extensible 3D graphics for Web authors*, ser. Morgan Kaufmann series in interactive 3D technology. Elsevier/Morgan Kaufmann, 2007.
- [9] Xj3D, "The xj3d project," <http://www.xj3d.org/>, outubro 2015.

Jogo Educacional de Caça ao Tesouro Baseado em Realidade Aumentada

Treasure Hunt Educational Game Based on Augmented Reality

Hytalo Tavares

Departamento de Estatística e Informática
Universidade Federal Rural de Pernambuco
Recife, Brasil
hytalo.augusto@hotmail.com

João Paulo Lima

Departamento de Estatística e Informática
Universidade Federal Rural de Pernambuco
Recife, Brasil
jpsml@deinfo.ufrpe.br

Yvonne Carvalho

Centro de Informática
Universidade Federal de Pernambuco
Recife, Brasil
yccal@cin.ufpe.br

Resumo—O uso de novas tecnologias vinculadas às atividades didáticas dentro do ambiente escolar é uma prática cada vez maior, e a Realidade Aumentada é uma das tecnologias que é defendida por especialistas como uma ferramenta educacional. Nesse contexto, este artigo apresenta um jogo educacional de caça ao tesouro para dispositivos móveis que faz uso da tecnologia de Realidade Aumentada. Um experimento também foi realizado usando o jogo desenvolvido para ensino de história em uma turma do nono ano de uma escola primária privada.

Palavras-chave—realidade aumentada; software educacional; dispositivos móveis; caça ao tesouro

Abstract—The use of new technologies linked to educational activities within the school environment is a growing practice, and Augmented Reality is one such technology that is advocated by experts as an educational tool. In this context, this paper presents a treasure hunt educational game for mobile devices that makes use of Augmented Reality technology. An experiment was also performed by using the developed game for history teaching in a ninth grade class of a private elementary school.

Keywords— augmented reality; educational software; mobile devices; treasure hunt

I. INTRODUÇÃO

Snyders [1] comenta que François Fénelon, educador do século XVIII, cita a escola como um lugar de aborrecimentos, e não de prazeres, e explica que os jovens esperam impacientes o momento que podem sair dela e se dedicar aos jogos, onde encontram prazer.

Desde o surgimento da escola formal, muitas mudanças ocorreram, mas poucas foram capazes de realizar o desejo

expresso por Fénelon no século XVIII e trazer para a escola o prazer que os jovens sentem fora dela, nos jogos.

Professores dos anos finais do Ensino Fundamental II, independentemente de suas áreas, enfrentam dificuldades metodológicas para tornar a sala de aula um espaço de aprendizagens significativas, capaz de despertar a curiosidade e o envolvimento das turmas no desenvolvimento das atividades propostas. Esses profissionais veem-se diante de alunos apáticos, desinteressados, que estabelecem uma relação puramente instrumental com o ensino, tomando-o como uma espécie de “mal necessário” [1].

No ensino de história a estrutura tradicional de ensino e as demandas sociais expressam-se de maneira particular, pois fazem com que os estudantes se deparam com formatos de ensino verticalizados. Assim, transmite-se uma concepção da história como o estudo de um passado sem vínculos com o presente, como um conhecimento pronto e acabado, a ser transmitido de forma fria e mecânica [2]. O reflexo dessa prática é a ausência do gosto pelo estudo, que se revela muitas vezes no baixo engajamento dos estudantes e, por consequência, no baixo rendimento nas avaliações e demais atividades educacionais.

Subverter essa lógica tradicional e tornar as experiências de sala de aula prazerosas e significativas, capazes de despertar a curiosidade dos educandos e de envolver as turmas nas atividades propostas é um desafio para os professores. Os jogos possuem um estímulo para que se possa haver uma melhor compreensão do conteúdo, assim como o crescimento e o desenvolvimento intelectual do aluno, fundamentais para que possa alcançar a responsabilidade e a maturidade. Sendo uma

forma de abordar o conteúdo com os alunos, os motivando a estudar de uma maneira mais atrativa [3].

Através dos sistemas de Realidade Aumentada (RA) é possível a realização de experiências em que ambientes 3D gerados por computador sejam misturados com imagens do mundo real, aumentando as informações do ambiente. A utilização da RA contribui com a dinâmica dos jogos através de um novo ambiente onde os alunos possam interagir com os objetos virtuais que serão dispostos no ambiente real. Este fator possibilita atrair ainda mais a atenção dos alunos, buscando trabalhar de uma maneira melhor o conteúdo abordado.

Nesse contexto, o objetivo do presente trabalho consiste em utilizar a RA como forma de despertar o interesse e aumentar a participação dos educandos no processo de ensino-aprendizagem. Acredita-se que com uso da RA será possível trazer o gosto e o prazer de que fala Fénelon aos estudos, visto que a literatura da área nos fornece suporte teórico neste sentido.

O trabalho trata da utilização da RA vinculada a um jogo de caça ao tesouro para dispositivos móveis, onde podem ser abordados conteúdos de diversas disciplinas no formato de perguntas do tipo múltipla escolha. Após o desenvolvimento do aplicativo, foi realizado um experimento onde o jogo foi utilizado no contexto do ensino de história para estudantes do nono ano de uma escola primária privada.

O presente artigo está organizado da seguinte forma. A Seção II apresenta conceitos sobre RA e elementos relacionados à implementação da tecnologia neste trabalho. A Seção III fala sobre o uso de jogos no processo de ensino-aprendizagem e exemplos de como a RA pode auxiliar na educação. A Seção IV apresenta o jogo de Caça ao Tesouro desenvolvido com o auxílio da tecnologia de RA. Na Seção V são apresentados os resultados obtidos na avaliação com alunos. A Seção VI traz as conclusões e os trabalhos futuros a serem realizados.

II. REALIDADE AUMENTADA

A. Conceitos e Definições

RA é uma variação de ambientes virtuais, ou Realidade Virtual (RV), como é mais comumente chamado. A RV cria um ambiente inteiramente sintético, no qual o usuário é impossibilitado de ver o que se passa em seu entorno, possibilitando assim alto grau de imersão. Em contraste, RA permite ao usuário ver o mundo real, com os objetos virtuais sobrepostos ou compostos com o mundo real. Portanto, RA complementa a realidade, em vez de substituí-la completamente. Assim, expressa a ideia de que os objetos virtuais e reais coexistem no mesmo espaço [3].

Outra definição trazida para RA é a sobreposição de objetos virtuais em um ambiente real, onde esse argumento parte de um conceito maior conhecido como Realidade Misturada. Que podemos definir como uma combinação do ambiente real e virtual, recebendo assim duas denominações: RA quando existe a predominância do ambiente real e Virtualidade Aumentada (VA) quando existe predominância do ambiente virtual [5]. Na

Figura 1 podemos observar a ilustração do diagrama do Contínuo Realidade-Virtualidade.



Fig. 1. Diagrama do Contínuo Realidade-Virtualidade [5].

Uma característica também presente na RA é a possibilidade de interação do usuário com os objetos virtuais, podendo utilizar as próprias mãos e descartando o uso de equipamentos tecnológicos complexos [6].

B. Metaio SDK

Para que um sistema de RA possa funcionar corretamente, se faz necessário a utilização de um software. Dentre os softwares existentes, o Metaio SDK foi o escolhido para a realização deste trabalho, pois um dos objetivos é ser utilizado em dispositivos móveis, como *tablets* e *smartphones*, e o Metaio SDK permite o desenvolvimento de aplicativos de RA para tais plataformas.

O Metaio SDK possui uma estrutura modular que inclui o componente de captura, de interface de sensor, de renderização, de rastreamento e a interface do Metaio SDK. A interface fornece interação entre a aplicação e os outros quatro componentes modulares. Sob essa configuração, os detalhes de implementação são encapsulados e o usuário não precisa se preocupar com a captura, renderização, sensores ou rastreamento. As principais funcionalidades são realizadas por meio de APIs que se comunicam com as outras partes do SDK, o que resulta na fácil implementação de aplicações de RA [7].

III. JOGOS EDUCACIONAIS

Por muito tempo existiu uma discussão sobre como o vídeo game influencia de forma negativa as crianças e adolescentes. Mas nos últimos anos vem aumentando o interesse sobre como os jogos podem influenciar de forma positiva os jogadores e de que forma eles podem ser utilizados na educação. Porém, para que eles possam ser utilizados junto à educação é preciso que eles englobem os conteúdos das disciplinas dos alunos ou promovam o desenvolvimento da capacidade cognitiva e intelectual. Com isso, existe um grande interesse por parte das instituições de ensino e professores em buscar formas de utilizar os jogos para a apoiar a aprendizagem [8].

Existem diversos tipos de jogos educacionais que tratam dos mais diversos tipos de disciplinas e conteúdo, como jogos matemáticos que trabalham as operações básicas [9] (adição, subtração, multiplicação e divisão) e as formas geométricas [10]. Existe também adaptações de jogos já conhecidos, como quebra-cabeças [11] e jogo da memória [12], onde se utiliza da

mecânica desses jogos atrelada ao conteúdo de alguma disciplina.

A RA também vem sendo utilizada no desenvolvimento de atividades educacionais. Como exemplo temos: os livros aumentados, que permitem aos alunos visualizarem projeções de ilustrações no formato 3D; aplicações que podem ser utilizadas na disciplina de história, permitindo a interação dos alunos com ambientes distintos daqueles em que estão presentes, possibilitando um conhecimento mais aprofundado do conteúdo abordado; jogos que contribuem com a aprendizagem da leitura por parte das crianças, permitindo a formação de frases e sentenças; aplicações que tratam dos conteúdos de química, permitindo ao aluno montar e visualizar as moléculas dos elementos da tabela periódica; aplicações que podem ser utilizadas na disciplina de biologia ou até mesmo na universidade, onde se pode visualizar a estrutura interna de um corpo humano e interagir a partir de seus gestos [13].

IV. CAÇA AO TESOURO

A difusão dos dispositivos móveis na sociedade foi o fator que impulsionou sua escolha na hora de desenvolver essa ferramenta voltada para os mesmos. E a utilização destes dispositivos no processo de ensino e aprendizagem possibilita o acesso aos conteúdos educacionais em diferentes lugares e horários [14].

Para o desenvolvimento deste trabalho utilizou-se o Metaio SDK e a plataforma alvo foi o sistema operacional móvel Android. Utilizando da linguagem de programação Java juntamente com a IDE Eclipse.

Para utilizar o jogo, se faz necessário o uso de um dispositivo Android munido de uma câmera, para que assim os alunos possam visualizar os marcadores e as imagens sobrepostas nos mesmos.

Com o aplicativo aberto, basta que a câmera do smartphone visualize o marcador, para que seja exibida na tela uma série de imagens interativas. A primeira imagem mostrada é a de um baú, seguida de quadros contendo perguntas (Fig. 2), e respostas em múltipla escolha (Fig. 3). Além destas, o usuário recebe informações de direcionamento relacionadas aos passos do jogo (expostas em formas de setas) e em caso de erro nas perguntas, ilustrações e textos que servem de dicas para os alunos. Os baús, perguntas, alternativas e demais imagens utilizadas no jogo estão vinculadas a marcadores. A função dos marcadores (Fig. 4) é fazer com que o sistema possa identificá-los e expor sobre eles as respectivas imagens vinculadas aos mesmos.

Quem foi o primeiro presidente do Brasil?

Fig. 2. Modelo de pergunta.

Marechal Deodoro
da Fonseca

Fig. 3. Modelo de resposta.

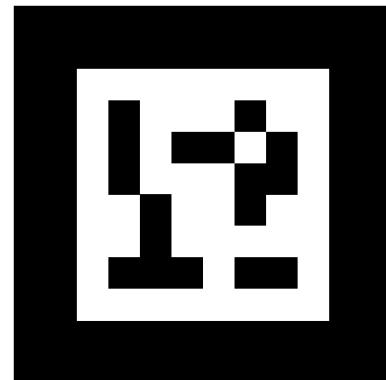


Fig. 4. Exemplo de marcador.

A disposição dos marcadores no ambiente é feita de forma sequencial, onde cada marcador possui um número de identificação. Mas os marcadores não precisam estar alinhados lado a lado, já que se trata de um jogo de caça ao tesouro, onde o propósito é fazer com que os alunos interajam com o ambiente em busca do marcador, sendo que ele será apenas indicado a qual direção seguir, direita ou esquerda.

O jogo pode conter quantas perguntas forem necessárias e essas perguntas e suas alternativas estão representadas através de imagens no formato PNG. A partir do momento em que um aluno direciona sua câmera para um marcador, ele irá visualizar um baú fechado (Fig. 5) e através do toque na tela sobre o baú, lhe será apresentada uma pergunta de múltipla escolha (Fig. 6).



Fig. 5. Modelo de baú fechado.



Fig. 6. Modelo de pergunta de múltipla escolha visualizada por aluno.

Essas perguntas e suas alternativas não seguem uma sequência obrigatória, sendo ordenadas de forma aleatória. As perguntas recebem sua ordem assim que o jogo é executado, e suas alternativas a partir do momento em que um aluno passa de uma pergunta para a outra. Dessa forma é possível que diferentes alunos possam utilizar o jogo, sem que a mesma pergunta seja selecionada para todos os alunos ou que as alternativas estejam sempre na mesma ordem.

A partir do momento em que o aluno responde uma pergunta tocando na alternativa correta, será selecionado aleatoriamente qual o próximo baú a ser encontrado. Ele será guiado até o próximo baú através de setas (Fig. 7), que direcionam o lado que o aluno deve seguir (esquerda ou direita).

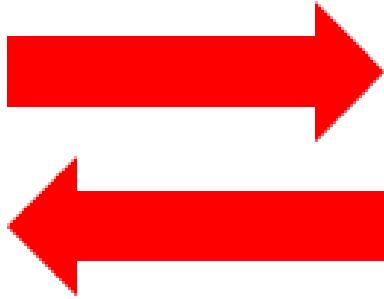


Fig. 7. Modelos de setas para a direita e para a esquerda.

Quando o aluno tiver respondido corretamente todas as perguntas, ele irá visualizar um mapa (Fig. 8), com um modelo de uma planta da escola (que pode ser editado), identificando onde está o marcador com o baú que guarda o tesouro. Ao clicar no mapa, o aluno irá habilitar uma nova imagem para os marcadores: um baú vazio aberto (Fig. 9), fazendo assim uma alusão de que o verdadeiro tesouro ainda não foi encontrado. Os alunos só conseguem visualizar o baú que guarda o tesouro (Fig. 10) após responder corretamente todas as perguntas e encontrar o último marcador, que tem sua localização revelada no mapa.

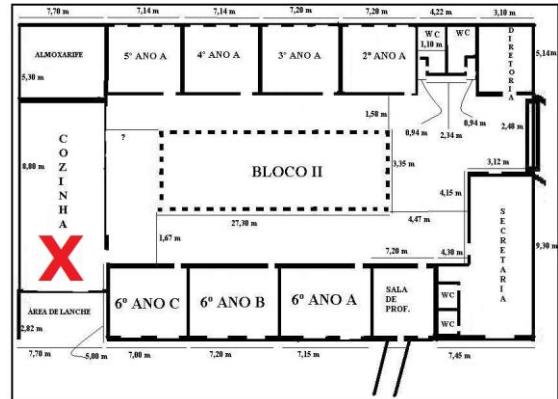


Fig. 8. Modelo de mapa.



Fig. 9. Modelo de baú vazio.



Fig. 10. Modelo de tesouro.

Para tornar o jogo mais versátil e autoral, foi desenvolvida uma aplicação, cujo o objetivo é permitir que o professor possa criar as imagens que irão conter as perguntas e respostas automaticamente. Ao iniciar a aplicação o professor deverá informar quantas perguntas deseja inserir no jogo, digitar as perguntas, as alternativas corretas e incorretas e selecionar uma imagem como dica. Isso irá se repetir até que ele tenha fornecido a quantidade de perguntas que informou no início da aplicação. Em seguida, todos esses dados serão salvos em uma pasta no computador, sendo necessário apenas a transferência desta pasta para o dispositivo onde está instalado o aplicativo. Desta forma pretende-se tornar o jogo uma ferramenta mais prática e mais acessível às diversas áreas.

Espera-se que com este jogo os conteúdos trabalhados pelos professores passam ser tratados de forma mais lúdica.

V. RESULTADOS

O jogo foi apresentado a uma turma de 12 alunos de uma escola da rede privada que cursam o nono ano do Ensino Fundamental II (antiga 8^a série), na faixa dos catorze e quinze anos de idade. A proposta pedagógica da escola é baseada no Interacionismo Sócio Discursivo de Jean-Paul Bronckart [15] e que tem na coordenação e direção escolar abertura para experimentação e inovações metodológicas que visem a melhoria do processo de ensino-aprendizagem.

As perguntas utilizadas na vivência foram previamente elaboradas pela professora da turma, sobre a Primeira Guerra Mundial, assunto que foi visto em sala de aula no mesmo bimestre. Assim, por parte da professora, a atividade teve também como objetivo fazer com que os alunos testassem seus aprendizados vivenciando os alunos de forma lúdica.

A vivência iniciou-se com um momento de preparação. Os marcadores foram espalhados no espaço aberto da escola e o aplicativo foi compartilhado via *bluetooth* com os alunos (Fig. 11). Depois que todos os estudantes instalaram o aplicativo em seus dispositivos móveis, iniciou-se a atividade da Caça ao Tesouro.



Fig. 11. Compartilhamento e instalação do aplicativo com os alunos.

Com relação à dinâmica e à percepção dos usuários de como lidar com o jogo, o mesmo se demonstrou intuitivo, não havendo necessidade de orientações externas sobre o uso do aplicativo. Até mesmo os alunos que desconheciam RA e que nunca haviam utilizado a tecnologia antes, ao abrirem o dispositivo e verem marcadores colados nas paredes da escola perceberam como a atividade deveria funcionar.

Houve, no entanto, algumas dificuldades por parte dos participantes. A primeira dificuldade enfrentada foi referente às setas de orientação espacial utilizadas no jogo (o que pode ser atribuído à utilização do mesmo em um ambiente aberto). A segunda dificuldade encontrada foi a visualização das perguntas e dicas por parte dos usuários que utilizaram telefones com telas menores, visto que a visualização das imagens das dicas foi prejudicada por esta razão. Destaca-se,

porém, que essas dificuldades não impediram que todos conseguissem seguir jogando.

O uso por parte dos alunos foi marcado por muito entusiasmo, curiosidade, interação e dinamismo (Fig. 12). Os alunos destacaram, por meio de entrevistas e observações (Fig. 13), que apreciaram a atividade, repetiriam a mesma e destacaram o fator diversão como o mais marcante do jogo.



Fig. 12. Aluno em meio à correria da vivência da Caça ao tesouro.

“Muito massa essa aula. Todos os professores deviam fazer isso”.
(Fala espontânea de aluna no encerramento da atividade)

Fig. 13. Fala feita por aluna, percebida por meio de observação.

Por fim, o que se pôde constatar é que o jogo consegue despertar nas pessoas imersão, derivada da interação com o aplicativo de RA e com o caráter competitivo e dinâmico da mesma. Destaca-se também que, mesmo com a euforia da disputa, manter a atenção na leitura das perguntas e respostas foi essencial para que os alunos encontrassem a resposta correta.

VI. CONCLUSÃO E TRABALHOS FUTUROS

A partir da utilização do jogo em um ambiente escolar e tratando de um assunto curricular dos alunos, pôde-se concluir que o trabalho alcançou seus objetivos para uma primeira versão e que, apesar de algumas dificuldades encontradas pelos alunos, todos ficaram muito animados com a atividade e na expectativa de a realizarem novamente.

Outro ponto a ser destacado é a potencialidade do jogo para ser utilizado também em atividades extraescolares, em dinâmicas em grupo e até mesmo em atividades individuais.

Com relação a RA, evidenciou-se que atividades utilizando tecnologias tendem a atrair ainda mais a atenção das pessoas. E nesse caso a RA se destaca, pois o uso de imagens virtuais interativas registradas com o mundo real dos usuários tende a ser efetivo no requisito de cativar atenção.

Como trabalhos futuros, o primeiro passo é corrigir os problemas que foram encontrados durante a avaliação da

aplicação, para que novos testes possam ser feitos e verificar como será a recepção dos alunos com relação à nova versão do jogo. Outras perspectivas estão relacionadas ao aprimoramento da flexibilidade de criação de conteúdos para o jogo, pois é importante que ele possa permitir aos professores uma maior facilidade nesse ponto, podendo, inclusive, aumentar a quantidade de materiais e conteúdos abordados. E para um futuro mais distante, estuda-se a possibilidade de desenvolver uma ferramenta de autoria de jogos educacionais com RA, permitindo que o professor crie aplicações de acordo com suas necessidades. Ou seja, permitir que o professor possa gerar seu próprio jogo educacional com RA, seja ele um caça ao tesouro ou alguma outra atividade.

O propósito de um sistema autoral passa por algumas fases. Onde o primeiro passo será aprimorar a criação de conteúdos para o caça ao tesouro por parte dos próprios professores, ou seja, eles serão os responsáveis por criar as imagens utilizadas como perguntas e respostas, de forma que eles não precisaram utilizar um editor de imagens para isso, apenas informar os textos e identificar a alternativa correta. Permitir diferentes tipos de *layouts* para essas imagens à serem geradas, assim como aceitar imagens 3D e vídeos a serem utilizados como dicas. Os passos seguintes passam por adicionar novos tipos de jogos e atividades a serem realizadas com RA, mantendo a proposta de que serão os próprios professores a criar os materiais a serem utilizados na aplicação, ampliando, possivelmente, para que seja também permitido a criação de imagens em 3D no próprio *software*. E com um ponto de vista mais longínquo, permitir que os professores possam desenvolver seus próprios aplicativos de RA, sem que seja necessário nenhum tipo de conhecimento prévio em programação ou a intervenção de algum programador.

A realização de uma avaliação quantitativa mais formal, detalhada e criteriosa junto aos usuários é algo que será utilizada nas próximas fases do projeto. Onde deve ser feita uma avaliação com os professores envolvidos na atividade, para verificar como eles veem a utilização da ferramenta [16]. A partir deste tipo de avaliação será possível verificar como e o quanto o Caça ao Tesouro pode contribuir no processo de ensino e aprendizagem.

REFERÊNCIAS

- [1] G. Snyders, Alunos Felizes: Reflexão sobre a Alegria na Escola a partir de Textos Literários. Rio de Janeiro: Paz e Terra, 1993.
- [2] D. Saviani, Pedagogia Histórico-Crítica: Primeiras Aproximações. São Paulo: Cortez: Autores associados, (Coleção polêmicas do nosso tempo; vol. 40), 1991.
- [3] Verri, Juliana Bertolino, and Angela Endlich. "A utilização de jogos aplicados no ensino de Geografia." Revista Percurso 1.1 (2009): 65-83.
- [4] Azuma, R. T. (1997). A survey of augmented reality. *Presence*, 6(4), 355-385.
- [5] Zorral, E. R., BUCCIOLI, A. A. B., & Kirner, C. (2006). Usando Realidade Aumentada no Desenvolvimento de Quebra-cabeças Educacionais. In SVR2006-VIII Symposium on Virtual Reality (pp. 221-232).
- [6] Zorral, E. R., Cardoso, A., Kirner, C., & Júnior, E. L. (2006). Realidade aumentada aplicada em jogos educacionais. In V Workshop de Educação em Computação e Informática do Estado de Minas Gerais-WEIMIG.
- [7] Metaio. Metaio SDK Framework. Disponível em: <http://dev.metaio.com/sdk/documentation/metaio-sdk-framework/index.html>. Acesso em: 19 de setembro de 2015.
- [8] Savi, R., & Ulbricht, V. R. (2008). Jogos digitais educacionais: benefícios e desafios. *RENOTE*, 6(1).
- [9] Racha Cuca. Calculando. Disponível em: <http://rachacuca.com.br/jogos/calculando/>. Acesso em: 27 de maio de 2015.
- [10] Escola Games. Formas Geométricas. Disponível em: <http://www.escolagames.com.br/jogos/formasGeometricas/>. Acesso em: 27 de maio de 2015.
- [11] Racha Cuca. Quebra-Cabeça. Disponível em: <http://rachacuca.com.br/raciocinio/quebra-cabeca/>. Acesso em: 27 de maio de 2015.
- [12] Matematicazup. Memória geométrica. Disponível em: <http://matematicazup.com.br/jogos-de-matematica-memoria-geometrica/>. Acesso em: 27 de maio de 2015.
- [13] da Silva, M. M. O., Roberto, R. A., & Teichrieb, V. (2012). Um estudo de aplicações de realidade aumentada para educação.
- [14] N. Paviani, "Aprendizagem na perspectiva da teoria do interacionismo sociodiscursivo de Bronckart," *Revista Espaço Pedagógico*, vol. 18, n. 1, pp. 58–73, Janeiro/Junho 2011.
- [15] de Oliveira, Leandro Ramos, and Roseclea Duarte Medina. "Desenvolvimento de objetos de aprendizagem para dispositivos móveis: uma nova abordagem que contribui para a educação." *RENOTE* 5.1 (2007).
- [16] Silva, M., Roberto, R., & Teichrieb, V. (2013). Evaluating an Educational System Based on Projective Augmented Reality. In Anais do Simpósio Brasileiro de Informática na Educação (Vol. 24, No. 1).

Face Detection:

New methods based on skin detection

Marcelo Urbano Pereira

Centro de Matemática, Computação e Cognição
Universidade Federal do ABC
Santo André, Brasil.
marcelo.urbano@ufabc.edu.br

Fernando Teubl Ferreira

Centro de Matemática, Computação e Cognição
Universidade Federal do ABC
Santo André, Brasil.
fernando.teubl @ufabc.edu.br

Abstract— We live in a world in which the digital information is present around us, tasks such as human identification in scanned images, for a variety of purposes. In daily life we are surrounded by appliances that use of facial recognition for various purposes, such as in the area of biometrics, surveillance, entertainment, among others. Recognize the face is becoming an important task in many applications. In this paper is proposed the contextualization of the state of art of some recent methods for face detection based on detection of skin, revealing the methods which became consolidated concepts for identification of face and new approaches of recent methods, is also described the advantages and limitations of each method.

Keywords—component; face detection; Adaboost; Haar-Like Features; color skin detection.

I. INTRODUÇÃO

O reconhecimento facial é um tópico importante da visão computacional e é muito utilizado nos dias atuais em muitos dispositivos eletrônicos que fazem parte do cotidiano moderno (como câmeras fotográficas, smartphones, videogames, entre outros).

Uma das etapas fundamentais é a detecção de faces, utilizado principalmente em sistemas de reconhecimento das faces em imagens ou vídeos, utilizado em aplicativos tal como de vigilância, em que os vídeos podem identificar pessoas em situações suspeitas. Identificar com exatidão as possíveis faces em uma imagem tem sido foco de diversos pesquisadores, e muitos métodos foram propostos para determinar com mais precisão a área que contém a informação do rosto.

É difícil para um computador segmentar a face de uma imagem. Zhang [1] classifica os métodos de detecção em quatro categorias: (a) o método baseado no conhecimento humano para definir regras de detecção de faces; (b) o método que analisa as características invariantes tais como pose e iluminação; (c) o método de correspondência de modelos de faces pré-armazenados e (d) o método baseado na aprendizagem da aparência com base em treinos em vários modelos de faces, sendo este muito eficaz e factível de ser implementado devido ao avanço do poder de processamento e armazenamento computacional atual.

Alguns métodos bem sucedidos buscam identificar as faces em imagens digitais a partir da cor da pele humana. A proposta deste artigo é de abordar o estado da arte dos métodos recentes de detecção de face baseados em detecção de pele. A

continuidade do artigo esta organizado da seguinte forma. A seção 2 mostra o uso de métodos que se tornaram conceitos consolidados para identificação de face, e que são usados nas abordagens dos cinco métodos recentes descritos na seção 3. A seção 4 relata as vantagens e limitações dos métodos abordado na seção anterior. A seção 5 conclui o artigo.

II. MÉTODOS DE DETECÇÃO DE FACES

Os principais métodos pesquisados identificam, inicialmente, as regiões das imagens com alta probabilidade de conter pele. Em sequencia, estas regiões de pele são analisadas para identificar as regiões da face. Segundo Li [2] a detecção de faces refere ao processo de definição das características da face como a posição, tamanho e formato da face de uma imagem de entrada. Nos últimos anos apareceram muitos métodos de detecção de face. Um dos métodos recentes propostos é o Adaboost, proposto por Viola [3], que foi baseado no Haar-like Feature (HLF).

A ideia principal do Adaboost é treinar classificadores fracos de um conjunto de treinamento para obter alguns classificadores fracos adequados, que serão candidatos para tornarem-se classificadores fortes. A classificação é referenciada por pesos, que é calculado pela classificação de acertos em uma classificação geral. Em seguida, são treinados os novos conjuntos de dados modificados com a próxima camada, que por sua vez realizará novas classificações. Finalmente, agrupam-se os melhores classificadores de cada camada para o classificador de decisão final. O classificador Adaboost elimina alguns recursos desnecessários e se focar no treinamento em si, que é a chave para o processo de detecção da face.

O HLF é proposto no trabalho de Viola [3] e calcula a característica da face em uma imagem, características definidas e configuradas em seu algoritmo, que pode aumentar a eficiência computacional na detecção de faces empregado com outros métodos (como o Adaboost). No entanto, o HLF é um recurso grosso, é sensível à borda, linha e só pode descrever uma estrutura específica. Também é necessário muito tempo para essa detecção das características.

A lógica Fuzzy é outra técnica empregada na busca de extração de informações de regiões. Foi introduzida com sucesso em diversos campos. Chowdhury [4] descreve sistemas de lógica Fuzzy tais como: sistema de processo automático de controle, diferentes tipos de classificação de dados, sistemas

especialistas, análise de decisão e visão computacional. As regras Fuzzy são caracterizadas por uma coleção de diferentes funções e regras de operações lógicas.

Baseada nas técnicas apresentadas, serão analisados alguns métodos de detecção de imagens em que o elemento principal a se definir são as áreas de faces com base na detecção da pele.

III. MÉTODOS RECENTES DE DETECÇÃO DE FACES BASEADOS EM DETECÇÃO DE PELE

A. Detecção Multi View da Face com base em Transformação da Característica

Han [5] propõe trabalhar com o HLF juntamente com o Adaboost, e visa aprimorar o conceito de ensinar e treinar classificadores para construir o melhor classificador fraco de um grande número de diversas faces.

Características como posições e tamanhos dos olhos de cada face vertical são diferentes, mas isso é solucionado com a escolha de um tipo de HLF específico que cobri a região sobre uma pequena variação da face.

A robustez do HLC pode ser utilizada para lidar com variações de faces rotacionadas na fase de treinamento. Um conjunto de treinamento original consiste de várias faces, mas é preciso aumentar o conjunto de treinamento das características obtidas do HLC girando-os para um conjunto de treinamento expandido.

As faces rotacionadas no plano contribuíram para detectar aspectos de faces semelhantes na fase de teste, mostrada na fig. 1.

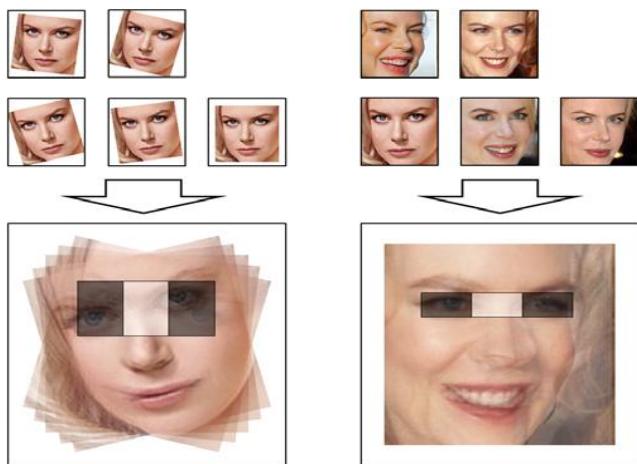


Fig. 1: Fase de treinamento de faces rotacionadas [5].

Faces rotacionadas fora do plano não podem ser facilmente criadas por transformações rígidas como de rotação, então é necessário encontrar faces em diferentes pontos de vista correspondendo à rotação fora do plano para construir o conjunto de treinamento original adequado. O Adaboost não é eficiente em classificar amostras como positivas e negativas. Contudo, Han [5] demonstra que taxa de convergência do Adaboost cai proporcionalmente dependendo do grau do ângulo de rotação ao longo da execução, evidenciado na fig. 2.

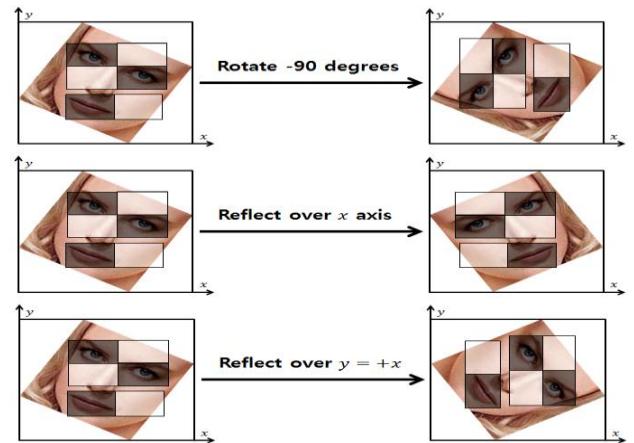


Fig. 2. Detecção de características com variação dos ângulos de rotação [5].

A proposta do Método de Transformação de Características (*Feature Transform Method*) é obter classificadores fortes de classificadores fortes já existentes. (gerados pelo Adaboost) para detectar a face.

Cada classificador forte gerado pelo método de transformação de característica tem uma estrutura em cascata constituindo-se de classificadores fracos. Todos os classificadores fracos podem ser combinados para gerar um único classificador forte unificado, o que otimiza o processo de detecção.

B. Detecção de Face rápida com base na cor da pele

Em Hu [6], é abordado um método simples e rápido de detecção de rosto com base na cor de pele. É proposto e aplicado corte e separação das regiões da pele e da região da não pele. Em seguida, as regiões escolhidas são extraídas da região de pele a partir de uma decisão de ser ou não regiões de face, detectadas pelos métodos de detecção Adaboost (denominado de VJ detector por Hu[6] em seu artigo), e criado por Viola [7].

Para extrair todos os segmentos das regiões em que há pele humana na região caracterizada como de não pele, o modelo de segmento de cor da pele deve ser confiável e eficaz para adequar-se não apenas as diferentes cores de pele, mas também imagens com diferentes condições de iluminação. A fig. 3 evidencia os diferentes tipos de amostras.



Fig. 3. Diferentes amostras de tom de pele para detecção [6].

A distribuição dos diferentes tons de pele das pessoas é o primeiro item a ser analisado. A abordagem de Hu [6] utiliza o

espaço de cores YCbCr, o que inclui o componente de luminância (Y) e os componentes de cor cromáticas ou puras (Cb e Cr). O espaço YCbCr pode ser facilmente obtido a partir do espaço RGB. Classificam-se então diversas amostras de pele sem considerar as diferentes raças e tonalidades de pele. Em seguida, criam-se histogramas do Cb e Cr das amostras a partir das indicações das curvas de distribuição normal (de um modelo Gaussiano) que são avaliadas de acordo com intervalo de confiança de 99%.

A partir de valores dos intervalos de Cb e Cr são detectadas as ROIs (*Region of Interest* – Região de Interesse), para minimizar pseudo ROIs de tamanho pequeno que não possuem faces na imagem. A imagem é convertida para uma representação binária e aplicam-se operações morfológicas (ex.: dilatação e erosão) para remover aglomerados indesejados de tamanho menor. As áreas que potencialmente contém pele são segmentadas da área de não pele. Por fim, a posição do centro de cada ROI (com a área de pele) é submetida ao detector de face, e o centro da face é detectada (fig. 4).

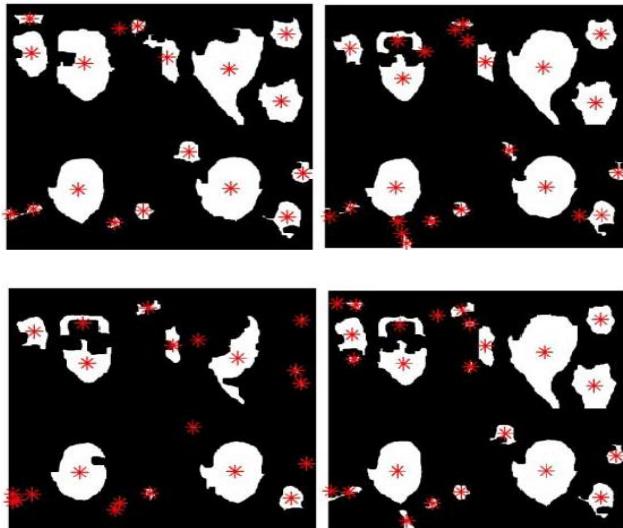


Fig. 4. Resultado da detecção e centralização das ROIs adequados e eliminação das possíveis não faces [6].

C. Detecção de pele humana e Reconhecimento Facial usando Lógica Fuzzy e Eigenface

Chowdhury [4] propõe um sistema de detecção de pele com base na lógica Fuzzy para detectar a presença ou a ausência de ser um humano em uma determinada matriz denominada *eigenvector*. O sistema localiza as áreas de pele humana na imagem no espaço de cor HSV. Chowdhury [4] ainda comenta que a cor da pele humana é totalmente diferente da cor de outros objetos e, por esta razão, as informações desta área são muito importantes para a detecção de face, reconhecimento de gestos e de identificação pessoal.

Um *eigenvector* de uma matriz quadrada A é um vetor v não zero que, ao multiplicar a matriz por v , produz uma constante múltipla de v . O multiplicando é denotado por λ , e é chamado de *eigenvalue* de A correspondendo a v . Chowdhury [4] apud Turk [8] descreve *eigenfaces* como o nome de um conjunto de *eigenvectors* quando eles são usados na visão

computacional no reconhecimento de rosto humano para a identificação humana.

Na abordagem do reconhecimento de face baseada na lógica fuzzy e *eigenface* primeiramente adquiri um conjunto inicial de imagens faciais (que é usado como conjunto de treinamento). Calcula-se a média de todas as imagens, assim como o *eigenvector*, o *eigenvalue* e o *eigenface*. Calcula-se uma assinatura para cada imagem, e subtraí o valor da média de um índice de imagem que foram adquiridas randomicamente. Ao comparar o valor da assinatura com o valor subtraído, a face ou faces podem ser reconhecidas se estiverem presentes na imagem, como apresentado na fig. 5.



Fig. 5. Imagem original (esquerda) e imagem pós detecção de pelo usando lógica Fuzzy. (direita) [4].

D. Método de detecção de Face baseado em cor de pele e bordas

No trabalho de Ghimire [9] é proposta uma abordagem do uso de características que não variam, com base na cor da pele e na informação da borda para detecção da face. O método consiste primeiramente em trabalhar a imagem de entrada para lidar especialmente com imagens sem restringir com condição de iluminação (em que a pele é afetada pela condição da luz). É necessária a conversão da imagem do espaço de cor RGB para o espaço de cor HSV.

A melhora da imagem é realizada através de uma função de transferência não linear baseada de abordagem local do melhoramento da cor, o que cria um ambiente de iluminação uniforme na imagem no canal V (valor) do HSV. Então é aplicado melhoramento de luminância por transformação não linear em certos pixels e melhoramento o contraste por convolução Gaussiana (fig. 6).



Fig. 6. Imagem modificada com melhora na cor e iluminação no espaço de cor HSV [9].

Após o processo de adaptação da luminância no canal V, os outros canais H (matiz) e S (saturação) originais são convertidos novamente para o RGB.

Neste momento, dois processos distintos e simultâneos são realizados: (1) conversão de cor para a escala de cinza e um algoritmo de detecção de borda (usado o *Canny Edge Detection*, mostrado na fig. 7) e, (2) segmentação de cor do espaço RGB para os componentes Cb e Cr (do espaço de cor YCbCr) e remoção de ruído usando o índice de percentagem do tom de pele.



Fig. 7. Imagem resultante do Canny Edge Detection e aplicada transformações de dilatação e erosão.

As imagens resultantes dos processos distintos são combinadas para ressaltar e agrupar os componentes que assemelham-se com a cor da pele, utilizado no processo de classificação de faces e não faces. A classificação tem como base a estrutura geométrica e características das formas para definir uma não face.

A anatomia da face sugere a taxa de limite da região a ser destacada, no qual será analisada a partir de uma janela que tem o tamanho 20% menor em escala que a região destacada da possível face. Se o centro da janela estiver fora da região significa que essa região não é considerada uma face. Para finalizar o processo de detecção, utiliza-se a técnica de cascata e, então, é repetida para classificar todas as regiões como face ou não face, o que possibilita a detecção pela cor da pele. O resultado é apresentado na fig. 8.



Fig. 8. Imagem original (esquerda) e imagem depois da detecção de face por cor de pele e bordas (esquerda).

E. Detecção de rosto com base na cor de estimativa probabilística

A abordagem proposta por Azad [10] tem como base a técnica de estimativa probabilística. Primeiramente é estimada a distribuição da intensidade de pele, o que utiliza um conjunto de treino de imagens com fotos de pele pura para distinguir cores de pele e cores de não pele na imagem. Em seguida, é feito a média a partir de um desvio padrão das cores. As informações são computadas para cada canal de cor RGB. Em seguida, utiliza-se o modelo Gaussiano em cada um dos canais (computando a distinção entre cada canal RGB e a distribuição de pele).

Para aumentar a qualidade da detecção de pele, Azad [10] propõem realizar uma fase de pré-processamento a partir de técnicas de *Histogram Equalization*. Em seguida, é feito o ajuste do *threshold* da imagem para detectar a regiões de pele. Obtida essas regiões, são realizadas operações morfológicas de dilatação para remover ruído, e erosão para remover buracos. Assim, são obtidas as características faciais, tais como olhos, orelha, boca de uma imagem binária.

É fundamental o ajuste do *threshold* para áreas escuras, como a área da boca.

Por fim, para detectar a face, é calculado um triângulo com dois vértices na estimativa dos dois olhos e o terceiro vértice na estimativa da boca (fig. 9).

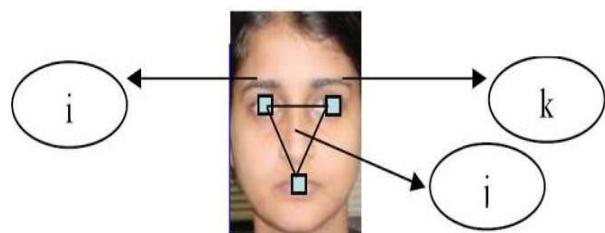


Fig. 9. Aplicação dos pontos de base para olhos e boca para distinção das faces e não faces [10].

Para calcular a distância entre os pontos determinado no centro do olho direito/esquerdo e na boca de 90 - 110% de uma

distância Euclidiana entre os pontos e o triângulo desenhado. Neste momento, a face é detectada, como mostrado na fig. 10.

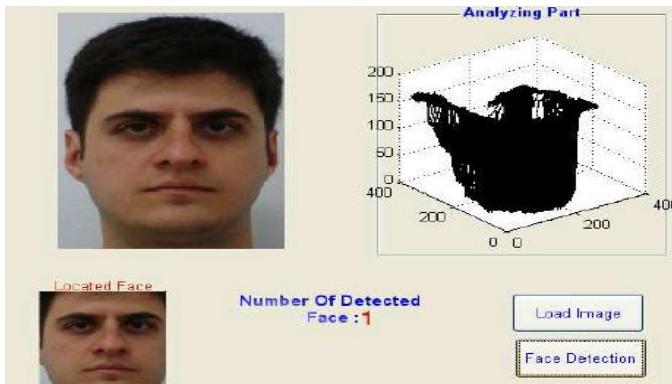


Fig. 10. Resultado da detecção da Face no Sistema desenvolvido por Azad [10].

IV. VANTAGENS/LIMITAÇÕES DOS MÉTODOS RECENTES

Na abordagem de Han [5] sobre a perspectiva da detecção *Multi View* da Face, tem se observado a vantagem do tempo de treinamento que é drasticamente reduzida na fase de formação de Adaboost. Com este método de transformação, apenas dois classificadores fortes da vertical são necessários, e um classificador forte da face rotacionada de pequenos graus na rotação para transformar e gerar novos classificadores fortes muito rapidamente.

O método de Hu [6] se mostrou muito eficiente para reduzir a quantidade de tempo que o Adaboost usa para determinar se a região analisada é uma face. Contudo, o método demonstra gastar mais tempo na obtenção do centro da região de interesse que outros métodos.

A combinação da lógica Fuzzy e *eigenface* de Chowdhury [4] aumenta a precisão do sistema de identificação de faces, mas não demonstrou eficiência quando a imagem possui muita ou pouca condição de iluminação. O método falha na identificação da variação de cor da pele (características étnicas).

A abordagem de Ghimire [9] não será superior que outros métodos em imagens com boas condições de luz. Ele se destaca quando existe uma variação muito extrema na iluminação (pouca ou muita luz).

A adaptabilidade na abordagem de Azad [10] demonstra identificação na maioria dos tipos de pele. Para isso, é necessário estágios de treino, baixa sensibilidade à iluminação e ajustando do *threshold* para ter um rendimento razoável.

V. CONCLUSÃO

Este artigo abordou métodos recentes de detecção de face em imagens, em que a cor da pele é a informação principal para identificar a presença de faces em uma imagem. Foi analisada a especificidade dos métodos abordados.

Nota-se que os métodos de detecção por cor de pele são bons e rápidos para algumas ocasiões, pois a informação da cor

da pele pode ser facilmente rastreada e extraída se os métodos de treinos forem bem definidos. Porém, observa-se que em situações de variação extrema de iluminação na imagem (muito escura ou clara), alguns algoritmos não conseguem ser eficientes, e outros não conseguem ter bom desempenho quando a imagem apresenta boa iluminação. É observado que, nos métodos analisados neste artigo, não se têm uma abordagem que tenha alto desempenho em todas as condições descritas.

Por fim, novos métodos para detecção de face são desenvolvidos e aprimorados muito rapidamente. A evolução crescente no processo de detecção de faces culmina no surgimento de novos sistemas de reconhecimento facial mais preciso e de alto desempenho.

REFERENCES

- [1] C. Zhang and Z. Zhang. A survey of recent advances in face detection. Technical report, Microsoft Research, 2010.
- [2] R. Li, C. F. Li, "Adaboost face detection based on improved covariance feature," Journal of Computers, vol. 9, no. 5, p. 1077–1082, 2014.
- [3] P. Viola, M. Jones, "Rapid object detection using a boosted cascade of simple feature," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, p. 511-518, 2002.
- [4] A. Chowdhury, S.S. Tripathy, "Human Skin Detection and Face Recognition using Fuzzy Logic and Eigenface" in International Conference on Green Computing Communication and Electrical Engineering (ICGCCEE), p. 1 – 4, 2014.
- [5] D. Han, Jiwhan Kim, J. Ju, I. Lee, J. Cha and Junmo Kim, "Efficient and fast multi-view face detection based on feature transformation," in 16th International Conference on Advanced Communication Technology (ICACT), p. 682 – 686, Pyeongchang, 2014
- [6] X. Hu, S. Peng, J. Yan, N. Zhang, "Fast face detection based on skin color segmentation using single chrominance Cr" in 7th International Congress on Image and Signal Processing (CISP), p. 687 – 692, 2014
- [7] P. Viola, M.J. Jones, "Robust real-time face detection," International Journal of Computer Vision, vol. 57(2), p. 137-154, 2004.
- [8] Turk, M.A. and Pentland, A.P., "Face recognition using eigenfaces", Computer Vision and Pattern Recognition, 1991. Proceedings CVPR '91. IEEE Computer Society Conference, Volume 2, p. 586-591.
- [9] D. Ghimire, J. Lee, "A robust face detection method based on skin color and edges," Journal of Information Processing Systems, vol. 9, no. 1, p. 141–156, 2013.
- [10] R. Azad, F. Davami, "A Robust and Adaptable Method for Face Detection Based on Color Probabilistic Estimation Technique". International Journal of Research in Computer Science, 3 (6): p. 1-7, 2013

Parallel Rendering:

Methods Based on Sort-First and Sort-Last

Marcelo Urbano Pereira

Centro de Matemática, Computação e Cognição
Universidade Federal do ABC
Santo André, Brasil.
marcelo.urbano@ufabc.edu.br

Fernando Teubl Ferreira

Centro de Matemática, Computação e Cognição
Universidade Federal do ABC
Santo André, Brasil.
fernando.teubl @ufabc.edu.br

Abstract— Computer Graphics computing is an area that is in constantly evolution. Traditionally, the use of multicore graphics processors are expensive and limited to large research centers. The parallel rendering technique based on PC cluster, using intercommunication between graphics nodes to render models creates scalable distributed rendering system with high performance. The proposal of this article is to study parallel rendering methods such as Sort-first and Sort-Last. This article will present a comparative approach of these techniques.

Keywords—component; Graphic Computing; Parallel Rendering; Sort-First; Sort-Last.

I. INTRODUÇÃO

A Computação Gráfica é uma área da Computação que está em constante evolução. Por mais rápido que seja a evolução do hardware, os aplicativos gráficos sempre necessitarão de hardwares mais poderosos para prover uma imagem com melhor qualidade. Esse crescimento se deve pelas demandas das aplicações de Realidade Virtual, Simuladores, Visualização científica, Jogos, entre outros.

Tradicionalmente, visualizar modelos gráficos complexos exige hardwares caros, o que restringe tais práticas em poucos centros de pesquisas. Por outro lado, o desempenho gráfico e processamento dos computadores pessoais (PCs) estão em constante evolução [1].

As técnicas de renderização paralela [2] baseada em aglomerados de PCs conectados por uma rede de intercomunicação podem trabalhar de modo cooperativo com o objetivo de processar modelos complexos de forma distribuída. Esta prática pode proporcionar alta capacidade de processamento gráfico com um custo reduzido se comparar com sistemas gráficos centralizados, conforme abordado por Molnar [3].

A proposta deste artigo é apresentar um estudo sobre os métodos de renderização paralela tipo *Sort-first* e *Sort-Last*. O artigo esta organizado da seguinte forma: a Seção 2 mostra os conceitos consolidados sobre a renderização paralela que guiaram este trabalho. A seção 3 detalha a arquitetura do método *Sort-First* e apresenta alguns trabalhos relacionados. A seção 4 detalha o método *Sort-Last*, e também apresenta trabalhos relacionados. A seção 5 realiza consideração sobre as duas técnicas, apresentando uma abordagem comparativa e um

algoritmo híbrido com a utilização dos métodos juntos. Por fim, a seção 6 conclui este artigo.

II. RENDERIZAÇÃO PARALELA

A renderização gráfica é um trabalho computacionalmente intensivo, e tem como principal objetivo gerar imagens virtuais de alta qualidade em um curto intervalo do tempo, o que possibilita aplicações interativas e de tempo real. Molnar [1] mostra que as demandas de aplicações gráficas são amplas, tais como: visualização científica, CAD, Simulação de Veículos e Realidade Virtual. Tais aplicativos podem exigir centenas de Mflops de desempenho e gigabytes por segundo de largura de banda de memória, muito além da capacidade de uma simples unidade de processamento.

O paralelismo computacional têm se mostrado uma tática importante para a construção de sistemas gráficos de alto desempenho. O paralelismo pode ser visto de diversos modos. Molnar [1] cita exemplos de paralelismo: funcional (*pipelining*), que pode executar rapidamente cálculos críticos; paralelismo de dados, que permitir resultados para serem computados de uma vez; e paralelismo de objeto e por pixel ou porção da tela (paralelismo de pixel ou imagem).

Molnar [1] ainda demonstra de maneira simplificada o uso de *pipeline* para renderização gráfica, que consiste em duas partes principais: o processamento geométrico (transformações, *clippings*, luz, dentre outros) e a rasterização (*scan-conversion*, *shading*, e visibilidade). Este *pipeline* é demonstrado na fig. 1.

O processamento geométrico pode ser paralelizado se atribuir a cada processador um subconjunto de objetos na cena, conhecido como primitivas. Na rasterização, o paralelismo pode ser obtido ao atribuir para cada processador uma porção dos cálculos de pixel, como apresentado por Molnar [1].

A tarefa da renderização é calcular o efeito de cada primitiva em cada pixel. Devida a natureza arbitrária da modelagem e das transformações de visões, uma primitiva pode cair em qualquer lugar, dentro ou fora da tela.

Molnar et al. [1] apresenta um sistema de classificação baseado no ponto do *pipeline* de renderização no qual é realizado o mapeamento entre o espaço do objeto e o espaço da imagem. Podemos entender a renderização como um problema de primitivas de *sorting* para a tela.

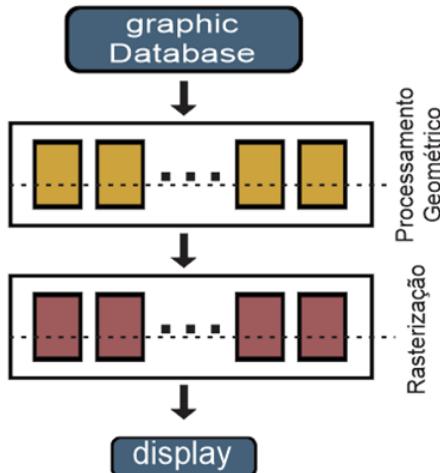


Fig. 1. *Pipeline* gráfico em um sistema completamente de renderização paralela, figura criada com base em Molnar [1].

Para sistemas que utilizam renderizadores completamente paralelos, o *sort* envolve a redistribuição de dados entre processadores. A responsabilidade para primitivas e pixels é distribuída. A localização deste *sort* determina a estrutura resultante no sistema de renderização paralela.

Em geral, o *sort* pode ser aplicado em qualquer parte do *pipeline* de renderização: pode ser aplicado durante o processamento geométrico (*Sort-First*[1,4]); entre o processamento geométrico e a rasterização (*Sort-Middle*[1]), que não será abordado neste trabalho; ou durante a rasterização (*Sort-Last*[1,5]). O *Sort-First* diz respeito à redistribuição das primitivas antes dos seus parâmetros de espaço de tela ser conhecidos. Já no *Sort-Last* a redistribuição se refere aos pixels, amostras e fragmentos de pixels. Nas próximas seções iremos detalhar dois métodos mais utilizados em sistemas de renderização paralelo o *Sort-First* e *Sort-Last*.

III. SORT-FIRST

O *Sort-First* tem o papel de distribuir primitivas no início do *pipeline* da renderização, durante o processamento geométrico, para os processadores poderem fazer os cálculos de renderização restantes, mostrada na fig. 2.

Molnar [1] apresenta a arquitetura do *Sort-First* que consiste em dividir a tela em regiões disjuntas de modo que cada processador (renderizador) se torna responsável por todos os cálculos de renderização de suas respectivas regiões.

Na técnica de *Sort-First*, pode-se atribuir primitivas para os renderizadores de forma aleatória: cada renderizador faz transformações suficientes para determinar em qual região ou regiões cada primitiva cairá. Geralmente é computado um *bounding box* (estrutura delimitadora) do espaço de tela da primitiva. Este processo é chamado de pré-transformação e pode ou não envolver especificamente transformações na primitiva.

Em alguns casos, a primitiva pode cair na região da tela de renderizadores. Em outros casos, podem situar-se na própria região em que reside. Deste modo, é necessário que a imagem seja redistribuída por uma rede de comunicação para o

renderizador apropriado, que executará o restante do processamento geométrico e dos cálculos de rasterização para essas primitivas.

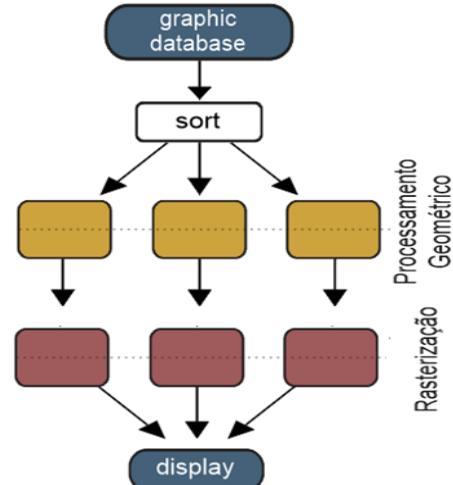


Fig. 2. *Sort-First* - figura criada com base em Eilemann [6].

As distribuições das primitivas no início do processo de renderização é a característica do *Sort-First*. Molnar [1] salienta que o processo envolve conjuntos de custos operacionais (*overhead*), desde que o desempenho de alguns processamentos geométricos estejam em renderizadores errados. Os resultados desses cálculos precisam ser recomputados por novos renderizadores.

O *Sort-First* requer muito menos comunicação de largura de banda do que as outras abordagens se na primitiva for aplicada a técnica de Tessellated (a técnica de Tessellate decompõe primitivas grandes em primitivas menores, tipicamente usada em polígonos ou malhas de polígonos para gerar maior definição e qualidade na primitiva).

Correa [7] mostra o uso do *Sort-First* e sistema de aglomerados de PCs. Correa mostra que o *Sort-First* age durante o processamento geométrico, após um estágio de pré-transformação que determina em qual bloco na tela cada primitiva se localizara. O sistema com a arquitetura *Sort-First* apresentou um bom desempenho, o que permitiu renderizar modelos grandes e de forma interativa em tempo real, como visto na fig. 3.

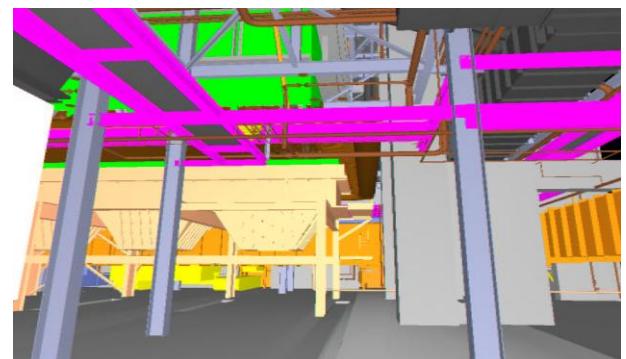


Fig. 3. Modelo de planta urbana usada no trabalho de Correa [7], usando 13 milhões de faces de triângulos.

Abraham [8] mostra em seu trabalho o uso de aglomerados baseado em PCs para melhorar o desempenho da renderização de cenas complexas, através de taxa de quadros em aplicações de realidade virtual.

Abraham opta por *Sort-First*, no qual a tela é partitionada em partes independentes, e são renderizadas por diferentes nós computacionais de renderização distintos. Esta abordagem evidencia a principal vantagem desse tipo de arquitetura: pouco requerimento de comunicação entre os nós gráficos. Entretanto, Abraham salienta a deficiência do *Sort-first* acerca da sobreposição das primitivas que precisam ser redundantemente renderizadas por mais de uma das partes disjuntas (fig. 4).

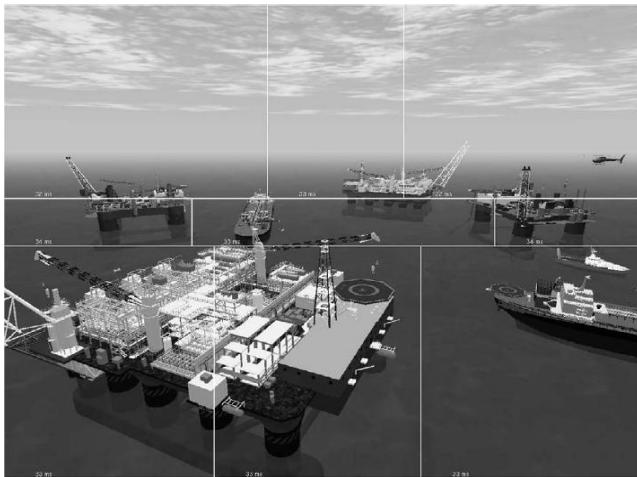


Fig. 4. Visualização de modelo virtual de 9 nós utilizando Sort-First[8].

A arquitetura do *Sort-first* pode ser explorada no processamento de nós de todo *pipeline* de renderização para uma porção da tela. Mas as primitivas ficam sujeitas em poder aglomerar-se em certas regiões, o que concentra o trabalho em poucos renderizadores.

IV. SORT-LAST

O *Sort-Last* consiste em aplicar o *sorting* no fim do *pipeline* da renderização, depois das primitivas serem rasterizadas em pixels, amostras, ou fragmentos de pixels, conforme apresentado na fig. 5.

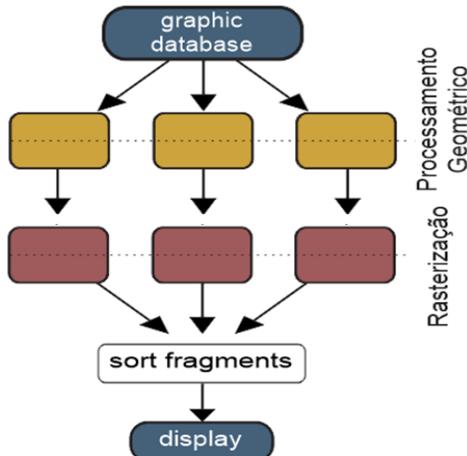


Figura 5. Sort-Last - figura criada com base em Eilemann [6].

Segundo Molnar, os sistemas de *Sort-Last* atribui cada processador (renderizador) arbitrariamente subconjuntos de primitivas, para que cada renderizador compute valores do pixel para seu subconjunto, não importando onde ele caia na tela. Os renderizadores então transmitem estes pixels em uma rede interconectada para os processadores desta composição. Por fim, é definida a visibilidade do pixel para cada renderizador.

Cavin [9] nos mostra que a arquitetura do *Sort-Last* permite que os renderizadores operem de modo independente até o estágio de visibilidade, no final do pipeline de renderização. Para aplicações interativas ou de tempo real, renderizar imagens de alta qualidade resulta em uma alta taxa de dados.

Em seu trabalho, Cavin apresenta análise de renderização de *pipeline* por *Sort-Last* aplicado em renderização de volume e polígonos. Cavin salienta que na renderização *Sort-Last* o conjunto de dados é decomposto e distribuídos através de diferentes nós de um cluster de computadores pessoais (PC). Para cada frame, cada nó renderiza uma imagem completa de dados atribuídos a uma GPU (*Graphics Processing Unit*, ou Unidade de Processamento Gráfico).

Calvin mostra que a renderização *Sort-Last* associada a aglomerado de PC apresenta um sistema escalável de confiança e estável, mas é preciso boa largura de banda para comunicação dos nós gráficos para não limitar o desempenho.

Eilemann [6] apresenta em seu trabalho a proposta do algoritmo *Direct Send*, baseado no *Sort-Last*, com modificações. O *Direct Send* divide a tarefa de recolher a imagem final para n blocos do espaço de tela para evitar a troca de imagens tamanho real entre os n canais de composição. Cada bloco está interligado no aglomerado gráfico, composto por um canal, os blocos são eventualmente montados em conjunto para formar a imagem final, como mostra a fig. 6.

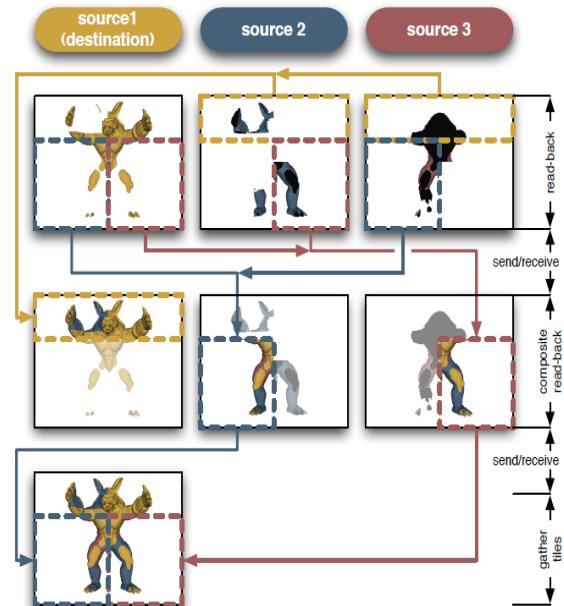


Fig. 6. Renderização Sort-Last usando Direct Send de três canais. As seis primeiras imagens mostram os buffers de renderização de cada bloco para cada canal e a imagem final com a imagem totalmente composta [6].

O *Direct Send* associado ao *Sort-Last* adaptado pode escalar o tempo de renderização até vários nós de processamento paralelo, especialmente com o tamanho do modelo médio, como visto na fig. 7.

Nota-se que um maior número de nós sobrecarrega a sincronização entre os nós, e pode se tornar um fator limitante para o paralelismo.

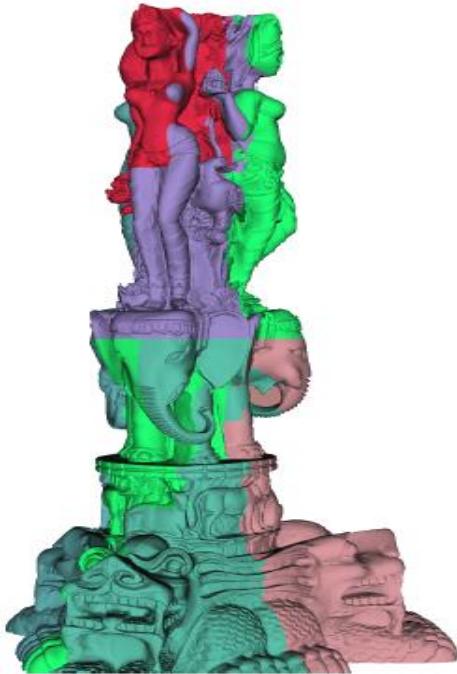


Fig. 7. Modelo médio de seis Nós do tipo *Sort-Last*, usando cores diferentes para cada Nô [11].

V. SORT-FIRST E SORT-LAST

Moloney [10] apresenta em seu trabalho um comparativo entre o *Sort-First* e *Sort-Last*. Ele salienta a dificuldade em comparar os métodos, pois muitos parâmetros influenciam no desempenho da renderização.

Moloney atenta para a queda de velocidade com que o *Sort-Last* trata dados grandes, mas o *Sort-Last* se mostra mais rápido quando o volume diminui e o numero de nós de renderização é menor.

Já o *Sort-Last* se mostra mais eficiente na comunicação entre os nós gráficos, o que proporciona taxas mais estáveis de quadros por segundo. A possibilidade do *Sort-First* permitir adaptações eficientes de muitos algoritmos de renderização de volume para sistemas paralelos.

Samanta [11] propõem a criação de um modelo híbrido do *Sort-First* e do *Sort-Last*. Samanta desenvolveu um algoritmo dependente de visão que particiona dinamicamente tanto a tela em blocos 2D e os polígonos 3D em grupos com o objetivo de equilibrar a carga de renderização entre os PCs para minimizar a largura de banda e composição dos blocos com uma redistribuição *peer-to-peer* dos pixels.

Samanta [12] apresenta em seu trabalho a possibilidade de construir um sistema de renderização paralela rápido, barato e com bom desempenho.

O sistema de Samanta apresenta um *pipeline* de 3 fase, mostrada na fig. 8:

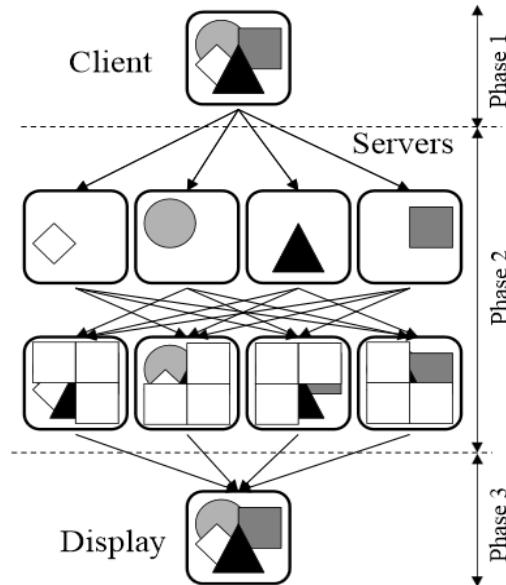


Fig. 8. Arquitetura do sistema híbrido *Sort-First* e *Sort-Last*, com fases de cliente, servidor e exibição proposto por Samanta [11].

- Na primeira fase, o cliente executa um algoritmo de particionamento que se decompõe ao mesmo tempo o modelo poligonal 3D em grupos disjuntos, atribuindo a cada grupo para um servidor PC diferente;
- Na segunda fase, cada servidor processa o grupo das primitivos 3D, que acessa o dado em volta do quadro do *buffer* na memória, atribuindo os dados a outro servidor para compô-los em seu *buffer* de quadro local para formar uma imagem completa de seu bloco. Cada servidor obtém valores de cor de pixels dentro de seu bloco e os envia para uma tela de PC;
- Na terceira fase, o visor do PC recebe subimágenes de todos os servidores, e agrupa-as para formar uma imagem completa para exibição.

O sistema de Samanta utiliza a comunicação *peer-to-peer* para associar todos os pixels entre servidores, técnica muito utilizada em aglomerados de PCs.

Na proposta de Samanta é importante ressaltar que a largura de banda agregada pelos pixels recebidos na tela da máquina é mínima. Cada pixel é enviado para a tela apenas uma vez para cada imagem na terceira fase do *pipeline*.

O modelo apresentado por Samanta se demonstrou uma técnica com bom desempenho e boa escalabilidade com modelos com grande quantidade de objetos em cena, como visto resultado da renderização na fig. 9.

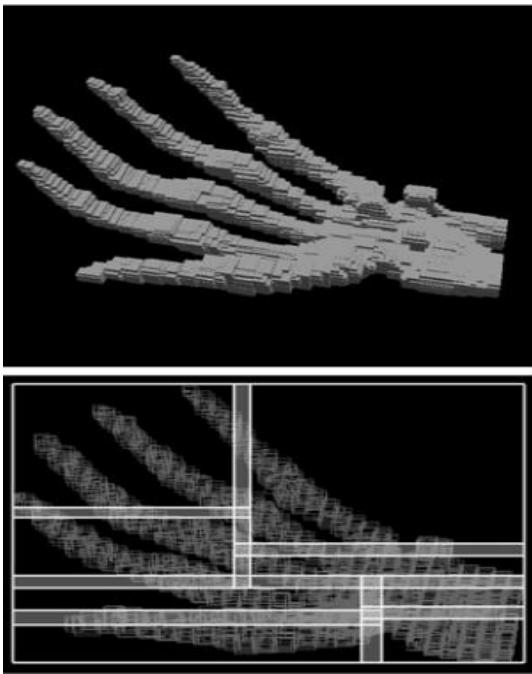


Fig. 9. Visualizações de blocos que sobrepõe em partições criadas com o algoritmo híbrido, modelo composto por 6.521 objetos [11].

VI. CONCLUSÃO

Neste artigo foram abordados métodos de renderização paralela. Fica evidente a necessidade de renderizar enorme volume de dados independente da forma desses modelos a serem visualizados. Com o avanço cada vez mais rápido da computação gráfica, é demandada também a rapidez na evolução de tratamento de renderização de grande volume de dados, como na indústria do entretenimento (filmes, animações, games), campo profissional (simuladores, CAD, visualização da informação) e demanda científica.

A computação gráfica associada à renderização paralela é uma alternativa eficiente. Vale salientar que o problema da renderização em si é a base para a classificação de algoritmos de renderização paralela, mas para cada caso específico o algoritmo de *sorting* depende da implementação e da aplicação de cada problema de renderização específico a ser estudado.

Contudo a limitações importantes de serem observadas como o alto consumo de banda da rede, ficando limitada a velocidade dessa banda, ficando prejudicado o tratamento de aplicações em tempo real.

REFERENCES

- [1] Molnar S., Cox M., Ellsworth D., Fuchs, “A Sorting Classification Of Parallel Rendering”. In ACM SIGGRAPH ASIA 2008 courses (New York, NY, USA, 2008), SIGGRAPH Asia ’08, ACM, pp. 35:1–35:11.
- [2] F. Crow, “Parallelism in Rendering Algorithms,” Proc. Graphics Interface 88, Morgan Kaufman, San Mateo, Calif., 1988, pp. 87-96.
- [3] S. Molnar, “Image-Composition Architectures for Real-Time Image Generation”, doctoral dissertation, TR 91-046, University of North Carolina at Chapel Hill, Oct. 1991.
- [4] Carl Mueller. “The sort-first rendering architecture for high-performance graphics”. ACM SIGGRAPH Computer Graphics (Special Issue on 1995 Symposium on Interactive 3-D Graphics), 1995.
- [5] K. Moreland, B. Wylie, and C. Pavlakos, “Sort-Last Tiled Rendering for Viewing Extremely Large Data Sets on Tiled Displays,” IEEE Symp. Parallel and Large Data Visualization and Graphics, 2001.
- [6] S. Eilemann and R. Pajarola, “Direct Send Compositing for Parallel Sort-Last Rendering,” Proc. Eurographics Symp. Parallel Graphics and Visualization (EGPGV), 2007
- [7] W.T. Correa, J.T. Klosowski, and C.T. Silva, “Out-of-Core Sort-First Parallel Rendering for Cluster-Based Tiled Displays,” Proc. Eurographics Workshop Parallel Graphics and Visualization (EGPGV ’02), pp. 89-96, 2002.
- [8] F. R. Abraham, W. Celes, R. Cerqueira, and J. L. E. Campos. A load-balancing strategy for sort-first distributed rendering. In Proceedings of SIBGRAPI ’2004. IEEE Computer Society, 17–20 Oct. 2004.
- [9] X. Cavin and C. Mion, “Pipelined Sort-Last Rendering: Scalability, Performance and Beyond,” Proc. Eurographics Symp. Parallel Graphics and Visualization (EGPGV), 2006.
- [10] B. Moloney, M. Ament, D. Weiskopf, and T. Moller, “Sort First Parallel Volume Rendering,” IEEE Trans. Visualization and Computer Graphics, vol. 99, Sept. 2010.
- [11] R. Samanta, T. Funkhouser, K. Li, and J.P. Singh, “Hybrid Sort- First and Sort-Last Parallel Rendering with a Cluster of PCs,” Proc. Eurographics Workshop Graphics Hardware, pp. 97-108, 2000.
- [12] [12] R. Samanta, T. Funkhouser, K. Li, and J. P. Singh. Sort-first parallel rendering with a cluster of PCs. SIGGRAPH 2000 Technical Sketch, August 2000.

A Influência da Taxa de Quadros no Desempenho de Tarefas em Sistemas de Realidade Aumentada

Joana S. Nascimento* e Silvio R. R. Sanches†

Universidade Tecnológica Federal do Paraná - UTFPR

Departamento de Computação, Cornélio Procópio

Email: joana_nascimento_@hotmail.com*, silviosanches@utfpr.edu.br†

Resumo—Aplicações de Realidade Aumentada como as baseadas em marcadores fiduciais requerem interação em tempo real do usuário com os elementos virtuais exibidos na cena. A qualidade gráfica dos objetos virtuais exibidos nas aplicações e o fato de algumas delas serem executadas em plataformas de baixo poder computacional ou via web tornam requisitos como a taxa de quadros um fator importante. Essa taxa deve se manter em níveis aceitáveis para garantir a qualidade da interação e satisfação do usuário. O objetivo desta pesquisa é, por meio de experimentos subjetivos, identificar a taxa de quadros mínima que aplicações de Realidade Aumentada, baseadas em marcadores e executadas em dispositivos móveis, devem atingir para que não cause desconforto e dificuldade de interação.

Abstract—Augmented Reality applications based on fiducial markers require real-time user interaction with virtual elements rendered in the scene. The graphic quality of these virtual objects generated by applications and the fact that some of them run on low computing power devices or via web platforms make requirements as the frame rate an important factor. The frame rate should be maintained at acceptable levels to ensure the quality of interaction and the user satisfaction. The objective of this research is to perform subjective experiments to identify the minimum frame rate for Augmented Reality mobile applications based on fiducial markers to avoid discomfort and difficulty of interaction.

I. INTRODUÇÃO

O avanço da eficiência computacional e da tecnologia de Realidade Virtual (RV) possibilitaram o desenvolvimento de aplicações que permitem a combinação, em tempo real, de quadros de vídeos capturados por uma câmera e ambientes reais interativos. Como resultado desse progresso, a Realidade Aumentada (RA) surgiu com o objetivo de proporcionar ao usuário a possibilidade de interagir com objetos virtuais, dentro de um ambiente real. Nesse ambiente, o usuário, utopicamente, não seria capaz de diferenciar o que é real do que é virtual [1]. Entre as aplicações de RA, existem as que são acessíveis somente por meio de plataformas sofisticadas [2], outras, no entanto, podem ser executadas em plataformas populares [3].

Uma das tecnologias que permitem o desenvolvimento de aplicações de RA se baseia no uso de marcadores fiduciais [4], que são imagens impressas em uma superfície plana cujo desenho se assemelha a um código de barras. Esses marcadores normalmente possuem contornos formados por uma borda retangular. O software manipula os padrões reconhecidos na imagem com esses marcadores, calcula e mantém a coerência de seu posicionamento em relação ao visor da câmera [4].

Desse modo, a imagem final incluirá um objeto virtual sobreposto ao marcador. Esse processamento, atualmente, pode ser considerado simples e pouco custoso computacionalmente [1].

Os requisitos para interação, navegação e comunicação em um ambiente de RA requerem taxas altas de quadros por segundo e respostas rápidas, em função de sua natureza interativa. Esses requisitos podem ser influenciados principalmente pela qualidade gráfica dos objetos virtuais presentes no sistema. Para aplicações que envolvem controle interativo, são necessários tempos de resposta rápidos. Segundo [5], os atrasos não devem ser inferiores a 0,1 segundo e, quanto maior a frequência de movimento de algum objeto da cena, maior deverá ser a taxa de quadros e menor o atraso, para que se obtenha a sensação confortável de interação.

Aplicações de RA, inclusive as baseadas em marcadores fiduciais, requerem interação em tempo real do usuário com os elementos virtuais. Além da qualidade gráfica, como algumas dessas aplicações são executadas em plataformas de baixo poder computacional ou via web, requisitos como tempo resposta e taxa de quadros devem ser mantidos em níveis aceitáveis para garantir a qualidade da interação. Para certificar-se que o desempenho do usuário não seja afetado, é importante identificar qual a taxa de atualização mínima exigida por essas aplicações.

Uma vez que o avanço dos elementos virtuais presentes nas aplicações tornam-se cada vez mais realistas, consumindo tempo considerável de processamento, e que as aplicações de RA baseadas em marcadores fiduciais são executadas em diversos ambientes, como sistemas desktops e web, a expectativa quanto à taxa de atualização pode variar entre os usuários dessas aplicações.

Inserido neste contexto, o objetivo desse trabalho consiste em, por meio do desenvolvimento de um protótipo e da condução de experimentos subjetivos, identificar a taxa de atualização mínima exigida em aplicações de RA baseadas em marcadores fiduciais. Encontrar este limite (para uma ou várias aplicações) pode auxiliar desenvolvedores a encontrarem o equilíbrio entre qualidade gráfica e desempenho.

II. TRABALHOS RELACIONADOS

Historicamente, existe uma busca constante por novas formas de interação do usuário com a aplicação, seja ela um jogo ou uma ferramenta. Há um vasto campo de novas possibilidades quando somamos, por exemplo, RA com jogos [6], uma vez que é possível combinar características da RA com a

dinâmica dos jogos. Dessa forma, pode-se sintetizar elementos virtuais com técnicas de jogos convencionais.

Em diversas áreas existem muitas pesquisas que investigam o problema da taxa de atualização, entre elas, o trabalho de [7] que mostra um estudo sobre o protocolo de segurança TLS (Transport Layer Security) e uma ferramenta de criptografia AES (Advanced Encryption Standard). Ambos são aplicados em vídeos em tempo real enviados pela rede, com um propósito de que a taxa de quadros seja pouco afetada.

O impacto da taxa de quadros na percepção da qualidade de um vídeo foi avaliado no trabalho apresentado em [8] em que foi desenvolvido um modelo que demonstra como pode ser evitada a redução dessa taxa. No trabalho foram aplicadas avaliações subjetivas para obtenção de *feedback* preciso.

Em [9] foi apresentada uma métrica para a avaliação da qualidade dos dados de um vídeo, levando em consideração o impacto da taxa de quadros. Naquele trabalho foi utilizado um estudo subjetivo utilizando vídeos de baixa e alta qualidade, que eram diferentes em termos de taxa de quadros e distorção. A métrica produzida, que utiliza um valor da relação sinal-ruído da sequência de vídeo, além de outros aspectos, como a taxa de quadros, mostrou resultados satisfatórios comparada a outras existentes na literatura.

Em aplicações que utilizam a rede, em que larguras de banda variam consideravelmente entre os receptores, o tempo que um vídeo leva para ser codificado antes do envio deve ser pequeno o suficiente para manter a aplicação com uma taxa aceitável. Essa taxa, que pode variar inclusive conforme o tamanho do quadro, é um fator determinante para a satisfação do usuário.

Nessa situação cada quadro deve possuir qualidade adequada, conforme detalhado no estudo apresentado em [8] que visa escolher as resoluções espaciais, temporais, e amplitudes adequadas de modo a alcançar uma harmonia entre a qualidade da imagem e do movimento dentro do conteúdo vídeo. Naquele trabalho, os autores demonstraram que a taxa de quadros é um fator que influencia a qualidade.

No estudo apresentado em [10] foi discutido como a percepção da qualidade de um vídeo pode mudar conforme a variação da taxa de quadros. Foi demonstrado que a degradação da qualidade do vídeo está ligada a redução desse fator.

Entre os trabalhos que investigam a influência da variação da taxa de quadros no contexto de RA, pode ser destacado o estudo realizado em [11]. No trabalho, os autores implantaram sistemas de RA que se adaptam em tempo real, considerando algumas limitações como: taxa de quadros, tamanho da imagem e atrasos. O projeto tinha como foco a qualidade de experiência de usuários, procurando valores ideais para esses requisitos.

III. TAXA DE QUADROS

Nos sistemas convencionais de RA, inclusive os baseados em marcadores, existem três principais tarefas que são executadas sequencialmente: aquisição da imagem, posicionamento da câmera para rastreamento e renderização [12]. Entre essas três tarefas, o cálculo da câmera representa uma

porção significativa do total de processamento, pois exige grande poder computacional. Em consequência disso, pode haver uma decadência na taxa de quadros, afetando a qualidade da aplicação [13].

A taxa de quadros depende também de fatores como complexidade gráfica e textura. Desse modo, ela está relacionada com a complexidade do ambiente e com a quantidade de objetos dinâmicos presentes na cena. Com isso, os atrasos podem ser percebidos durante a execução da aplicação [1].

Os filmes, por exemplo, possuem um forte compromisso com a qualidade de objetos virtuais gerados, requerendo muita potência e tempo computacional para gerar os quadros. O tempo de processamento durante a produção é consideravelmente maior que o tempo de exibição da cena após estar concluída. No entanto, nos filmes, não existe compromisso com o tempo de execução e todo o processamento pode ser realizado em fase de pós-produção utilizando a melhor qualidade gráfica possível.

Uma vez que a RA tem como requisito a interação em tempo real, a qualidade dos objetos virtuais deve ser cuidadosamente avaliada. Todo o processamento necessário para construir o ambiente da cena deve ser realizada a cada quadro, que deve ser exibido em tempo real ao usuário. Desse modo, a qualidade desses objetos tende a ser menor em RA, quando comparada com as produções cinematográficas, para que a taxa de quadros exibida se mantenha em níveis aceitáveis [1].

Os videogames, por sua vez, possuem maior relação com sistemas de RA, devido ao fato de necessitarem de interação. A princípio, os elementos gráficos dos jogos eram em 2D, porém com a mudança para os jogos em 3D, foram agregados todas as características de RA (elementos reais e virtuais) em alguns jogos [1], [6]. A taxa de atualização de quadros também é fator preponderante em jogos de RA.

IV. DESENVOLVIMENTO DO TRABALHO

Para realização dos testes que permitirão identificar a taxa de atualização mínima exigida por aplicações de RA baseadas em marcadores fiduciais, foi desenvolvido um método baseado em um protótipo que é executado em dispositivo móvel.

A. Desenvolvimento do Protótipo utilizado nos Experimentos

A primeira etapa é o desenvolvimento de uma aplicação que reproduz um ambiente de RA baseado em marcadores fiduciais. A cada etapa da execução da aplicação, diferentes taxas de quadro são simuladas. O protótipo foi desenvolvido com o motor de jogo Unity3d, uma vez que se trata de um ambiente de alto nível e possibilita referências de bibliotecas externas e maior facilidade para integrar bibliotecas de RA.

A biblioteca Vuforia, que pode ser integrada com o Unity3d, também foi utilizada. A Vuforia possibilita o desenvolvimento de interfaces de RA, identificando marcadores configurados no sistema no conteúdo da imagem capturada por uma câmera. A aplicação foi compilada para execução no sistema operacional Android.

B. Aplicação dos Experimentos

Os voluntários recrutados para os experimentos são alunos de graduação de diferentes cursos de tecnologia. Ao todo, 17 (dezessete) voluntários participaram da pesquisa. A quantidade de voluntários foi definida com base em recomendações apresentadas por métodos de avaliação subjetiva de qualidade de vídeo [14], [15], [16]. Com o propósito de obter resultados consistentes, esses participantes não possuíam nenhuma formação técnica ou trabalhavam na área de qualidade de vídeo ou desenvolvimento de sistemas de RA.

O objetivo dos usuários na aplicação consiste em colidir dois objetos virtuais, uma esfera e um cubo, cada um deles associados a marcadores distintos, como exibido na figura 1. Na cena, a esfera permanece quicando durante todo o tempo de execução da aplicação. O cubo, por sua vez, permanece fixo no marcador.



Figura 1. Tela apresentada ao usuário durante a execução dos experimentos subjetivos. Existem dois marcadores: o primeiro associado à bola, que permanece imóvel e o segundo, associado ao cubo, que deve ser movimentado pelo participante.

A aplicação foi dividida em cinco fases, cada uma exibe uma bola utilizada em um esporte diferente: futebol, basquete, golfe, beisebol e tênis. Cada fase é apresentada ao usuário com uma taxa de quadros específica, conforme mostrado na tabela I

Tabela I. OBJETOS VIRTUAIS EXIBIDOS DURANTE A APLICAÇÃO E AS RESPECTIVAS TAXAS DE ATUALIZAÇÃO CONFIGURADAS NO MOMENTO EM QUE CADA OBJETO ERA APRESENTADO QUICANDO SOBRE O MARCADOR.

Objeto Virtual	Taxa de Quadros
Bola de Futebol	18 quadros por segundo
Bola de Basquete	10 quadros por segundo
Bola de Golfe	4 quadros por segundo
Bola de Beisebol	1,5 quadros por segundo
Bola de Tênis	1 quadro por segundo

Sempre que uma colisão é provocada, a esfera troca de textura e o cubo, que originalmente é branco, passa a exibir a cor vermelha (Figura 2) e um botão com o rótulo “Próximo” é exibido para que uma nova fase seja iniciada.

Quando uma fase é encerrada, o usuário mostra a tela do celular para o instrutor do teste, que verifica se a taxa de quadros média corresponde à que foi configurada previamente no sistema. Em seguida, o usuário é instruído a clicar no botão “Próximo”, o cubo recebe a cor branca novamente e uma nova fase se inicia.

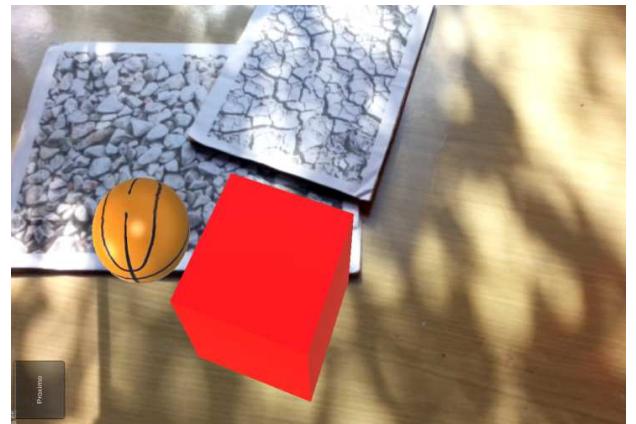


Figura 2. Tela apresentada quando o usuário consegue provocar a colisão entre os objetos virtuais por meio da movimentação do marcador associado ao cubo. A nova fase é iniciada com uma nova textura (outro tipo de bola) e outra taxa de quadros, quando o botão “Próximo” é clicado.

Antes do início do experimento, os voluntários foram instruídos sobre como movimentar um dos marcadores para provocar a colisão. Uma série de recomendações foram especificadas, com o objetivo de garantir uniformidade no ambiente e no procedimento de execução. Tais especificações são:

- Estar a 30 centímetros de distância dos marcadores;
- Estar em ambiente claro;
- O celular deve estar na altura do peito virado na horizontal;
- O marcador associado à bola deve estar na posição horizontal e o associado ao cubo, na vertical;
- Assim que houver colisão entre os elementos virtuais, colocar o marcador associado ao cubo no local inicial;
- Não movimentar o marcador associado à bola;
- O usuário deve permanecer sentado.

Durante a bateria de testes o usuário é instruído a informar o momento em que ocorre as seguintes situações:

- Quando o ambiente do jogo esteja causando desconforto/lentidão ou qualquer característica na animação que faça a cena parecer não natural.
- A partir de qual bola ou para quais bolas a tarefa estava difícil de ser realizada.

C. Análise dos Resultados

Posteriormente aos experimentos, foi realizada a análise dos dados obtidos. Conforme descrito na seção IV-B, para melhor qualificar o desempenho da aplicação o usuário deve responder exatamente em qual bola começou a insatisfação ou o incomodo com a animação da aplicação. Os resultados obtidos dessa questão podem ser observados na figura 3.

A partir dos resultados, pode ser observado que a insatisfação dos usuários é notada em atualizações menores que 4 quadros por segundo.

Animação começou a ficar ruim

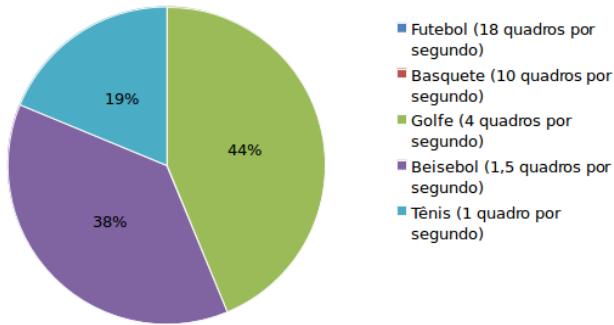


Figura 3. Gráfico mostrando o percentual obtido das respostas dos usuários em relação à pergunta “Quando o ambiente do jogo causa desconforto ou apresenta qualquer característica na animação que faça a cena parecer não natural?”. A pergunta era respondida durante a execução da aplicação.

A partir desses resultados, pode-se concluir que uma aplicação de RA baseada em marcadores fiduciais executada em dispositivos móveis, como um celular com telas menores e pouco poder de processamento, deve preservar essa taxa de atualização. A complexidade e a quantidade de objetos virtuais exibidos na cena, que são fatores que influenciam a taxa de quadros, deve ser ponderada de forma que o custo computacional permita a execução dentro desses patamares.

O segundo questionamento ao usuário considerava “A partir de qual bola ou para quais bolas a tarefa estava difícil de ser realizada”. Essa pergunta tem como objetivo identificar em que momento a tarefa se tornava exaustiva demais para ser realizada por conta dos poucos quadros da animação.

Intuitivamente, os efeitos causados na aplicação representariam um prejuízo na capacidade do usuário de posicionar o marcador na posição correta. Os resultados obtidos desse questionamento podem ser observado na figura 4

Ficou difícil de realizar a colisão

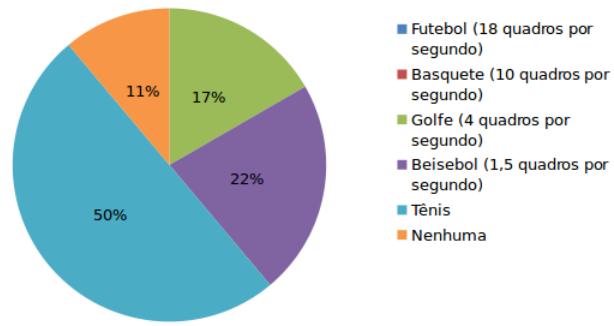


Figura 4. Gráfico mostrando o percentual obtido das respostas dos usuários em relação à pergunta “A partir de qual bola ou para quais bolas a tarefa estava difícil de ser realizada”. A pergunta era respondida durante a execução da aplicação.

Pode-se observar que a grande insatisfação da maioria dos usuários foi quando as variações dos quadros da aplicação estavam em torno de 1 quadro por segundo, situação na qual os quadros chegavam ao seu extremo com menor poder processamento. Nesse ponto a cena travava muito e, para

grande parte dos usuários, realizar a tarefa de colisão foi extremamente difícil.

Nota-se também que 40% dos usuários sentiram esse desconforto antes, nas bolas de golfe e beisebol. Na primeira, a taxa de quadros estava em torno de 1,5. Também foi possível constatar que 12% dos usuários, mesmo com taxas baixas, não sentiram desconforto algum em realizar a colisão.

A partir desse experimento foi possível concluir que o desconforto provocado pela baixa taxa de atualização não está diretamente relacionado com a facilidade de realizar tarefas dentro ambiente de RA. A taxa de quadros mínima exigida para um aplicativo de RA executada em plataforma Android, fazendo uso de marcadores, é aproximadamente cinco quadros por segundo. Nesse patamar, tornam-se satisfatórios tanto o conforto em relação à naturalidade da cena da aplicação quanto a facilidade de realizar tarefas deslocando marcadores dentro do ambiente de RA.

V. CONCLUSÃO

Apesar de existirem estudos anteriores que mostram que a taxa de quadros deve ser mantida perto de 8 quadros por segundo, neste trabalho foi demonstrado que esse valor pode variar dependendo do tipo da aplicação RA e até mesmo ser dependente da plataforma em que a aplicação será executada. Foi demonstrado neste estudo que o desconforto provocado pela baixa taxa de atualização não está diretamente relacionado com a facilidade de realizar tarefas dentro do ambiente de RA.

REFERÊNCIAS

- [1] R. Tori, C. Kirner, and R. Siscouto, Eds., *Fundamentos e tecnologia de realidade virtual e aumentada*, 1st ed. Porto Alegre: Editora SBC, 2006.
- [2] B. Furht, Ed., *handbook of Augmented Reality*. Springer, 2011.
- [3] S. R. R. Sanches, D. M. Tokunaga, V. F. Silva, A. C. Sementille, and R. Tori, “Mutual occlusion between real and virtual elements in augmented reality based on fiducial markers,” in *Proceedings of IEEE Workshop on Applications of Computer Vision – WACV 2012*. Washington, DC, USA: IEEE Computer Society, jan 2012, pp. 49 –54.
- [4] H. Kato and M. Billinghurst, “Marker tracking and hmd calibration for a video-based augmented reality conferencing system,” in *IWAR ’99: Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality*. Washington, DC, USA: IEEE Computer Society, 1999, pp. 85–94.
- [5] C. Kirner and M. S. Pinho, “Introdução á realidade virtual,” 1997, accessed at May, 2010.
- [6] B. H. Thomas, “A survey of visual, mixed, and augmented reality gaming,” *Comput. Entertain.*, vol. 10, no. 3, pp. 3:1–3:33, Dec. 2012. [Online]. Available: <http://doi.acm.org/10.1145/2381876.2381879>
- [7] N.-D. Khalifa and H. ELmahdy, “The impact of frame rate on securing real time transmission of video over ip networks,” in *Networking and Media Convergence, 2009. ICNM 2009. International Conference on*, March 2009, pp. 57–63.
- [8] Y.-F. Ou, T. Liu, Z. Zhao, Z. Ma, and Y. Wang, “Modeling the impact of frame rate on perceptual quality of video,” in *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, Oct 2008, pp. 689–692.
- [9] A. Ukhanova, J. Korhonen, and S. Forchhammer, “Objective assessment of the impact of frame rate on video quality,” in *Image Processing (ICIP), 2012 19th IEEE International Conference on*, Sept 2012, pp. 1513–1516.
- [10] Y.-F. Ou, Z. Ma, and Y. Wang, “Modeling the impact of frame rate and quantization stepsizes and their temporal variations on perceptual video quality: A review of recent works,” in *Information Sciences and Systems (CISS), 2010 44th Annual Conference on*, March 2010, pp. 1–6.

- [11] D. Perritaz, C. Salzmann, and D. Gillet, “Quality of experience for adaptation in augmented reality,” in *Proceedings of the 2009 IEEE international conference on Systems, Man and Cybernetics*, ser. SMC’09. Piscataway, NJ, USA: IEEE Press, 2009, pp. 888–893.
- [12] H.-S. Jang, J.-Y. Jeong, Y.-H. Kim, Y.-J. Yoon, and S.-J. Ko, “Augmented reality with high frame rate for low computational power devices,” in *Consumer Electronics - Berlin (ICCE-Berlin), 2011 IEEE International Conference on*, Sept 2011, pp. 274–275.
- [13] D. Wagner, T. Langlotz, and D. Schmalstieg, “Robust and unobtrusive marker tracking on mobile phones,” in *Mixed and Augmented Reality. ISMAR 2008. 7th IEEE/ACM International Symposium on*, 15-18 2008, pp. 121 –124.
- [14] F. Kozamernik, V. Steinmann, P. Sunna, and E. Wyckens, “Samviq – a new ebu methodology for video quality evaluations in multimedia,” *SMPTE Motion Imaging Journal*, vol. 114, no. 4, pp. 152 – 160, april 2005.
- [15] ITU-R, “Recommendation ITU-R BT.500-12 – methodology for the subjective assessment of the quality of television pictures,” International Telecommunications Union, Geneva, Switzerland, BT Series Broadcasting service (television) BT.500-12, 2009.
- [16] ITU-T, “Recommendation ITU-T BT.P.910 – subjective video quality assessment methods for multimedia applications,” International Telecommunications Union, Series P: Telephone Transmission Quality, Telephone Installations, Local Line Networks (Audiovisual quality in multimedia services) P.910, 2008.

Towards Automatic Evaluation of Personality Tests

Avaliação Automática de Testes de Personalidade

Caio José dos Santos Brito¹, João Marcelo X. N. Teixeira², Veronica Teichrieb¹

Voxar Labs - Centro de Informática - Universidade Federal de Pernambuco¹

Departamento de Estatística e Informática - DEINFO - Universidade Federal Rural de Pernambuco²

{cjsb, jmxnt, vt}@cin.ufpe.br

Abstract—Personality testing has become quite popular in job selection in an attempt to predict human behavior and deduce if the person tested has the qualification for the job. In Brazil, a commonly used test was introduced by Agostinho Minicucci and it is able to predict personality by the ability of writing vertical traces distributed in lines over time. The measurement of the test results is done manually by a psychologist, which demands a lot of time and work effort. To overcome those problems, a solution is proposed using image processing to measure the results of the test automatically. The solution was validated with 3 psychologists, being able to calculate the productivity of the tested person, the size of each line stroke, the angle of a line stroke, the distance between 2 consecutive strokes, the distance between two lines of strokes, the distance between lines and borders, the direction of a line of strokes and the regularity of the writing. The solution was able to measure with precision most of the variables, except the trace irregularity which had a 76% compatibility with psychologists manual classifications. This was expected due to the fact it is a qualitative analysis that can vary between different psychologists as well.

Keywords—personality test; image processing; line stroke

Resumo—Testes de personalidade tornou-se bastante popular na seleção de emprego em uma tentativa de predizer o comportamento humano e deduzir se a pessoa testada tem a qualificação para o trabalho. No Brasil, um teste comumente utilizado foi introduzido por Agostinho Minicucci e é capaz de prever a capacidade de personalidade por vestígios de escrever traços verticais distribuídos em linhas ao longo do tempo. A medição dos resultados do teste é feita manualmente por um psicólogo, o que exige uma grande quantidade de tempo e esforço de trabalho. Para superar esses problemas, uma solução é proposta usando processamento de imagens para medir os resultados do teste automaticamente. A solução foi validada com 3 psicólogos, sendo capaz de calcular a produtividade da pessoa testada, o tamanho de cada traçado da linha, o ângulo traço, a distância entre 2 traços consecutivos, a distância entre duas linhas, a distância entre as linhas e fronteiras, a direção de uma linha de traços e da regularidade da escrita. A solução foi capaz de medir com precisão a maioria das variáveis, exceto a irregularidade do traço que obteve uma compatibilidade de 76% com as classificações manuais dos psicólogos. Isto era esperado devido ao fato de que é uma análise qualitativa que pode variar entre diferentes psicólogos.

Palavras-chave – teste de personalidade; processamento de imagem; traços

I. INTRODUÇÃO

A palavra "teste" se refere a qualquer método padronizado e sistemático com propósito de obter informações sobre o comportamento humano [1]. A típica definição de um teste psicológico geralmente irá cobrir três características [2]:

- Um teste psicológico é um exemplar de um comportamento;
- Uma amostra é obtida em condições padrões;
- Regras de pontuação ou informações quantitativas são estabelecidas a partir do comportamento da amostra.

As duas áreas tradicionais para a aplicação de testes de personalidade têm sido ambientes clínicos e ocupacionais. Recentemente, testes de personalidade tornaram-se populares no campo da psicologia dos esportes, onde eles são usados para ganhar intuições sobre fatores que afetam o desempenho do atleta [1].

Testes de personalidade também têm uma forte tradição em seleções de emprego. [3] indica que testes de personalidade são capazes de indicar características específicas do empregado que possam ser úteis diante de situações no emprego e tais testes podem ser melhores indicadores do que carta de recomendação, pois as empresas têm medo de serem precessadas por difamação. [4] mostra que, em 2003, 30% das empresas americanas usaram teste de personalidade para ajudar na seleção de emprego. O trabalho de [5] ressalta importância e benefícios do uso teste psicológico nas empresas. Além disso, [6] indica as vantagens que esses testes trazem em um custo benefício a longo prazo.

Um teste comum para a análise psicológica foi desenvolvido na Espanha por Salvador Escala Milá e introduzido no Brasil por Agostinho Minicucci. O teste palográfico é usado por psicólogos em muitas áreas, por exemplo, seleção de emprego e aquisição de carteira de motorista. Ele é capaz de avaliar a personalidade de alguém por sua capacidade de escrever traços verticais ao longo do tempo [7]. Psicólogos experientes constatam que eles levam cerca de 15 a 20 minutos para analisar esse tipo de teste. Por isso, em dias onde há um grande número de avaliações, tal tarefa pode ser bastante cansativa.

Visando este cenário, propomos uma solução capaz de avaliar automaticamente o teste palográfico e, para o nosso conhecimento, não há uma solução disponível semelhante aquela proposta. Percebemos que mesmo os aspectos subjetivos do teste podem ser avaliados através de algoritmos, em vez de a psicóloga ter que realizar observação manual. Focamos a solução no processamento e análise de traços de linha, em outras palavras, uma biblioteca de terceiros (OpenCV) é utilizado para pré-processamento. O uso do sistema pode encurtar o tempo necessário para avaliar os testes, uma vez que fornece resultados em tempo real (leva menos de um segundo para processar completamente um teste). Os resultados da solução foram comparados com a avaliação manual de três psicólogos diferentes e eles mostraram estar corretos na maioria dos casos.

O restante desse artigo está organizado da seguinte forma. A seção 2 lista trabalhos relacionados a processamento de imagem aplicado na solução de diversos problemas e, especificamente, para a análise psicológica e de personalidade. A seção 3 detalha a solução proposta, e como a análises quantitativas e qualitativas são executadas. A seção 4 analisa os resultados obtidos com o sistema e eles são validados com os dados de referência de três psicólogos diferentes. E a Seção 5 conclui o artigo e fornece direcionamentos para trabalhos futuros.

II. TRABALHOS RELACIONADOS

Técnicas de processamento de imagem têm sido utilizadas por muitas áreas nos últimos anos. Empresas bancárias, forenses e psicólogos se beneficiam de diferentes métodos desenvolvidos. Esta seção lista os principais métodos pesquisados e os correlacionam com os problemas que enfrentam.

Em empresas bancárias, sistemas de verificação automática têm como objetivo comum reconhecer informações escritas à mão. De modo a reconhecer o texto escrito à mão, um passo de pré-processamento é necessário para que ele possa segmentar o conteúdo desejado a partir do resto da imagem (conteúdo gerado artificialmente, tais como palavras e símbolos impressos). O trabalho apresentado em [8] faz uso de propriedades e formas do próprio texto para identificar e distinguir elementos artificiais de aqueles escritos à mão no conteúdo. Uma vez que não usa qualquer informação a priori sobre o documento a ser analisado, a solução é mais genérica e eficiente.

Outra etapa de pré-processamento relacionada com o reconhecimento da informação manuscrita em cheques bancários é a correção da distorção. A fim de diminuir a variabilidade do texto capturado e facilitar o reconhecimento, o mesmo pode ser enquadrado e, em seguida, analisado por um processo de reconhecimento. O trabalho apresentado em [9] propõe uma abordagem morfológica para detecção e correção de textos distorcidos escritos em cheques bancários. Seu foco é facilitar o processo de reconhecimento de datas escritas em Português.

Medicina forense é uma área que também se beneficiou da análise de caligrafia. A fim de verificar a veracidade da assinatura de uma pessoa, diferentes abordagens podem ser adotadas. Por exemplo, o trabalho apresentado em

[10] compara a eficiência de quanto o uso de Wavelets diminuem as taxas de rejeição falsas. Combinando-as com funções de coordenadas $x(t)$ e $y(t)$, com base em casos de treinamento, os autores foram capazes de evitar que falsificações fossem aceitas nos estudos de caso.

A psicologia pode se beneficiar também a partir de algoritmos de processamento de imagem. Estudos indicam correlações entre escrita e comportamento pessoal, de modo que é possível prever personalidade com base no que uma pessoa projetam através da escrita. Grafologia é um método científico capaz de identificar, avaliar e compreender a personalidade com base em traços e padrões revelados pela escrita. O trabalho apresentado em [11] propõe um método para prever personalidade com base na linha base, a pressão da caneta, a letra "t", a curva inferior da letra "y" e a inclinação da escrita como encontrados na escrita de um indivíduo.

Semelhante ao trabalho anterior, há o HABIT (Handwritten Analysis Based Individualistic Traits Prediction) [12], uma ferramenta em Matlab para a previsão *offline* da personalidade com base em características como a inclinação da linha base, a pressão da caneta, inclinação das letras e tamanho da escrita. Como uma alternativa para trabalhos anteriores, [13] propõe um método para prever o comportamento da pessoa por meio de análise automatizada de caligrafia usando características como tamanho, inclinação e pressão, linha base, número de quebras, margens, velocidade de escrita e espaçamento entre as palavras. 80% dos estudos de caso analisados estão de acordo com a classificação manual de personalidade feita por especialistas.

Lógica Fuzzy (LF) também pode ser aplicada para a identificação do comportamento humano baseada em manuscritos. O trabalho apresentado em [14] propõe um método baseado em LF para a previsão de personalidade com base na linha de base e pressão da caneta. Com base nos testes realizados, o sistema pode oferecer uma taxa de precisão de 82%, devido à sua mudança dinâmica de caligrafia.

Os quatro trabalhos anteriores com foco em previsão de personalidade foram baseados em uma análise geral sobre todo o texto escrito. A fim de tornar mais precisas predições de personalidade, diferentes aspectos do texto pode ser analisado, tais como a assinatura da pessoa. Ao isolar a assinatura, é possível diminuir a complexidade do problema (desde agora uma porção menor do texto tem de ser avaliada) e melhorar a confiabilidade do sistema (uma vez que as assinaturas fornecem menos variação do que toda uma página inteira com texto escrito à mão). O trabalho apresentado em [15] utiliza técnicas de redes neurais artificiais (RNA) para correlacionar diferentes assinaturas de personalidades diferentes. Usando algoritmos multi-estrutura, foi possível obter uma precisão de 87-100% na classificação de personalidade, enquanto estiver usando RNAs com menos texto apresenta este número igual a 52-100%.

Nosso trabalho também lida com previsão personalidade utilizando técnicas de processamento de imagem. A fim de tornar o problema menos complexo e a solução

mais confiável, em vez de avaliar a totalidade do texto a partir de uma página ou mesmo uma única assinatura, o processo baseia-se em padrões de linhas propostas por [7] no seu teste de personalidade. Tal teste é amplamente utilizado nas admissões profissionais (especialmente no Brasil) e foi provado ser confiável e comumente utilizado no campo da psicologia [7] [16] [17]. Nossa solução inclui um sistema totalmente automatizado capaz de analisar os traços escritos pela pessoa durante a avaliação e dá como resultado as descrições da personalidade com base nas características encontradas pela análise de processamento de imagem. As seções a seguir fornecem detalhes sobre como o sistema proposto foi implementado e os resultados obtidos.

III. SOLUÇÃO PROPOSTA

Para iniciar o teste palográfico, um pedaço de papel é dado à pessoa a ser avaliada. Este papel contém, além do espaço para o preenchimento de algumas informações pessoais (em sua parte superior), quatro linhas verticais escritas para servir como referência, três delas na primeira linha e uma na segunda linha. Todos os traços têm um comprimento de 7 mm e são espaçados em 2,5 mm uns dos outros e a distância entre a primeira e segunda linhas é de 4 mm.

As instruções para o teste são: a pessoa testada deve desenhar traços o mais semelhante possível com as já impressas no papel. Novos traçados devem ser feitos de cima para baixo e da esquerda para a direita, tentando manter a mesma distância entre eles. Assim que a escrita chegar ao fim de uma linha, o padrão deve ser continuado numa nova linha imediatamente abaixo. A cada 30 segundos, há uma sinalização, e uma linha traço horizontal deve ser feito (-") e o teste deve continuar após o sinal. Esta indicação é repetida cinco vezes. Após a conclusão do teste, um psicólogo tem que avaliá-lo com base em algumas características traços feitos pela pessoa. Essas informações são:

- Número total de traços e produtividade por intervalo;
- distância entre traços na mesma linha;
- Angulação do traçado;
- Tamanho do traço;
- Direção de cada linha;
- distância entre linhas consecutivas;
- distância entre o primeiro traço de cada linha e a borda esquerda;
- distância do último traço de cada linha e a borda direita;
- Regularidade da escrita.

A solução proposta pode ser dividida em 3 fases: pré-processamento, análise quantitativa e análise qualitativa. Cada uma delas será detalhada a seguir.

A. Pré-processamento

O primeiro passo para a solução proposta é encontrar cada traçado na imagem. Para fazer isso, a biblioteca cvBlob foi usada [18]. Ele implementa um esquema de segmentação com base em [19] para etiquetar cada blob encontrado e provou ser confiável, dada a simplicidade da imagem.

Um traçado S tem uma série de informações de pixel (P_x, P_y): o topo do traço $T(S)$, a parte inferior do traço $B(S)$, o ponto mais à esquerda $W(S)$ e o ponto mais à direita do mesmo $E(S)$ conforme mostrado na Fig. 1. Além disso, o ponto central pode ser calculado usando a equação 1 e cada pixel do contorno de S pode ser encontrado.

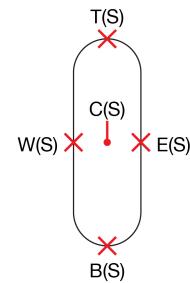


Figura 1. Elementos de um traço.

$$C(S) = \frac{1}{nbP} \sum_1^{nbP} S^i \quad (1)$$

nbP é o número de pixels de um traço e (S^i) é o i-ésimo pixel do traçado.

Depois que todos os traços são encontrados, eles precisam ser classificados como válidos e separados em linhas. A segmentação é baseada na distância entre $T(S)$ e $B(S)$. A mesma precisa ser maior que a distância entre $W(S)$ e $E(S)$. Caso a condição seja satisfeita, o traço está próximo da vertical e é considerado um traço válido. Os traçados são agrupados em linhas L_i . Para checar se um traço S_i pertence à mesma linha que S_{i-1} , a coordenada y de $C(S_i)$ precisa estar na extensão da coordenada y de $C(S_{i-1})$, como mostra a equação 2.

$$\begin{aligned} L = \{S_i \mid & C(S_{i-1})_x < C(S_i)_x \\ & e \\ & B(S_{i-1})_y \leq C(S_i)_y \leq T(S_{i-1})_y, 1 \leq i \leq n\} \end{aligned} \quad (2)$$

n é o número de traços no teste.

Após todos os traços serem segmentados e agrupados entre linhas, a análise quantitativa e qualitativa pode ser iniciada.

B. Análise Quantitativa

A análise quantitativa refere-se às definições diretas sobre características dos traços obtidos a partir da fase de pré-processamento. Essas definições são usadas de maneira

que a avaliação de teste deve dar os mesmos resultados, mesmo no caso de ser realizada por diferentes psicólogos. A maioria das análises quantitativas baseia-se em contagem, tamanho, ângulo do traçado, como descrito a seguir:

1) *Produtividade*: A produtividade pode ser medida com base no número total de traços escritos.

2) *Ritmo*: O ritmo se refere à variação entre o número de traços em cada intervalo de tempo em comparação com a produtividade.

3) *Tamanho de um traço*: O tamanho do traço representa simbolicamente a auto-estima e o nível de confiança da pessoa que está sendo testada em relação a seus próprios valores morais. Para calcular a distância entre $T(S)$ e $B(S)$ a equação 3 foi utilizada.

$$\text{tamanho}(S) = \sqrt{(T(S)_x - B(S)_x)^2 + (T(S)_y - B(S)_y)^2} \quad (3)$$

4) *Distância entre linhas*: Esta métrica refere-se à necessidade de contato com o exterior, a vontade de liberdade, de isolamento e concentração de uma pessoa. Para cada linha do teste, a distância entre o traçado foi calculado pela média das distâncias entre $B(S_i)$ e $B(S_{i-1})$, como mostra Fig. 2. A distância entre dois pixels pode ser calculada pela equação 4, similar à equação anterior 3.

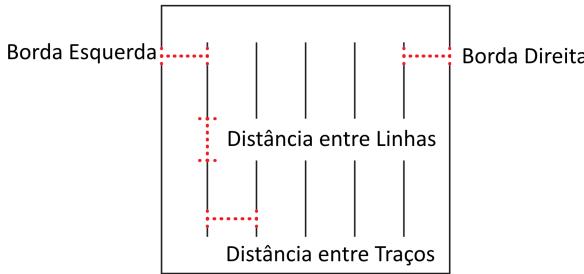


Figura 2. Distância entre traços, linhas e borda.

$$\text{distancia}(P_i, P_j) = \sqrt{((S_j)_x - (S_i)_x)^2 + ((S_j)_y - (S_i)_y)^2} \quad (4)$$

P_i e P_j são dois pixels na imagem.

5) *Inclinação do traço*: Essa medida prediz se a pessoa testada precisa de contato com outras pessoas e tenta se identificar como a pessoa interage com elas.

A inclinação de um traço S é calculada pelo complemento do ângulo relacionado ao arco-tangente da inclinação da linha criada por $B(S)$ e $T(S)$ como parâmetros, como mostra a equação 5.

$$\theta(S) = 90^\circ - \frac{180}{\pi} \arctan\left(\frac{T(S)_x - B(S)_x}{T(S)_y - B(S)_y}\right) \quad (5)$$

6) *Distância entre linhas*: Esta métrica avalia a relação interpessoal, em outras palavras, se a pessoa está aberta a um relacionamento ou não. A fim de calcular a distância entre as duas linhas, as distâncias médias entre $B(S)$ e $T(S)$ do primeiro e do último traço de duas linhas são consideradas, como mostra a equação 6 e pode ser visualizado na Fig. 2.

$$\text{distanciaLinha}(L, L') = \frac{1}{2}(\text{distancia}(L^0, L'^0) + \text{distancia}(L^l, L'^l)) \quad (6)$$

onde L e L' são duas linhas diferentes, L^0 é o primeiro traço de L , L^l é o último traço de L , L'^0 é o primeiro traço de L' e L'^l é o último traço de L' .

7) *Distância para a borda*: A distância para a borda está relacionada à capacidade da pessoa em se organizar com o ambiente e consigo mesma. Em relação à borda esquerda, a medida é feita calculando a distância entre o primeiro traço da linha e a origem na coordenada x, como pode ser visto na equação 7. Analogamente, a distância para a borda direita é calculada utilizando o último traço da linha em relação ao canto superior da imagem, como pode ser visto na equação 8.

$$\text{bordaEsquerda}(L) = C(L^0)_x \quad (7)$$

$$\text{bordaDireita}(L) = D - C(L^l)_x \quad (8)$$

D é a largura da imagem, a qual indica a coordenada x do canto superior da imagem.

C. Análise Qualitativa

A análise qualitativa representa avaliações que são dependentes do ponto de vista do psicólogo. Direção da linha, apesar de ter definições específicas de ângulos mínimos e máximos para determinar a classificação da linha, pode gerar avaliações duvidosas, uma vez que a maioria dos psicólogos só realiza análise visual e não usa instrumentos para medir os ângulos entre os traços. Esta variação de resultados é ainda mais comum em relação à análise de regularidade do traçado, em que não há uma definição direta de um traço irregular. O conceito é demonstrado usando amostras de imagem, mas o psicólogo deve julgar qual dos traços pertence às classificações regulares e irregulares. Os algoritmos desenvolvidos para a análise qualitativa são descritos a seguir.

1) *Direção da linha*: A direção da linha mapeia o humor, coragem e vontade de uma pessoa. Ela também tenta deduzir a maturidade e a constância em relação à personalidade, convicções e comportamento. A direção de uma linha L é calculada com base no baricentro de três traços: O primeiro traço L^0 , o último traço L^l e o traço L^m localizado no meio da linha. Duas linhas r e s são criadas usando esses pontos. A linha r é formada conectando o centróide do primeiro traço com o centróide do traço do meio e a linha s é criada conectando o centróide do traço

do meio com o centróide do último traço, como pode ser visto na Fig. 3.

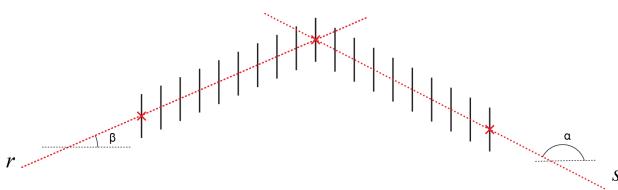


Figura 3. Inclinação da reta calculada usando 3 pontos principais (primeiro centróide, centróide do traço no meio da linha e o último traço).

Dependendo da inclinação de r e s , a linha pode ser classificada como: Horizontal, Crescente, Decrescente, Concava ou Convexa. A inclinação dessas retas pode ser definida pelas equações 9 e 10.

$$inclinacao(r) = \frac{C(L^m)_y - C(L^0)_y}{C(L^m)_x - C(L^0)_x} \quad (9)$$

$$inclinacao(s) = \frac{C(L^l)_y - C(L^m)_y}{C(L^l)_x - C(L^m)_x} \quad (10)$$

Em caso de ambas as linhas terem uma inclinação entre -1,9 e 1,4 graus, a linha de traços é classificada como uma linha horizontal. Caso contrário, se ambas as inclinações forem positivas, a linha é crescente. Caso ambas as inclinações forem negativas, a linha é decrescente. E se a inclinação de uma reta for positiva e outra tiver inclinação negativa, a linha é côncava; caso contrário, é convexa.

2) Regularidade do Traço: Para determinar se um traço é irregular em sua escrita, um solução estatística é proposta. Em um traço, uma linha r é criada utizando os pontos $T(S)$ e $B(S)$, então a média é calculada como na equação 12, sendo usada a distância $d(C, r)$ entre um traço e os pontos do contorno do mesmo e r (equação 11). O próximo passo é calcular o desvio padrão SD das distâncias entre pontos de contorno e r (equação 13). Caso o desvio padrão seja um valor alto, significa que a variação da distância entre o contorno do traço e r é alta, classificando o traço como irregular.

$$d(C, r) = \frac{|aC_x + bC_y + c|}{\sqrt{a^2 + b^2}} \quad (11)$$

$$media(contorno(S), r) = \frac{1}{nbC} \sum_{nbC} d(C, r) \quad (12)$$

$$SD(contorno(S), r) = \frac{1}{nbC - 1} \sum_{nbC} (d(C, r) - media(contorno(S), r))^2 \quad (13)$$

$r : ax + by + c = 0$ e C é um pixel do $contour(S)$ com nbC pixels nele.

De acordo com os testes realizados um valor limiar de 7 proporcionou bons resultados para a classificação de traços irregulares. Traços com desvio padrão superior a 7 são considerados irregulares.

IV. ANÁLISE DOS RESULTADOS

A fim de verificar a solução proposta, foram analisadas amostras que possuem posições, tamanhos e inclinações diferentes. Nossa solução tem uma limitação em relação à imagem, que deve ser fornecida como uma imagem retificada, sem distorção, e o papel deve estar na posição correta (cabeçalho no topo) pois o cálculo é feito com os traços no ângulo e na posição correta. A solução foi capaz de classificar os agrupamentos de pixels como traços válidos, agrupá-los em linhas e a análise quantitativa mostrou-se ser capaz de extrair as medições diretas do teste. Nas análises qualitativas, o método foi capaz de capturar a direção das linhas e os traçados irregulares dos testes, como mostrado na Fig. 4 e na Fig. 5. Para provar que o método é confiável, três psicólogos foram convidados a analisar o mesmo teste para comparar seus resultados com o resultado da solução proposta. Em relação à direção da linha, os psicólogos e solução proposta obtiveram resultados idênticos. Já para a análise de regularidade do traçado, a solução proposta mostrou uma taxa de confiabilidade de 76% quando comparado com os psicólogos.

V. CONCLUSÃO

Neste trabalho, propomos uma solução para avaliação automática de um teste de personalidade. O teste palográfico foi escolhido, o qual é baseado em parâmetros quantitativos e qualitativos, tais como a produtividade, tamanho e ângulo de cada traçado da linha, distância entre traços da linha e fronteiras, a direção de traços da linha e regularidade do desenho. Os fundamentos teóricos por trás do teste de personalidade se dão de modo que a pessoa projeta a si própria no papel utilizando o traçado. A aplicação do teste leva cerca de 5 minutos, enquanto a sua avaliação manual, por um psicólogo leva cerca de 15-20 minutos. O objetivo principal deste trabalho foi a implementação de um avaliador teste de personalidade totalmente automático, capaz de diminuir o tempo necessário para a análise de tais testes. A solução foi bem sucedida uma vez que era capaz de fornecer uma análise confiável em menos de um segundo.

O relatório gerado pela solução proposta foi comparado com as avaliações manuais de psicólogos. A solução entrega um resultado equivalente para todas as análises quantitativas e também para análise direção de linha. Em relação à classificação de linhas irregulares, a solução mostrou uma taxa de similaridade de 76%. Isso era esperado, porque a definição de uma linha irregular não é um conceito comum, mesmo entre psicólogos, o que significa que as suas avaliações também divergem em uma percentagem semelhante à do sistema.

Como trabalho futuro, pretendemos fazer o avaliador automático disponível como um serviço online gratuito, de

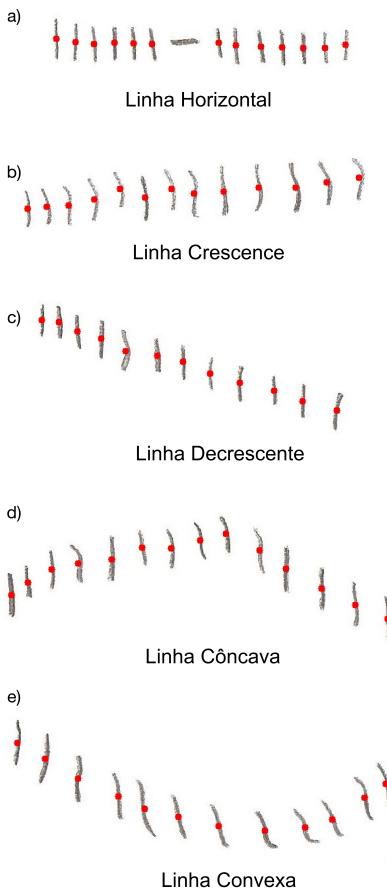


Figura 4. Padrões de linhas diferentes usados para validar a solução proposta e comparados com a avaliação dos psicólogos.

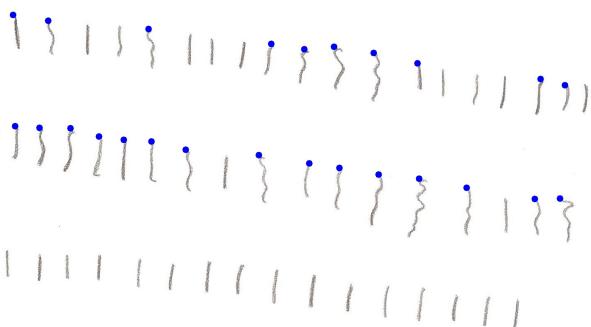


Figura 5. Traçado com diferentes irregularidades usados para validar a solução proposta e comparados com a avaliação dos psicólogos. Traços marcados com um círculo azul são classificados como irregulares pelo sistema.

uma forma os psicólogos poderiam fazer upload de testes digitalizados para o sistema e dar feedbacks baseados nas avaliações automáticas realizadas. Tais feedbacks poderiam ser usados para refinar os parâmetros de classificação da solução proposta e melhorá-la usando algum método de aprendizagem supervisionada.

REFERÊNCIAS

- [1] J. Athanasou, *Adult Educational Psychology*. Cengage Learning Australia, 1999. [Online]. Available: <http://books.google.com.br/books?id=xfW3AAACAAJ>
- [2] K. R. Murphy and C. O. Davidshofer, "Psychological testing," *Principles, and Applications*, Englewood Cliffs, 1988.
- [3] S. J. Stabile, "Use of personality tests as a hiring tool: Is the benefit worth the cost, the," *U. Pa. J. Lab. & Emp. L.*, vol. 4, p. 279, 2001.
- [4] M. Heller, "Court ruling that employer's integrity test violated ada could open door to litigation," *Workforce Management*, vol. 84, no. 9, pp. 74–77, 2005.
- [5] S. N. P. Dalbosco and J. da Silva Consul, "A importância da avaliação psicológica nas organizações," *Revista de Psicologia da IMED*, vol. 3, no. 2, 2011.
- [6] S. Laher and K. Cockcroft, "Psychological assessment in post-apartheid south africa: the way forward," *South African Journal of Psychology*, vol. 44, no. 3, pp. 303–314, 2014.
- [7] I. C. B. Alves and C. Esteves, *O teste palográfico na avaliação da personalidade*. Votor, 2004.
- [8] J. Eduardo Bastos Dos Santos, B. Dubuisson, and F. Bortolozzi, "Characterizing and distinguishing text in bank cheque images," in *Computer Graphics and Image Processing, 2002. Proceedings. XV Brazilian Symposium on*. IEEE, 2002, pp. 203–209.
- [9] M. E. Morita, F. Bortolozzi, J. Facon, and R. Sabourin, "Morphological approach of handwritten word skew correction," in *Computer Graphics, Image Processing, and Vision, 1998. Proceedings. SIBGRAPI'98. International Symposium on*. IEEE, 1998, pp. 456–461.
- [10] A. Vergara da Silva and D. Santana de Freitas, "Wavelet-based compared to function-based on-line signature verification," in *Computer Graphics and Image Processing, 2002. Proceedings. XV Brazilian Symposium on*. IEEE, 2002, pp. 218–225.
- [11] H. Champa and K. AnandaKumar, "Automated human behavior prediction through handwriting analysis," in *Integrated Intelligent Computing (ICIIC), 2010 First International Conference on*. IEEE, 2010, pp. 160–165.
- [12] A. Rahiman, D. Varghese, and M. Kumar, "Habit: Handwritten analysis based individualistic traits prediction," *International Journal of Image Processing (IJIP)*, vol. 7, no. 2, p. 209, 2013.
- [13] V. Kamath, N. Ramaswamy, P. N. Karanth, V. Desai, and S. Kulkarni, "Development of an automated handwriting analysis system," *ARPON Journal of Engineering and Applied Sciences Volume 6*, no. 9, pp. 1819–660, 2011.
- [14] C. M. Jadhav and A. B. Valsang, "A fuzzy logic approach for identifying the human behavior based on handwriting," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 4, no. 10, pp. 162–165, 2014.
- [15] E. C. Djamal, R. Darmawati, and S. N. Ramdlan, "Application image processing to predict personality based on structure of handwriting and signature," in *Computer, Control, Informatics and Its Applications (IC3INA), 2013 International Conference on*. IEEE, 2013, pp. 163–168.
- [16] R. Primi and J. C. Alchieri, "Instrumentos de avaliação mais conhecidos/utilizados por psicólogos e estudantes de psicologia," *Psicologia: Reflexão e Crítica*, vol. 18, no. 3, pp. 390–401, 2005.
- [17] D. A. Senden, R. M. F. Lopes, R. F. L. d. Nascimento, L. B. Terroso, and I. I. Argimon, "Análise da personalidade de profissionais da área da saúde: um estudo pelo teste palográfico," *Aletheia*, no. 42, pp. 62–70, 2013.
- [18] C. C. Linan, "cvBlob," <http://cvblob.googlecode.com>. [Online]. Available: <http://cvblob.googlecode.com>
- [19] F. Chang, C.-J. Chen, and C.-J. Lu, "A linear-time component-labeling algorithm using contour tracing technique," *computer vision and image understanding*, vol. 93, no. 2, pp. 206–220, 2004.

Integração de Ambientes Virtuais 3D Colaborativos baseados em Aglomerados Gráficos

Diego Roberto Colombo Dias

Laboratório de Visualização Imersiva,

Interativa e Colaborativa – LaVIIC

Departamento de Computação

Universidade Federal de São Carlos – UFSCar

Email: diegocolombo.dias@dc.ufscar.br

Marcelo de Paiva Guimarães

Universidade Federal de São Paulo (UAB/UNIFESP)/

Programa de Pós Graduação da Faccamp

Email: marcelodepaiva@gmail.com

Luis Carlos Trevelin

Laboratório de Visualização Imersiva,

Interativa e Colaborativa – LaVIIC

Departamento de Computação

Universidade Federal de São Carlos – UFSCar

Email: trevelin@dc.ufscar.br

Abstract—Ambientes Virtuais 3D têm sido de grande auxílio para áreas de treinamento e simulação, mesmo em ambientes locais. No entanto, soluções que possibilitem o uso deste tipo de ambiente em colaboração entre vários usuário é uma tendência. Ambientes de multiprojeção permitem ao usuário um alto nível de imersão, porém, este tipo de ambiente, normalmente, possibilita apenas o uso local, devido as suas complexidades de implementação e alto custo. Este artigo apresenta uma arquitetura de comunicação para ambientes virtuais 3D colaborativos baseados em aglomerados gráficos, utilizando como infraestrutura o miniCAVE, um ambiente de multiprojeção de baixo custo, e o AixCAVE, baseada no projeto original do CAVE. É apresentada a integração de duas bibliotecas de desenvolvimento distintas: libGlass e VistaVRToolKit. A VistaVRToolKit é utilizada como solução local de renderização, uma solução intra aglomerado, enquanto a libGlass é utilizada como solução inter aglomerado.

Abstract—3D Virtual environments have been used on training and simulation areas, even in local environments. However, solutions that enable the use of this type of environment on a collaborative way is a trend. Multi projection environments allow an high level of immersion to the user, however, this type of environment, only allows local use, usually, due to its implementation complexity and high cost. This paper presents a communication architecture for 3D Collaborative Virtual Environments based on graphics clusters, using as infrastructure the miniCAVE, a low-cost multi projection environment, and the AixCAVE, which is based on the original CAVE project. The integration of two distinct development libraries has been developed: libGlass and VistaVRToolKit. VistaVRToolKit is used as a local rendering solution, an intra cluster solution, while libGlass is used as an inter cluster solution.

I. INTRODUÇÃO

A Realidade Virtual (RV) envolve tecnologias de ponta para o desenvolvimento de aplicações em função de novas possibilidades, oferecendo diversos tipos de interface, por meio de dispositivos multi-sensoriais, navegação em espaços

tridimensionais, imersão no contexto da aplicação e interação em tempo real. Paralelamente, a visualização científica em ambientes distribuídos vem tendo importância crescente, por permitir a interação de especialistas localizados remotamente, dando suporte a recursos gráficos em tempo real.

O Ambiente Virtual Colaborativo 3D, em inglês *3D Collaborative Virtual Environment* (3DCVE), é um tipo de ambiente que prove as características de ambientes virtuais desenvolvidos utilizando RV. No entanto, agregam novas características, tendo como principal a possibilidade de interação entre múltiplos usuários em um mesmo ambiente virtual, mesmo que estes estejam localizados em lugares diferentes e distantes. Contudo, este tipo de ambiente possui alguns requisitos referente a tempo real, que são difíceis de alcançar em ambientes colaborativos que baseados em redes de dados sobre a Internet. Apesar de as redes de dados terem evoluídos nos últimos anos, ainda existem desafios a serem resolvidos com relação às adversidades ocasionadas por elas.

Vários ambientes distribuídos de RV já foram desenvolvidos, por exemplo, Second Life [1], Active Worlds [2], IMVU [3], Eye gaze [4] e VizGrid [5]. Esses ambientes têm buscado resolver problemas complexos, como o suporte a centenas de usuários e a comunicação desses via vídeo. No entanto, não foram projetados para aplicações que exijam alto poder computacional do sistema local, como, por exemplo, a geração e exibição de imagens de alta resolução. Além disso, geralmente são voltados para um domínio específico de problema e, apesar de serem ambientes tridimensionais, não suportam dispositivos de entrada/saída 3D, como o *Cave Automatic Virtual Environment* (CAVE) ou *PowerWall*.

Por outro lado, existem ferramentas de desenvolvimento de ambientes de RV voltadas a aplicações que exigem alto poder computacional, como as que se baseiam em aglomerados gráficos de computadores, que são arquiteturas específicas para

RV. Dentre elas, pode-se citar o FlowVR [6], o Syzygy [7], o Chromium [8], e o Net Juggler [9]. No entanto, tais ferramentas limitam as aplicações a uma determinada localidade, ou quando não, permitem apenas a construção dessas para domínios específicos.

A necessidade de 3DCVE tem sido crescente. Problemas relacionados à logística tendem a ser beneficiados, visto que vários usuários podem participar de atividades colaborativas estando em locais distintos. A redução de gastos também é um incentivo ao uso de tais ambientes. Assim, estudos que visem identificar soluções para o desenvolvimento de 3DCVE, possuem alto impacto tanto na academia, quanto na indústria.

Este artigo está organizado da seguinte forma. Esta seção apresentou uma breve introdução. A Seção II apresenta um geral sobre sistemas de multiprojeção. A Seção III apresenta os objetivos deste artigo. A Seção IV apresenta os materiais e métodos utilizados, tais como as bibliotecas de desenvolvimento e a infraestrutura. A Seção V apresenta o modo como foi proposta e implementada a solução de integração entre a libGlass e a VistaVRToolKit. E, por último, as discussões a respeito do artigo são apresentadas na Seção VI.

II. SISTEMAS DE MULTIPROJEÇÃO

Os sistemas de multiprojeção, geralmente compostos por múltiplas telas, proporcionam ao usuário diferentes pontos de vista de um mesmo ambiente, possibilitando uma experiência altamente imersiva. Os sistemas de multiprojeção têm sido pesquisados a mais de uma década como solução de renderização de ambientes virtuais complexos [10], [11], [12], [13].

Uma das formas de se implementar um ambiente virtual distribuído, é por meio de aglomerados gráficos. Estes são compostos pela combinação de *hardware* e *software*. Os componentes de *software* que compõem um aglomerado gráfico podem ser classificados em 4 camadas: camada de aplicação, ferramentas de visualização, gerenciador de sistema do aglomerado e o sistema operacional. A Figura 1 apresenta as camadas da pilha de *software*.



Fig. 1: Pilha de Software

A camada de aplicação é representada por ambientes virtuais utilizados de forma distribuída. Os ambientes podem ser desenvolvidos especificamente para serem ambientes distribuídos ou podem ser portados, dependendo da biblioteca de gerenciamento do aglomerado gráfico.

A camada de renderização é composta por bibliotecas responsáveis pelo “desenho” do ambiente virtual. A renderização do ambiente virtual pode ser feita por bibliotecas nativas, tais como OpenGL e Java3D, ou por meio do uso de motores gráficos, tais como Ogre3D e Unity3D. Os motores gráficos tendem a facilitar o desenvolvimento de ambientes virtuais, visto que possuem primitivas específicas para interação, renderização e compatibilidade com vários formatos de ambiente. No entanto, não possuem suporte a ambientes de multiprojeção com características de distribuição, sendo possível por meio da utilização de bibliotecas/frameworks para este fim. Neste artigo, a solução de renderização utilizada é provida pela VistaVRToolkit. A solução é baseada no OpenSG [14].

A camada de gerenciamento do aglomerado gráfico é responsável pela comunicação entre os nodos do aglomerado. Ela é responsável por abstrair as formas de comunicação entre os nodos de um aglomerado gráfico, incluindo a intercomunicação entre aglomerados localizados em locais distintos.

A camada do sistema operacional é responsável pelo ambiente no qual as camadas superiores serão instaladas. Ela prove a comunicação entre as camadas superiores (*software*) e o *hardware*.

Neste artigo, é apresentado um modelo de comunicação que aborda a camada de gerenciamento do aglomerado gráfico, especificamente para um modelo inter-aglomerado. A camada de aplicação e renderização são gerenciadas pela VistaVR-Toolkit.

III. OBJETIVOS

O fator inicial de motivação de pesquisa deste artigo é que as atividades colaborativas tem sido aplicadas em várias áreas, já estando difundidas nos mais diversos ambientes de trabalho em grupo. Várias aplicações, como por exemplo, o Google Docs, acabaram com as barreiras de tempo e disposição geográfica existentes nas interações entre membros de um grupo, possibilitando a comunicação de maneira ubíqua. Porém, em se tratando de 3DCVE, as soluções encontradas na literatura são limitadas. Elas geralmente não suportam imagens de alta resolução ou o uso de qualquer dispositivos de multiprojeção, como um CAVE; ou são limitadas a um domínio específico de problema. A carência de soluções a esses problemas estimula a investigação na área, abrindo diversos pontos de pesquisa.

Outro fato motivador é que a RV possibilita a interação e o envolvimento imersivo de usuários com simulações por intermédio de dispositivos de visualização 3D e dispositivos de rastreamento corporal. Assim, a junção de atividades colaborativas com aplicações de RV é um caminho natural da evolução tecnológica. Por isso, espera-se que diversas áreas possam se beneficiar dos 3DCVE. Na área da RV, vários investimentos têm sido realizados pela indústria na produção de *hardware*, *software* e dispositivos de entrada/saída, o que tem gerado um crescimento acelerado desta classe de aplicação. Na área de atividades colaborativas, as grandes corporações têm realizado

vários investimentos para que os funcionários resolvam os problemas em grupo, mesmo em situações onde seus membros estejam localizados em locais distintos.

O objetivo principal deste artigo é propor a integração de duas diferentes soluções de *software* voltada ao desenvolvimento de 3DCVEs. A infraestrutura utilizada é composta pelo miniCAVE [15] e o AixCAVE [16], do *Virtual Reality Group* (VRG) – RWTH Aachen University. O propósito foi realizar a integração de duas soluções de *softwares* distintas no desenvolvimento de 3DCVEs: libGlass [17] e VistaVRToolkit [18].

IV. MATERIAIS & MÉTODOS

O ambiente miniCAVE, solução de baixo custo baseada no projeto original do CAVE, é utilizado como infraestrutura no decorrer deste artigo. Assim, os experimentos prévios já realizados, e os futuros, serão efetuados com parceria entre o LaVIIIC – UFSCar, o LSTR – UNESP e o VRG – RWTH Aachen University. A Figura 2 apresenta o miniCAVE. A Figura 3 apresenta o AixCAVE.

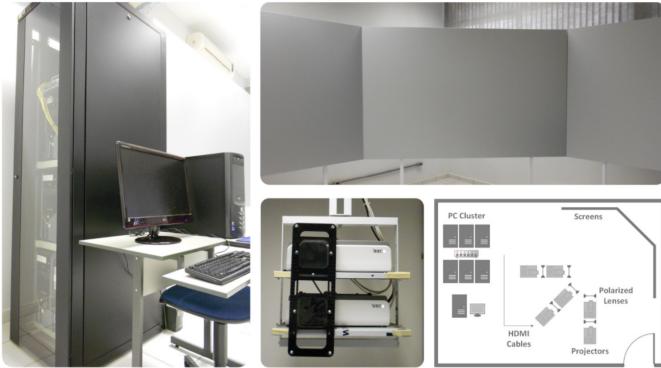


Fig. 2: MiniCAVE: Aglomerado Gráfico (7 computadores executando 6 instâncias (Clientes) e 1 Servidor), 3 telas, 6 projetores e lentes polarizadoras [19]

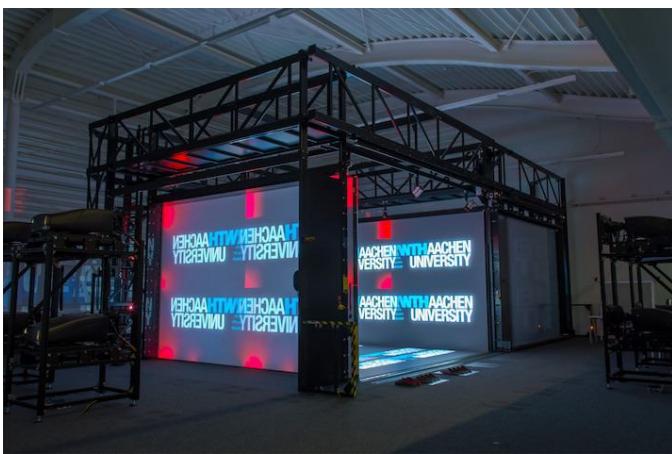


Fig. 3: AixCAVE: Aglomerado Gráfico (24 computadores executando 48 instâncias), 5 telas e 24 Projetores 3D Ativos Barco [16]

O principal objetivo é prover uma solução de software que permita a interconexão entre o miniCAVE [20] e o Aix-

CAVE [16], priorizando a qualidade de serviço das conexões. A verificação do status de conexão é realizada de forma dinâmica [21].

- **libGlass** [11][17][21]: é uma biblioteca voltada ao desenvolvimento de ambientes distribuídos em geral, computação em nuvem de alta performance e aplicações com pouco atraso, particularmente ambientes de RV e Realidade Aumentada. A libGlass foi desenvolvida na forma de componentes escaláveis, os quais podem ser usados por aplicações que necessitem distribuir dados entre nodos de um aglomerado gráfico. Dentre as principais características da biblioteca, pode-se destacar a interface amigável de desenvolvimento ao usuário.
- **VistaVRToolkit** [22]: é um *toolkit* que permite a integração de tecnologias relacionadas à RV, 3D e Visualização Científica. A VistaVRToolKit tem sido desenvolvida pelo VRG – RWTH Aachen University por mais de dez anos, em cooperação com o DLR Braunschweig. O *toolkit* é composto por múltiplos componentes.

V. MINICAVE & AIXCAVE – DESCRIÇÃO DA SOLUÇÃO

A Figura 4 define o comportamento esperado na comunicação entre os aglomerados gráficos. Em termos gerais, o processamento interno, como troca de mensagens, sincronismo, definição de pontos de vista e tratamento de interação ficam a cargo da VistaVRToolkit, contudo, sempre que a mensagem necessite ser enviada a outro aglomerado gráfico, esta será enviada utilizando a libGlass, provendo assim o modelo de conexão proposto e desenvolvido pelos autores.

O modelo de comunicação proposto entre os diferentes locais define que toda interação realizada por um usuário em um dos ambientes deve ser replicada a outros locais que tenham interesse neste estímulo. Assim, se um usuário gerar um estímulo no miniCAVE, este mesmo estímulo deverá ser realizado na AixCAVE, porém, sempre visando o menor atraso nesta tarefa.

Os ambientes virtuais utilizados diferenciam-se entre si tanto em *layout* físico, quanto no poder de processamento. Assim, algumas adequações tem de ser realizadas quanto ao posicionamento de câmeras. Toda a configuração do miniCAVE é feita seguindo como modelo os *scripts* de configuração já utilizados no AixCAVE. As Figuras 5 e 6, apresentam os diferentes *layouts*.

Para que as mensagens possam ser convertidas entre as diferentes soluções de desenvolvimento, um protocolo de comunicação é definido. Assim, sempre que uma mensagem necessite ser enviada entre diferentes aglomerados, ela deve ser empacotada em um objeto da libGlass. Da mesma forma, quando a mensagem é entregue ao aglomerado, esta deve ser desempacotada em um objeto da VistaVRToolkit. Os objetos libGlass são definidos como sendo do tipo que é empacotado, neste caso VistaVector3D e VistaQuaternion. Em termos gerais, sempre que um objeto VistaVRToolkit é criado, este deve ser instanciado como uma variável Shared (libGlass), sendo definido como modo de difusão Global. O termo Global na instanciacão do objeto permite que exista troca de mensagens entre *peers*, não restrinindo somente à arquitetura local Cliente-Servidor. O Código 1 descreve este comportamento em um protótipo simples.

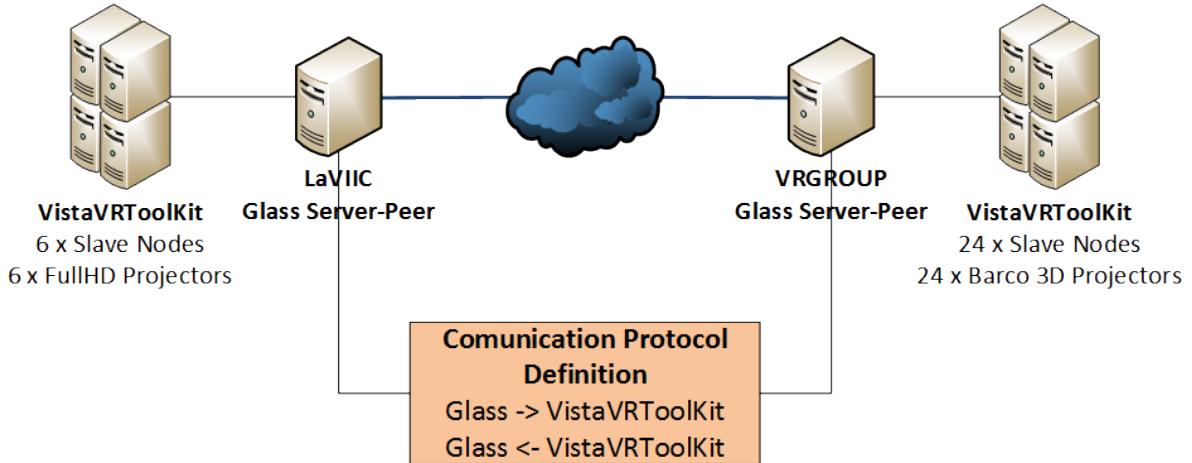


Fig. 4: Modelo de Comunicação – AixCAVE & miniCAVE

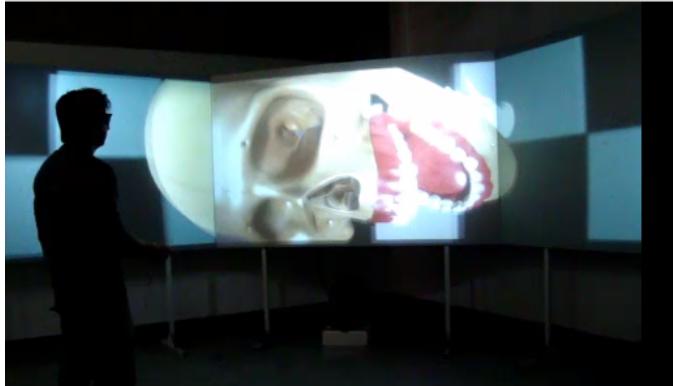


Fig. 5: miniCAVE [23]

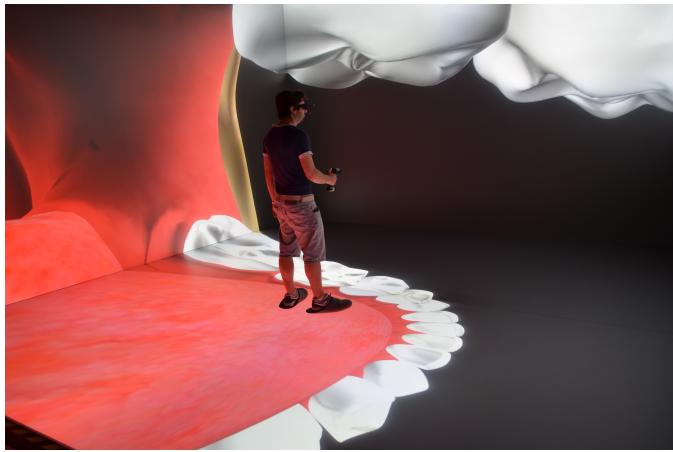


Fig. 6: AixCAVE

Código 1: Compartilhamento do Posicionamento de Câmera – Master

```
Shared<VistaVector3D, rawPack, rawUnpack> *glassTrans;
Shared<VistaQuaternion, rawPack, rawUnpack> *glassRotate;
```

```
void Handler::InitGlass() {g = new GlassClient(new TCPUDP(
    tcport, udpport), host);

glassTrans = new Shared<VistaVector3D, rawPack,
    rawUnpack> ("glassTrans", vistaVec, sizeof(
        VistaVector3D), Global);

glassRotate = new Shared<VistaQuaternion, rawPack,
    rawUnpack> ("glassRotate", vistaQuat, sizeof(
        VistaQuaternion), Global);
}

void Handler::HandleEvent(VistaEvent *pEvent) {

    if (isVistaVRMaster == true) {
        if (isGlassMaster == true) {
            if (vistaVec != GetTranslation()) {
                vistaVec = GetTranslation();
                *glassTrans = vistaVec;
                glassTrans->sendUpdate();
            }
            if (vistaQuat != GetRotation()) {
                vistaQuat = GetRotation();
                *glassRotate = vistaQuat;
                glassRotate->sendUpdate();
            }
        }
        if (isGlassMaster == false) {
            glassTrans->getUpdate();
            glassRotate->getUpdate();
            vistaVec = glassTrans->getData();
            vistaQuat = glassRotate->getData();
            SetTranslation(vistaVec);
            SetRotation(vistaQuat);
        }
    }
}
```

O protocolo de comunicação utilizado nos experimentos permite que sejam realizadas transformações relacionadas à translação e rotação de câmera. Assim, dois métodos são criados para cada uma das transformações (*get* e *set*). O Código 2 apresenta os métodos *get* e *set*, relacionados à

translação.

Código 2: Métodos Get e Set para Translação

```
VistaVector3D Handler::GetTranslation() {
    return m_pVirtualPlatform->GetTranslation();
}

void Handler::SetTranslation(VistaVector3D v3Trans) {
    m_pVirtualPlatform->SetTranslation(v3Trans);
}
```

As transformações relacionadas à rotação são tratadas da mesma forma que as translações, diferenciando apenas no tipo de objeto utilizado. O Código 3 apresenta os métodos *get* e *set* para as transformações de rotação.

Código 3: Métodos Get e Set para Rotação

```
VistaQuaternion Handler::GetRotation() {
    VistaQuaternion qRot = m_pVirtualPlatform->GetRotation();
    return qRot;
}

void Handler::SetRotation(VistaQuaternion q4Rot) {
    m_pVirtualPlatform->SetRotation(q4Rot);
}
```

Algo importante de se definir é a hierarquia para os tipos de nodos: master, slave, client e serverpeer. Um nodo masterVista (VistaVRToolkit) pode não ser, obrigatoriamente, um nodo masterclientGlass (libGlass), visto que todo ambiente intra-aglomerado possui um nodo masterVista, responsável pela troca de mensagens entre os nodos de renderização (slavesVista); no caso do nodo masterclientGlass, só pode existir um em um dado momento, visto que ele realiza o envio de mensagens síncronas; já os nodos clientGlass (libGlass), solicitam as mensagens com os valores atualizados aos seus *peers*, posteriormente difundindo aos nodos slaveVista, por meio do nodo masterVista. Os nodos serverpeers são responsáveis por efetuar a comunicação entre os ambientes inter-aglomerado. O plugin libGlass utilizado para a troca de mensagens é o Shared plugin (mensagens síncronas). A Figura 7 apresenta um diagrama de sequência descrevendo este comportamento.

Neste modelo, como a mesma configuração de câmera é compartilhada, apenas um usuário pode interagir por vez. Assim, apenas um dos master nodes pode realizar transformações em um dado momento. Na Figura 7, fica claro este tipo de comportamento no modo como os objetos clientGlass são instanciados: o clientGlass_1 como sendo um masterclient e o clienteGlass_2 como sendo um client. Quando um usuário desejar interagir com o ambiente, este deve efetuar uma solicitação, que é feita por meio de uma mensagem do tipo Event (libGlass) enviada a todos os clientGlass nodes. A Figura 8 apresenta um diagrama de sequência que exemplifica a requisição de um nodo clientGlass para se tornar um masterclient.

Para ser ter o controle sobre as trocas de mensagens entre os nodos do aglomerado gráfico local, são utilizados os modelos de *callbacks* providos pela VistaVRToolkit. Para o masterClientGlass, o tipo de evento utilizado foi o VistaSystemEvent::VSE_POSTAPPLICATIONLOOP, assim, a cada quadro os dados atualizados são enviados ao outro aglomerado gráfico utilizando as variáveis libGlass (glassTrans e

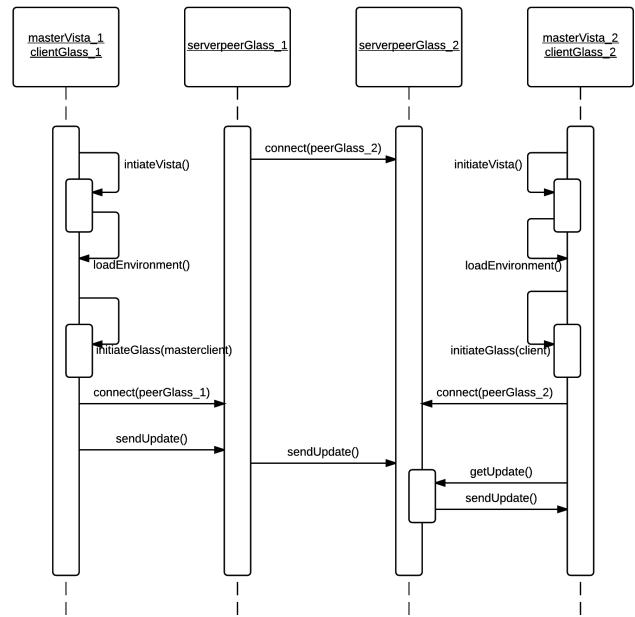


Fig. 7: Modelo Comportamental – VistaVRToolKit & libGlass

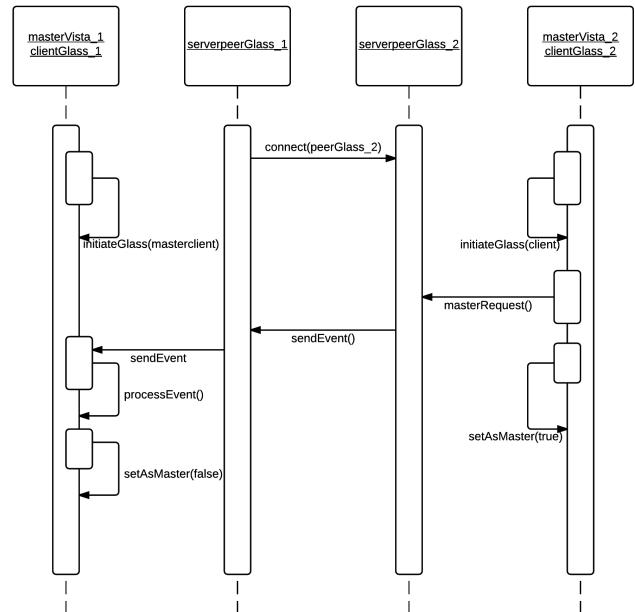


Fig. 8: Master – libGlass

glassRotate). Já para o clientGlass, o tipo de evento utilizado foi o VistaSystemEvent::VSE_PREAPPLICATIONLOOP. A diferença básica entre as duas transformações é o tipo de dados. No caso das translações é utilizado um vetor 3D, já para as rotações, é utilizado uma matriz 4D.

VI. DISCUSSÕES GERAIS

A integração de desenvolvimento utilizando duas soluções de *softwares* distintas foi realizada pelos autores. Como esperado, a solução permitiu o desenvolvimento de protótipos compatíveis com ambos ambientes de multiprojeção (AixCAVE e miniCAVE).

O VistaVRToolkit permite rápido desenvolvimento de protótipos, o que facilita ao extremo a concepção de novos ambientes virtuais por alunos ingressantes na carreira acadêmica. Assim sendo, os autores pretendem utilizar ambientes desenvolvidos utilizando a versão *opensource* da VistaVRToolkit disponibilizada no portal sourceforge.net [24] como solução local aos aglomerados gráficos.

Com relação aos resultados obtidos nos testes preliminares de conexão entre os ambientes, o estudo se mostrou eficiente. Por meio da solução desenvolvida, foi possível realizar visualização imersiva em colaboração entre os dois centros de pesquisas (RWTH e UFSCar). O estudo será continuado, visto que os autores irão utilizar a solução de *software* VistaVRToolkit em conjunto com a libGlass, o que permitirá a realização de experimentos futuros.

A análise da conexão é realizada por meio de um *plugin* Glass responsável por manter o QoS sobre as conexões [21]. Espera-se que, em um futuro próximo, outros artigos possam ser publicados em relação aos dados referentes à conexão entre o miniCAVE e a AixCAVE.

AGRADECIMENTOS

Os autores deste artigo agradecem o VRG – RWTH Aachen University e o Prof. Dr Torsten W. Kuhlen, por disponibilizar acesso à AixCAVE e ferramentas de desenvolvimento. Os autores também agradecem a CAPES, pelo apoio financeiro referente ao Processo PDSE/CAPES nº 2852/14-4, que possibilitou o período de pesquisa junto ao VRG – RWTH Aachen University.

REFERÊNCIAS

- [1] (2013) Secondlife. [Online]. Available: <http://secondlife.com/>
- [2] (2013) Activeworlds. [Online]. Available: <https://www.activeworlds.com/index.html>
- [3] (2013) Imvu. [Online]. Available: <http://pt.imvu.com/>
- [4] (2013) Eyegaze. [Online]. Available: <http://www.eyegaze.com/>
- [5] M. Okuda, Y. Karube, R. Matsukura, M. Watanabe, and T. Matsuzawa, “Development of a vizgrid volume communications environment,” in *Systems Modeling and Simulation*, K. Koyamada, S. Tamura, and O. Ono, Eds. Springer Japan, 2007, pp. 381–385.
- [6] J. Allard, V. Gouranton, L. Lecointre, S. Limet, E. Melin, B. Raffin, and S. Robert, “Flowvr: A middleware for large scale virtual reality applications,” in *Euro-Par 2004 Parallel Processing*, ser. Lecture Notes in Computer Science, M. Danelutto, M. Vanneschi, and D. Laforenza, Eds. Springer Berlin Heidelberg, 2004, vol. 3149, pp. 497–505.
- [7] B. Schaeffer and C. Goudeseune, “Szygy: native pc cluster vr,” in *Virtual Reality, 2003. Proceedings. IEEE*, 2003, pp. 15–22.
- [8] (2013) Chromium homepage. Chromium. [Online]. Available: <http://chromium.sourceforge.net/>
- [9] J. Allard, V. Gouranton, L. Lecointre, E. Melin, and B. Raffin, “Net juggler: running vr juggler with multiple displays on a commodity component cluster,” in *Virtual Reality, 2002. Proceedings. IEEE*, 2002, pp. 273–274.
- [10] C. Cruz-Neira, D. J. Sandin, T. A. DeFanti, R. V. Kenyon, and J. C. Hart, “The cave: audio visual experience automatic virtual environment,” *Commun. ACM*, vol. 35, pp. 64–72, June 1992. [Online]. Available: <http://doi.acm.org/10.1145/129888.129892>
- [11] B. B. Gnecco, M. de Paiva Guimaraes, and M. K. Zuffo, “Um framework para computação distribuída,” in *Simpósio Brasileiro de Realidade Virtual*. Ribeirão Preto: SBC, 2003.
- [12] T. A. DeFanti, J. Leigh, L. Renambot, B. Jeong, A. Verlo, L. Long, M. Brown, D. J. Sandin, V. Vishwanath, Q. Liu, M. J. Katz, P. Papadopoulos, J. P. Keefe, G. R. Hidey, G. L. Dawe, I. Kaufman, B. Glogowski, K.-U. Doerr, R. Singh, J. Girado, J. P. Schulze, F. Kuester, and L. Smarr, “The optiportal, a scalable visualization, storage, and computing interface device for the optiputer,” *Future Generation Computer Systems*, vol. 25, no. 2, pp. 114–123, 2009. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0167739X08001040>
- [13] F. Drolet, M. Mokhtari, F. Bernier, and D. Laurendeau, “A software architecture for sharing distributed virtual worlds,” in *Virtual Reality Conference, 2009. VR 2009. IEEE*, 2009, pp. 271–272.
- [14] (2015) Opensg. [Online]. Available: <http://www.opensg.org/>
- [15] D. C. Dias, A. La Marca, A. Moia Vieira, M. Neto, J. Brega, M. de Paiva Guimaraes, and J. Lauris, “Dental arches multi-projection system with semantic descriptions,” in *Virtual Systems and Multimedia (VSMM), 2010 16th International Conference on Virtual Systems and Multimedia*, ser. VSMM 2010, no. ISBN 978-1-4244-9027-1. Seoul: IEEE Xplore, oct. 2010, pp. 314 –317.
- [16] (2015) Aixcave at rwth aachen university. [Online]. Available: <http://www.itc.rwth-aachen.de/cms/IT-Center/Forschung-Projekte/Virtuelle-Realitaet/Infrastruktur/~fgqa/aixCAVE/>
- [17] M. P. Guimaraes, “Um ambiente para o desenvolvimento de aplicações de realidade virtual baseadas em aglomerados gráficos,” Ph.D. dissertation, Universidade de São Paulo, 2004.
- [18] T. van Reimersdahl, T. Kuhlen, A. Gerndt, J. Henrichs, and C. Bischo, “Vista: a multimodal, platform-independent vr-toolkit based on wtk, vtk, and mpi,” in *Proceedings of Fourth International Immersive Projection Technology Workshop (IPT2000)*, Ames, 2000.
- [19] M. Neto, D. Dias, L. Trevelin, M. de Paiva Guimarães, and J. Brega, “Unity cluster package – dragging and dropping components for multi-projection virtual reality applications based on pc clusters,” in *Computational Science and Its Applications – ICCSA 2015*, ser. Lecture Notes in Computer Science, O. Gervasi, B. Murgante, S. Misra, M. L. Gavrilova, A. M. A. C. Rocha, C. Torre, D. Taniar, and B. O. Apduhan, Eds. Springer International Publishing, 2015, vol. 9159, pp. 261–272. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-21413-9_19
- [20] D. R. C. Dias, “Sistema avançado de realidade virtual para visualização de estruturas odontológicas,” Master’s thesis, IBILCE/UNESP, São José do Rio Preto, 2011.
- [21] D. R. C. Dias, M. de Paiva Guimaraes, T. W. Kuhlen, and L. C. Trevelin, “A dynamic-adaptive architecture for 3d collaborative virtual environments based on graphic clusters,” in *The 30th ACM/SIGAPP Symposium On Applied Computing Proceedings*, 2015.
- [22] M. Schirski, A. Gerndt, T. van Reimersdahl, T. Kuhlen, P. Adomeit, O. Lang, S. Pischinger, and C. Bischof, “Vista flowlib - framework for interactive visualization and exploration of unsteady flows in virtual environments,” in *Proceedings of the Workshop on Virtual Environments 2003*, ser. EGVE ’03. New York, NY, USA: ACM, 2003, pp. 77–85. [Online]. Available: <http://doi.acm.org/10.1145/769953.769963>
- [23] M. de Paiva Guimarães, B. B. Gnecco, D. R. C. Dias, J. R. F. Brega, and L. C. Trevelin, “Graphical high level analysis of communication in distributed virtual reality applications,” *Procedia Computer Science*, vol. 51, pp. 1373 – 1382, 2015, international Conference On Computational Science, {ICCS} 2015Computational Science at the Gates of Nature. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1877050915011515>
- [24] (2015) Vista vr toolkit. [Online]. Available: <http://sourceforge.net/projects/vistavrtoolkit/>

Processamento de Linguagem Natural para a Avaliação de Situações de Emergência com Interface de Realidade Virtual

Matheus F. Sanches, Leonardo C. Botega, Allan C. M. Oliveira

Computing and Informations Systems Research Lab (COMPSI)

Centro Universitário Eurípides de Marília (UNIVEM), Marília, SP

matheussanches531@gmail.com, {botega, allan_oliveira}@univem.edu.br

Resumo —Consciência Situacional (*Situation Awareness - SAW*) é um conceito amplamente utilizado em áreas que demandam tomada de decisão crítica, como no domínio de gerenciamento de emergências. SAW está relacionada com o nível de percepção e entendimento que um indivíduo tem sobre eventos reais que ocorrem em cenários complexos, os quais devem ser geridos por sistemas críticos orientados a SAW. Uma questão desafiadora no design de tais sistemas, é determinar como o processo de aquisição pode ser desenvolvido para o enriquecimento de SAW, considerando ambientes com fontes de dados de inteligência humana (*Human Intelligence - HUMINT*), conhecidas por introduzir incertezas na análise da informação. No caso de analistas da Polícia Militar do Estado de São Paulo (PMESP), SAW é crucial para determinar quais recursos serão alocados ao atendimento de uma ocorrência. Neste contexto, o processamento de linguagem natural (*Natural Language Processing - NLP*) pode contribuir com o processo de obtenção e análise de dados de fala natural fornecidos pelos denunciantes, e fornecer ao sistema valiosos subsídios para a construção de um conhecimento situacional. Esse trabalho consiste no desenvolvimento de um processo de NLP capaz de obter, analisar e identificar informações relevantes ao domínio militar. Um estudo de caso é abordado para ilustrar a aplicabilidade do processo no contexto de avaliação de uma situação de roubo. Ao final, o produto do processo de avaliação de situações é apresentado utilizando técnicas de visualização de informações e um protótipo para o uso de técnicas de Realidade Virtual, demonstrando como o presente trabalho pode contribuir para a SAW de operadores da PMESP e consequentemente para a tomada de decisão.

Palavras-chave: consciência situacional, processamento de linguagem natural, sistemas de avaliação de situações.

Abstract - *Situational Awareness (SAW) is a concept widely used in areas that demand critical decision making, as in the emergency management area. SAW is related to the level of perception and understanding that an individual has about*

real events that occur in complex scenarios, which should be managed by critical systems oriented to SAW. A challenging point in the design of such systems is to determine how the acquisition process can be developed for SAW enrichment considering environments with data sources of human intelligence (HUMINT), known to introduce uncertainties in the analysis of information. In the case of analysts from Police of São Paulo (PMESP), SAW is crucial to determine which resources will be allocated to the care of a report. In this context, the natural language processing (NLP) can contribute to the process of obtaining and natural speech data analysis provided by the reporters, and provide valuable information to the system for the construction of a situational awareness. This work consists in developing a NLP able to obtain, analyze and identify relevant information to the military domain. A case study is discussed to illustrate the applicability of the process in the context of assessing a robbery situation. Finally, the product of situations assessment process is shown using information visualization and a prototype for the usage of Virtual Reality techniques, demonstrating how this work can contribute to SAW of PMESP operators and consequently for decision making.

Keywords—situational awareness, natural language processing, situation assessment systems.

I. Introdução

Sistemas de Avaliação de Situações (SAS) abastecidos por múltiplas fontes de dados heterogêneas e inteligência computacional, são recursos computacionais utilizados para reagir às mudanças ambientais e ajudar humanos especialistas a não somente desenvolver a percepção e o entendimento sobre o que está ocorrendo no ambiente em um determinado momento, mas também antecipar-se a eventos que podem emergir. Esta rotina é conhecida como Consciência Situacional (SAW), processo cognitivo que Sistemas de Avaliação de Situações buscam estimular utilizando rotinas de aquisição, inferência, avaliação e monitoramento de

informações (Endsley, 2001)(Llinas *et al*, 2004)(Kokar e Endsley, 2012).

Uma questão desafiadora para a comunidade de Avaliação de Situações, mais especificamente no domínio de gestão de emergências, é determinar como um processo semi-automatizado de análise de situações pode ser estruturado para acomodar o desenvolvimento de SAW sob dados e informações imperfeitas (Stanton,2001)(Salerno, 2002)(Blasch e Valin, 2012)(Laskeyetal, 2012).

Utilizando NLP para realizar a análise completa dos dados obtidos através das diversas fontes é uma maneira de melhorar e acelerar a tomada de decisão. Como o tempo para a tomada de decisão em situações de emergência é um fator extremamente importante foram estudados métodos para a análise em tempo real das informações enquanto ao mesmo tempo signifique menos erros e inconsistências. Através do NLP é possível determinar detalhes que um humano poderia deixar passar, como determinar o sexo de uma pessoa pela terminação das palavras, e apresentá-los de forma clara e intuitiva para a tomada de decisão.

Com dados mais precisos e completos gerados após esse processamento eles são introduzidos no sistema de tomada de decisão, que com informações mais completas e precisas tem melhores resultados, gerando uma tomada de decisão mais rápida, precisa e confiável.

Em cenários complexos, como os de gerenciamento de emergências, há várias fontes de dados disponíveis, tais como: áudio de chamadas à polícia ou bombeiros, mensagens em redes sociais, bancos de dados e imagens de câmeras de monitoramento. Cada escopo de aplicação tem particularidades das fontes e dos dados de entrada utilizadas para a realização da avaliação de uma situação. Neste processo, e mais adiante no estudo de caso, são abordadas gravações de áudio e mensagens provenientes de redes sociais, ambas categorizadas como fontes de dados HUMINT (inteligência humana). Ao final do processo busca-se obter objetos, atributos e situações preliminares, todos identificados e classificados de acordo com o domínio da aplicação.

Propõe-se com este trabalho um módulo de suporte a SAS, apresentando o desenvolvimento de um processo que, a partir de dados de áudio capturadas em língua Portuguesa, possa gerar informações relevantes para ajudar a identificar situações de emergência, classificando-as de forma coerente e disponibilizando- as para o uso do processo de avaliação e visualização com técnicas de Realidade Virtual.

Para que tal objetivo seja atingido, o processo de Aquisição e Pré-Processamento de Dados HUMINT é estruturado em estágios, tais como a obtenção dos dados, análise gramatical de sentenças, busca e identificação de informações relevantes e classificação de objetos e situações. Com esta rotina inicial, tal processo interno objetiva suportar preliminarmente o processo de avaliação de situações. As próximas subseções descrevem tais etapas em detalhes.

SAW, especialmente sistemas de emergência dependem diretamente da confiabilidade e qualidade das informações que estão sendo utilizadas para poder fornecer ao especialista que está utilizando o sistema, uma visão mais clara e precisa do que está acontecendo para que a tomada de decisão seja mais precisa e direta. Se informações incorretas ou incompletas são fornecidas ao especialista, a tomada de decisão fica comprometida. Dessa maneira este trabalho mostra maneiras de obter informações precisas para alimentar SAS com dados provenientes de fontes HUMINT.

II. Gerenciamento de Emergências

Sistemas de tomada de decisão para gerenciar emergências são desenvolvidos com base em características essenciais, tais como:

- A. Sistemas de tomada de decisão, baseados em consciência situacional, demandam de alta confiabilidade das informações, onde uma informação que apresenta limitações provocadas principalmente por problemas referentes à aquisição, podem causar inconsistências na tomada de decisão. Uma vez que algum tipo de inconsistência ou erro persiste durante todo o processo de avaliação de situação, graves falhas podem ser desencadeadas no mundo real. Apesar das informações para a tomada de decisão serem apresentadas a um operador que decide no que deve confiar, a influência das informações apresentadas pode gerar grandes problemas, caso forem apresentadas de maneira incorreta.
- B. Sistemas que tem SAW como base demandam de qualidade e confiabilidade das informações durante a análise, porém a análise não pode demorar para ter certeza ao apresentar um dado ao operador, que tomará a decisão baseado nos dados apresentados. Dados incompletos, incorretos ou tardios geram tomadas de decisões carentes de informações, assim situações que são muito complexas podem parecer pequenas e simples de serem resolvidas. Sistemas de tomada de decisão para situações de emergência demandam da resposta em um curto tempo de análise. Uma análise que demora um tempo arbitrariamente longo pode comprometer a tomada de decisão. Recursos alocados e decisões tomadas tardeamente podem resultar na ineficiência de todo o sistema.

III. A Avaliação de Situações de Emergência de Roubo: Estudo de Caso

A polícia da cidade de São Paulo (PMESP) recebe dezenas de milhares de ligações diariamente, contendo informações de todos os tipos, desde ligações pedindo informações sobre CEP de ruas, até denúncias efetivas de roubos e outros crimes.

As informações passadas para a polícia constituem um universo de dados que é impossível prever de maneira eficaz. Assim as análises estão orientadas a procurar apenas por informações relevantes a um crime em particular. Desta

maneira, o entendimento completo dos dados que são obtidos via denúncia seria momentaneamente impossível.

Para a realização dessa ação seria necessário ou um grande poder de processamento, ou um grande tempo para as análises das denúncias, algo que ainda não é contemplado de maneira viável. Todos os dados após serem processados são apresentados a um operador da polícia que tem como papel analisar os resultados obtidos e definir qual a melhor ação a ser tomada.

O desenvolvimento dessa pesquisa foi focado em uma das ocorrências mais recorrentes, o roubo na cidade de São Paulo.

Uma situação de Roubo é constituída pela existência de quatro elementos dentro de uma denúncia, sendo: o criminoso, vítima, objeto roubado e o local onde aconteceu o crime. Quando esses quatro elementos são encontrados é gerado como consequência uma situação de roubo.

As informações obtidas através da NLP são processadas para realizar a busca por objetos e atributos relevantes a esta situação. Quando dados relevantes são encontrados, os mesmos são empacotados em um objeto JSON, que é enviado aos próximos módulos de um processo completo de avaliação de situações, composto também por avaliação da qualidade das informações, fusão, representação e interface (visualização 2D e Realidade Virtual). A Figura 1 mostra o processo de avaliação de situações com as etapas de aquisição em destaque (cinza claro).

No estágio final, o processo completo de aquisição, composto pelas fases descritas a seguir, tem duração de execução aproximada de 02 segundos após os dados serem obtidos. Estes dados através da NLP são apresentados ao operador de polícia de forma 2D e 3D imersiva, utilizando técnicas de Realidade Virtual, para que ele possa ter um melhor entendimento da situação que está acontecendo no momento.

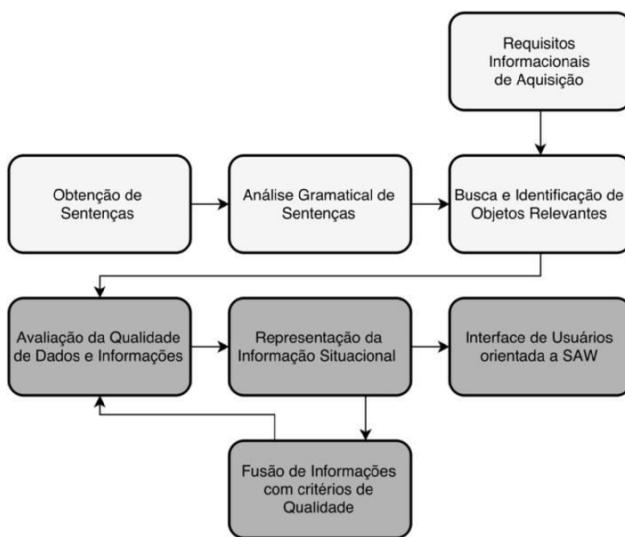


Figura 1. Processo de avaliação de situações de emergência, composto por aquisição, avaliação de qualidade, fusão de informação, representação e interface.

IV. Aquisição de Linguagem Natural Proveniente de Chamadas 190

O primeiro estágio da análise dos dados HUMINT consiste em realizar a extração da fala natural para texto.

Durante o processo de seleção da melhor ferramenta para realizar a transcrição da fala natural para texto foram testadas diversas ferramentas, como NLTK e Google Speech to Text.

Ao final dos ensaios a ferramenta escolhida foi a *SpeechTo Text* do *Google*. Uma vez que esta consegue realizar análise com uma precisão de 92% de acertos[7]. A escolha dessa ferramenta apresenta algumas vantagens, tais como:

- Grande precisão da análise, possibilitando 92% de acertos.
- Facilidade de implementação e uso, sendo usada com um serviço WEB.
- Alta velocidade na análise das palavras, com resultados em tempo real. A medida que se fala a frase a mesma já está sendo analisada e transcrita.

Em contrapartida, o uso de tal ferramenta apresentou as seguintes desvantagens:

- Vocabulário limitado. As palavras reconhecidas pela ferramenta não podem ser alteradas, assim algumas gírias ficam fora da análise até que seja identificada e indexada pelo Google.
- Dependência de conexão com a internet. Para as palavras serem analisadas elas precisam ser enviadas para os servidores deles e somente isso retornam como texto.

Em todos os ensaios realizados, a eficiência se mostrou bem acima da média quando comparada com as outras ferramentas. Reconhecendo inclusive nomes de ruas estrangeiras e palavras de outras línguas alternando com o texto.

V. Aquisição de Postagens de Rede Social

Atualmente uma grande parcela da população possui dispositivos capazes de acessar a internet. Onde muitas vezes relatam em redes sociais o que acontece ou o que viram antes mesmo de acionar os serviços de emergência disponíveis.

Dessa maneira uma importante fonte de dados poderia ser as redes sociais, como o Twitter, onde as pessoas postam algo logo após ele ter ocorrido.

Neste trabalho, módulos especializados na aquisição dos dados nessas redes sociais funcionam de modo a procurar por POSTS que possam remeter a situações de algum crime. Assim que essa é encontrada a localização do POST e o texto escritos são obtidos e tratados para a análise poder realizar sua função sem se preocupar com dados desnecessários.

Alguns dos pontos tratados são:

- Tratamento de gírias. As gírias escritas não são reconhecidas pelo analisador grammatical. Dessa maneira gírias como “vc” passa a ser “você”.
- Remoção da pontuação. A pontuação é removida para que o POST possa ser analisado da mesma forma que o texto obtido pela fala natural, uma vez que a fala natural não tem pontuação.

As postagens da rede social são avaliadas sob os mesmos processos na qual as denúncias 190 são submetidas, de acordo com a metodologia de Souza et al [6].

VI. Análise das Denúncias Adquiridas

Para realizar os testes apresentados neste artigo são utilizadas denúncias reais feitas a PMESP, que estavam inicialmente no formato de áudio, garantindo assim maior proximidade com o mundo real.

Após que o estágio de aquisição esteja concluído o sistema de análise tem acesso ao texto preparado que contém uma possível denúncia preparado para ser analisada.

O primeiro modelo de análise desenvolvido foi baseado um sistema unicamente de TAGs, SUB-TAGs e PALAVRAS. Onde a TAG significava uma categoria, a sub-tag significava o atributo dessa categoria e a palavra significava o item a ser procurado. Assim sempre que uma palavra analisada fosse igual a PALAVRA salva anteriormente o sistema emitiria um alerta disso e instanciaava um atributo dentro de um objeto contendo a palavra.

Exemplo de análise:

Sempre que a palavra “arma” era encontrada o atributo “arma” do objeto “criminoso” recebia “true”.

Esse sistema se mostrou muito eficiente com denúncias pequenas e pouco completas. Porém, com denúncias reais a dificuldade e a falsa instanciação dos atributos se tornava algo recorrente.

Para sanar os problemas encontrados na análise 1 foi introduzido um sistema baseado na formação das palavras em conjunto com o sistema de palavras-chave.

Essa análise é feita em 4 estágios.

1. Recebimento dos dados e análise grammatical.
2. Procura por informações relevantes
3. Instanciação de Objetos e atributos
4. Formatação dos dados e envio para o servidor.

Durante o estágio 1 os dados são recebidos por um web-service que controla o fluxo de dados dentro da aquisição.

Logo em seguida os dados são enviados ao web-service de análise grammatical que é feito pelo Cogroo. A função dessa análise grammatical é identificar os atributos de cada palavra,

como classe grammatical, variações e mais alguns itens. O Cogroo também consegue relacionar algumas palavras baseado na formação da frase. Facilitando assim durante a análise em buscas da palavras chave. Sabendo com quais palavras uma determinada palavra se relaciona o tempo de análise da sentença é reduzido significativamente. A Figura 2 apresenta um exemplo de saída de uma única palavra analisada pelo Cogroo.

```
{
  "tokens": [
    {
      "features": "F=S",
      "POSTag": "n",
      "lemmas": ["moto"],
      "lexeme": "moto"
    }
  ],
  "TAG": "NP"
}
```

Figura 2. Exemplo de saída de uma única palavra analisada pelo Cogroo.

Após o Cogroo realizar a análise inicial da frase os dados são enviados a outro web-service, que iniciará as análises em buscas das palavras-chave e suas relações.

Durante os estágios 2 e 3é dado início a busca por itens relevantes ao domínio militar, que foram definidos previamente através de pesquisas feita com profissionais da área.

Através da busca por palavras chave e a relação destas palavras é possível definir atributos para Objetos durante a análise. Utilizando esse sistema de análise, onde não é preciso saber qual a palavra que se espera, mas sim uma relação de palavras se tornou possível identificar palavras nunca antes analisadas ou desconhecidas.

Exemplo de relação: “fugiram”, “na”, “moto”. Quando as palavras “fugiram” e “na” são encontradas consecutivamente significa que a próxima palavra pode conter o meio de fuga dos criminosos, assim mesmo que a palavra “moto” não fosse uma palavra conhecida ela poderia ser analisada e o resultado seria que o atributo “fugaTipo” do objeto “Criminoso” seria “moto”.

Esse tipo de análise se mostrou muito mais eficiente e precisa que o método anterior, porém o trabalho de desenvolvimento é mais complexo, uma vez que para cada caso de análise um bloco de códigos precisa ser escrito, inserido e testado para garantir que esse não influencia negativamente nas análises já existentes.

Após esses dados serem analisados eles são encapsulados em um JSON e são enviados ao servidor e o próximo modulo, responsável pela fusão de dados, é avisado. Como resultado da

análise dos dados é gerado um objeto JSON (Figura 3) para ser enviado ao servidor que tem o seguinte formato:

Objeto -> Atributos -> [itens identificados]

```
{
  "id":1,"categoria":"roubo",
  "denúncias":[{
    "id":"1","categoria":"DESCONHECIDA","fonte":"190",
    "horaInicio": "2015-05-29 14:03:27","horaFim": "2015-05-29
    14:04:57", "hora":"14:04:57","gps":[-22.2208],"-
    22.2208],"-
    22.2208}],
  "origemMensagem":"14 3434-3434",
  "texto":"boa noite acaba de acontecer um roubo de carro aqui
  na domingos setti com a luis vives dois caras de moto apontaram
  uma arma para o motorista de um mercedes preto e mandaram sair
  do carro sem levar nada os dois fugiram acelerando em
  dire\u00e1o 00e7u00e3o ao metr\u00f4 00f4 klabin um deles usava uma
  camiseta vermelha e o outro tinha um bon\u00e9 preto",
  "eventos":"Adicionado C e L","certeza":0.0}],
  "situação":{
    "criminoso":{
      "sexo":"M","status":"fugindo",
      "fugaTipo":"acelerando","fugaPara":"metr\u00f4 00f4 klabin",
      "corVestimenta":["vermelha","preta"],"arma":"arma",
      "palavraRef": "caras","veiculo":"moto",
      "vestimenta":
      ["camiseta","bon\u00e9"]},
    "completude":42.85,"consistencia":null
  },
  "objeto":{
    "descricao": "carro","marca": "mercedes",
    "completude":42.85,"consistencia":null
  },
  "vitima": {
    "palavraRef": "motorista",
    "completude":42.85,"consistencia":null
  },
  "local": {
    "logradouro": "domingos setti ; luis vives dois",
    "gps": [-22.2208],"-
    22.2208],
    "completude":42.85,"consistencia":null
  },
  "certeza":0.0,"completudeTemporal":100,"eventos":["updTCrimino
  so","updLocal","addVitima"],"ignorarFlag":false,"ignorarMensagem":null
}
```

Figura 3. JSON com a situação identificada após a análise de um caso de teste baseado em uma denúncia real.

VII. Continuidade do Processo de Análise de Situações

Após os dados serem enviados ao servidor, os módulos seguintes são avisados para que a classificação, representação e apresentação da situação inferida seja executada.

O módulo de avaliação de qualidade realiza a análise dos dados logo após a inferência de uma nova situação, produzindo índices que quantificam a incerteza e qualificam a informação situacional.

Para detalhes sobre como é realizada a avaliação da qualidade da informação o leitor deve consultar o trabalho de Souza *et al* [6]

O módulo de fusão de dados realiza a busca por denúncias que tem características em comum e que possam contribuir para a geração de uma situação mais completa. Assim, duas denúncias distintas, que contém informações diferentes, porém dados complementares de uma mesma situação, podem ser integradas e gerar uma situação nova e mais completa.

Após estes processos, a situação é apresentada para o operador da polícia, para que seja feita a análise da situação e a tomada de decisão.

VIII. Interfaces de Visualização

Adquirir consciência situacional (SAW) é um processo desafiador. Estimular a percepção e promover a melhor compreensão dos elementos que compõem um cenário real muitas vezes torna-se um procedimento complexo. Em situações em que há risco à vida ou ao patrimônio, como em eventos de roubo, há necessidade de que o especialista conheça as entidades envolvidas e seus estados, para assim ter subsídios para uma tomada de decisão eficiente.

Muitas vezes tais entidades são complexas, dinâmicas e variadas, inferidas por procedimentos automatizados e também pelo humano operador. O objetivo da visualização 2D e 3D com Realidade Virtual, é agregar meios alternativos para a realização de análise visual da situação. A Figura 4 apresenta o primeiro dos métodos, composto por um grafo hierárquico que representa a situação de roubo, seus componentes (vítimas, criminosos, locais e objetos de roubo) e os atributos que descrevem tais componentes, organizados de dentro para fora, onde os atributos são as folhas do grafo. As cores indicam o nível de qualidade das informações. Cores próximas ao vermelho indicam baixa confiança do sistema nas informações inferidas e cores próximas ao verde indicam um nível superior de certeza nas informações.

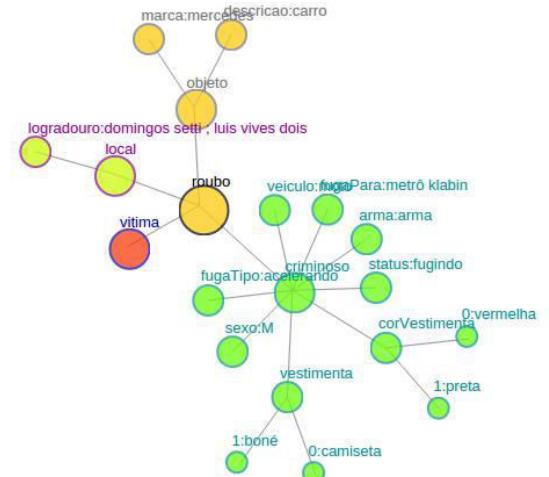


Figura 4. Representação gráfica da situação da Figura 3. Nós folha representam os atributos dos objetos identificados.

Em complemento, busca-se produzir um módulo de imersão e interação de especialistas com as informações de entidades, como pessoas, objetos e locais, dinamicamente geradas pelo sistema e pelo próprio operador a fim de prover melhores subsídios para a orientação de decisões.

Os ambientes dos eventos de emergência, inicialmente do roubo identificado e mostrado pela Figura 3, são reconstituídos em 3D para dar suporte ao emprego de técnicas de RV, para possibilitar que o especialista faça inferências sobre as informações situacionais que são estratégicamente posicionadas no ambiente. A Figura 5 apresenta a mesma situação da Figura 4 no ambiente 3D reconstruído a partir das informações situacionais da Figura 3 e com o uso da API Street View do Google, gerado em tempo-real pela extração das informações do objeto JSON (Figura 3).



Figura 5. Representação gráfica da situação da Figura 3 em formato de cena reconstruída à partir da interpretação automática do objeto JSON. Neste caso dados sobre os criminosos e o objeto roubado são automaticamente inseridos no local também identificado pelo NLP.

Futuramente, serão utilizados dispositivo de visualização não convencionais para imergir o especialista no local do evento de forma realista.

IX. Conclusões

Este trabalho teve como objetivo apresentar o desenvolvimento de um processo de NLP motivado por situações de emergências, mais especificamente de eventos roubo. Adicionalmente, foi ilustrado como esta abordagem se relaciona com outros módulos de um processo completo de avaliação de situações.

Os resultados obtidos com a análise de uma denúncia de roubo apresentaram os objetos e atributos identificados e caracterizados. Tais resultados foram também representados de duas maneiras: através de uma visualização em grafo bidimensional e utilizando uma cena 3D reconstruídas e

povoada com os elementos identificados pelo processo apresentado. Esta abordagem apresentou que o processo desenvolvido é capaz de fornecer subsídios suficientes para a composição de cenários de crime para serem explorados por técnicas de Realidade Virtual e assim complementar os processos de análise de situações, até o momento dependentes de visualizações 2D.

Considerando o processo de NLP, realizar a análise de um tipo específico de crime é algo viável. Porém, quando se enfrenta o desafio dos inúmeros crimes possíveis, a tarefa se mostra muito mais complexa.

A identificação de Objetos e Atributos em comum podem ser reutilizadas na análise de outros crimes, mantendo-se a busca por localidades, por exemplo. Porém para cada tipo de crime existem particularidades que devem ser analisadas e módulos de análises específicos devem ser preparados.

Durante o relato de uma denúncia, considerando que são dados HUMINT, existe a possibilidade do relato mudar repentinamente e passar a fornecer informações sobre outras situações. Esse tema é tratado como imprevisibilidade.

A imprevisibilidade não foi considerada no estágio atual de análise, uma vez que dados não esperados não são classificados *a priori*. Para realizar a análise contemplando a imprevisibilidade, é necessário um estudo mais aprofundado sobre este tema.

Como trabalhos futuros, para melhorar a identificação dos dados e a tomada de decisão militar, será proposto um módulo capaz de entender a estrutura da língua portuguesa de maneira a aprender a organizar os dados recebidos, visando melhorar o reconhecimento das palavras.

Desta maneira, não seria necessária uma alteração manual no sistema a cada novo tipo de denúncia (um outro crime, por exemplo). Esta nova abordagem deve também aprender com os dados que o sistema recebe diariamente e ser capaz de analisar os dados mesmo que inconsistentes, realizando assim um tratamento da imprevisibilidade.

Referências

1. Endsley, M. R. - The Challenge of the Information Age. In: Proceedings of the Second International Workshop on symbiosis of humans, artifacts and environment, Kyoto, Japan (2001)
2. Kokar, M. M., Endsley, M. R.: Situation Awareness and Cognitive Modeling. In: IEEE Intelligent Systems, vol. 27, no. 3, pp. 91–96, May. (2012)
3. Llinas, J., Bowman, C., Rogova, G., Steinberg, A.: - Revisiting the JDL data fusion model II. In: 7th International Conference on Information Fusion (2004)
4. Blasch, E.: High Level Information Fusion (HLIF): Survey of models, issues, and grand challenges. In: IEEE A&E Systems magazine, pp. 4–20 (2012)
5. Botega, L., Ferreira, L. C. N., Oliveira, P., Oliveira, A., Berti, C. B.: User Interface for Enhancing Situational Awareness in Emergency Management System. In: 17th International Conference on Human-Computer

- Interaction, Los Angeles. Lecture Notes in Computer Science (LNCS), (2015)
- 6. Souza, J. Botega, L., Santarém Segundo, J.E., Berti, C.: Conceptual Framework to Enrich Situation Awareness of Emergency Dispatchers. In: 17th International Conference on Human-Computer Interaction, Los Angeles. Lecture Notes in Computer Science (LNCS), (2015)
 - 7. Ciprian Chelba, Dan Bikel, Maria Shugrina, Patrick Nguyen, Shankar Kumar: Large Scale Language Modeling in Automatic Speech Recognition

Interação Através das Mão em um Ambiente de InfoVis 3D

Hand-based Interaction in a 3D InfoVis Environment

Nikolas Carneiro*, Brunelli Miranda†, Carlos Santos*, Tiago Araújo*, Anderson Marques*, Bianchi Serique*†

*Programa de Pós-Graduação em Ciência da Computação - PPGCC

†Faculdade de Computação – FACOMP

Universidade Federal do Pará, Belém, Pará, 66075-110

Resumo — Um ponto importante nos estudos em Visualização de Informação (InfoVis) são técnicas de interação, onde visa-se melhorar e propor novos mecanismos para configurar e explorar técnicas de InfoVis. Em visualizações tridimensionais a interação é ainda mais desafiadora, devido a falta de padrões de interação e questões como oclusão e navegação. Este trabalho propõem técnicas de interação baseadas em rastreamento de gestos e movimentação das mãos para exploração de InfoVis em 3D. O protótipo apresenta como estudo de caso a técnica de dispersão de dados tridimensional com uma base de dados simplificada de veículos. Com protótipo apresentado é possível realizar Zoom/Pan, Rotação, Filtro e Seleção na visualização apenas com a utilização de gestos e widgets inseridos no ambiente 3D.

Palavras-Chave — *Visualização de Informação; Rastreamento de Mão; Interação; 3D*

Abstract — An important point in Information Visualization (InfoVis) studies are techniques of interaction, which aims to improve and propose new mechanisms to configure and operate InfoVis techniques. In three-dimensional visualizations the interaction is even more challenging due to lack of interaction patterns and issues such as occlusion and navigation. This paper proposes interaction techniques based on gesture tracking and movement of hands for 3D InfoVis configuration and exploration. The prototype developed has as a case study a three-dimensional scatterplot technique with a simplified vehicle dataset. With the prototyped users can Zoom/Pan, Rotation, Filter and Selection in the visualization only with the use of gestures and widgets inserted in the 3D environment.

Keywords — *Information Visualization; Hand Tracking; Interaction; 3D*

I. INTRODUÇÃO

É comum perceber sistemas de Visualização de Informação (InfoVis) como dois componentes interligados, o componente de representação gráfica e o componente de interação. Esta divisão é identificada por Yi *et al.* [1] e por ser visto na interface gráfica de sistemas de InfoVis, ainda que estas estejam intrinsecamente relacionadas.

Separar a visão e os controles em ferramentas de InfoVis pode aumentar o esforço cognitivo em algumas situações, pois o usuário pode não perceber como se deram as modificações, uma vez que as ações podem ser realizadas em *widgets* distantes da visualização, e ter uma área dedicada somente a

controles reduz o espaço disponível para renderizar a visão, o que pode gerar maior oclusão e perda de detalhes na representação gráfica.

Essa questão já foi abordada de diversas maneiras, como a utilização de animação para mostrar as mudanças provenientes da interação ou a adição de visualizações informativas nos componentes de interação [2] para manter o usuário informado.

Uma possibilidade minimizar os efeitos negativos de ter áreas de visão e controle separados é minimizar, ou eliminar, a área de configuração trazendo o maior número de interações possíveis para dentro do espaço da visão, preferencialmente diretamente sobre a visualização. Isso indica não apenas transportar componentes de interface para a área da visão, indica tentar fazer com que as interações sejam o mais naturais possíveis quando ocorrendo diretamente sobre a visualização, como tocar um item para selecioná-lo ou fazer um movimento de “pinch” para dar zoom.

Com a disponibilização de novas tecnologias, como telas de toque e dispositivos de rastreamento de movimentos, é possível explorar ainda mais os mecanismos de interação no sentido de torná-los mais naturais e potencializar o ganho cognitivo na interação com InfoVis. Estes mecanismos também permitem explorar, no caso dos rastreadores de movimentos, ambientes 3D, propiciando que ferramentas de visualização tenham tanto a visão quanto a interação em três dimensões. Entretanto, acrescentam suas próprias particularidades que precisam ser tratadas no desenvolvimento de uma ferramenta de InfoVis, como a representação do mecanismo de interação e tratamentos de erros e sensibilidades excessivas que podem ocorrer.

Este trabalho visa explorar e propor técnicas de interação em ambientes totalmente 3D (visão e interação) baseadas em rastreamento de gestos. Essa proposta será construída através de um processo iterativo incremental, de forma que suas fases não sejam apenas complementares, mas também sirvam para validar, bem como corrigir se necessário, as fases anteriores. As fases previstas neste processo são: revisão de literatura, protótipo preliminar, avaliação de especialista, testes preliminares e testes de desempenho e com usuários, onde as duas primeiras são mostradas neste trabalho.

Na seção 2 há uma definição de InfoVis, na seção 3 são indicados trabalhos que nortearam o desenvolvimento deste, na

seção 4 é apresentado o protótipo desenvolvido e na seção 5 são feitas algumas considerações finais sobre o trabalho.

II. VISUALIZAÇÃO DE INFORMAÇÃO

Pode-se definir visualização como um tipo de comunicação da informação que utiliza representações gráficas [3]. Imagens têm sido utilizadas como forma de comunicação muito antes da linguagem escrita. Uma imagem simples pode conter muitas informações simultaneamente, e pode ser processada muito mais rapidamente por um ser humano quando comparada com uma página de texto (palavras). A razão disso está no sistema de percepção humana que processa as diversas informações da imagem em paralelo, enquanto a interpretação de um texto é limitada pelo processo sequencial de leitura do mesmo. Imagens podem também ser independentes de uma linguagem local, como um gráfico ou um mapa que podem ser entendidos por um conjunto de pessoas sem uma língua comum [4].

De acordo com Card *et al.* [5], visualização de informação é o uso de um ambiente computacional interativo, que possibilita representação visual de dados abstratos para ampliar a cognição. Este ambiente computacional interativo dá-se quando uma visualização ocorre em um computador, onde a esta pode ser manipulada de forma simples, direta e livre, incluindo ações como filtro de dados, navegação de dados, etc. A representação visual da informação é utilizada para representar atributos de dados abstratos tais como localização, tamanho, forma, cor para formar a imagem que permita a visualização de padrões, tendências e exceções. Dados abstratos dizem respeito a informações de dados quantitativos, processos e relacionamentos entre os dados. E por fim, o conceito de cognição ampliada procura, através da visualização e interação com os dados, estender a habilidade do usuário compreender de forma mais fácil os dados e seus relacionamentos. Estes dados podem ter várias dimensões e a utilização de visualizações em ambientes 3D permitem representar múltiplas dimensões de dados de maneira mais natural.

III. TRABALHOS RELACIONADOS

No desenvolvimento deste trabalho buscou-se encontrar na literatura estudos sobre interação em InfoVis com foco em interação 3D e interação por gestos ou movimentos. Dentre os trabalhos encontrados percebeu-se que há avaliações da interação pelo toque [6] ou pela representação de dispositivo apontador em um ambiente 3D. Para tal, selecionou-se alguns trabalhos para basear este, como apresentados abaixo.

Yi *et al.* [1] caracteriza sete diferentes técnicas de interação em ferramentas de InfoVis, baseado na intenção do usuário ao interagir com ferramentas, sendo estas:

- **Seleção:** a capacidade de marcar um ou mais itens com uma característica visual que os diferencie dos itens não selecionados. A seleção aparentemente funciona mais como interações que antecedem outras interações, permitindo uma identificação mais facilitada do comportamento dos itens selecionadas em resposta as interações subsequentes;
- **Exploração:** a possibilidade de analisar um subconjunto dos dados sem alterar a representação

gráfica dos mesmos e sem alterar os dados mostrados, ao exemplo das interações de *pan* e movimento de câmera;

- **Reconfiguração:** são técnicas pelas quais se altera a configuração espacial das visualizações sem alterar a codificação visual dos itens, a exemplo de operações de *sorting*;
- **Codificação:** são técnicas para alterar a representação visual dos itens (a exemplo de trocar o atributo que define a cor ou a cor que representa determinado atributo) ou a própria visualização (a exemplo de trocar um gráfico de pizza por um *treemap*);
- **Abstração:** são técnicas para manipular o nível de abstração das representações apresentadas, indo das informações apresentadas em um nível de zoom, passando pela navegação hierárquica que revela detalhes à medida que vai avançando (*drill-down*) chegando até a apresentação textual dos valores de um item;
- **Filtragem:** são técnicas pelas quais o usuário pode escolher o conjunto de dados a ser representado na visualização através de um ou mais critérios, sendo que dependendo da técnica aplicada os itens filtrados podem ser totalmente removidos da visualização ou apenas codificados de forma que sejam menos chamativos para o usuário. A responsividade do sistema é importante em filtragem e o conceito de *dynamic queries* (especificação de condições através da interface gráfica gráficas com resposta instantânea a interação) é um assunto bastante abordado na literatura;
- **Realce:** são técnicas que se referem a ressaltar a relação entre dois ou mais itens visíveis ou tornar visível um item relacionado a outro item, ou conjunto de itens, de interesse. Essas técnicas são mais usuais quando múltiplas visões são utilizadas, pois neste contexto pode ser difícil relacionar um mesmo item em diferentes visões, mas também podem ser aplicadas em visões únicas, por exemplo, ressaltando os vizinhos de um nó em um grafo.

[7] apresenta uma técnica para guiar o usuário quando realizando interações sensíveis a imprecisões do mecanismo de interação ou do usuário. Neste trabalho é indicado que determinados mecanismos, como o sensor de detecção e movimentos das mãos Leap Motion [8], necessitam de ajustes de interação devido a sensibilidade dos mesmos e do fator de *jitter* [7], aqui definido como a falta de precisão humana (ao por exemplo não conseguir ficar com o dedo parado sobre um item pequeno), como a predição de interações.

IV. PROTÓTIPO

A. Tecnologias Utilizadas

No desenvolvimento do protótipo buscou-se agregar tecnologias que agregassem as seguintes características: desempenho em renderização, capacidade de representar elementos em 2D e 3D, possibilidade de integração com diferentes meios de entrada (mouse, teclado, sensores, etc.) e,

preferencialmente, uma tecnologia difundida que viabilize documentação acessível com baixo custo de desenvolvimento.

Para entrada de gestos o dispositivo escolhido foi o Leap Motion, um sensor de movimento das mãos capaz de identificar ambas as mãos de um usuário e ainda outras mãos (ainda que devido ao espaço de alcance do sensor, a utilização de múltiplas mãos possa se tornar propensa a erros), identificação individual dos dedos de cada mão, além de identificar ferramentas e um conjunto pré-definido de gestos (que pode ser estendido através da adição de novos *scripts*). A entrada no protótipo se dá por movimento das mãos nas ações de clicar (tocar com o indicador) em um item, segurar e mover, bem como alguns gestos definidos para a ferramenta.

Para o motor da visualização (dispersão de dados), carregamento da base, desenho e controle dos *widgets* de interface foram utilizados scripts na plataforma Unity. O Unity é uma *engine* multiplataforma de desenvolvimento de jogos, capaz de portar conteúdo para PCs, consoles, dispositivos móveis e para web[9]. O Unity é utilizado no protótipo tanto como agregador de funcionalidades (como reconhecer e pegar a saída do Leap Motion) como é responsável pela lógica de desenho, carregamento da base, identificação de gestos e colisões através de scripts desenvolvidos no decorrer deste trabalho.

B. Arquitetura

Uma forma de mostrar como funciona o protótipo implementado é demonstrar como ocorre o fluxo dos dados dentro do mesmo durante uma execução. Este fluxo pode ser visto na Figura 1 e é descrito a seguir.

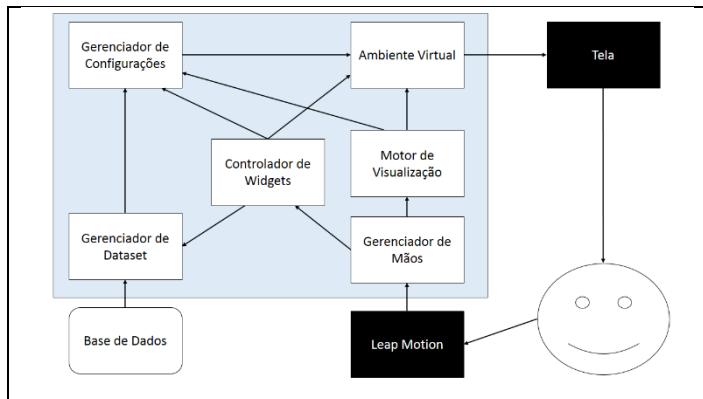


Fig. 1. Fluxo dos dados dentro do protótipo.

Na Figura 1, a área com preenchimento claro é a área do Unity que contém os códigos que foram gerados no decorrer deste trabalho, as duas caixas com preenchimento escuro são dispositivos físicos e o retângulo com os retângulos arredondados representa a fonte dos dados.

A Base de Dados é o fonte estática dos dados, ou seja, ela não sofre alterações nos dados, nem adição ou remoção destes enquanto a aplicação está sendo executada. Atualmente esta fonte de dados é um arquivo no disco rígido da máquina onde a aplicação está sendo executada, entretanto como a uma camada de software entre a Base de Dados e o resto do sistema, suporte a novas fontes de dados serão implementadas futuramente.

O Leap Motion é o sensor de rastreamento das mãos e serve como entrada do sistema, os dados captados a partir dele na forma de frames são passados a aplicação através do Gerenciador de Mão, conforme a Figura 1.

A Tela é o monitor do usuário e serve como saída do sistema, não há um requisito mínimo de tela (nem em tamanho ou resolução) definido para utilização do aplicativo, entretanto ao fazer uso de uma tela com um tamanho físico maior, e uma resolução maior, é possível visualizar mais itens ou ver mais detalhes sobre estes com um esforço menor do que o apresentado em uma tela menor.

O Gerenciador de Dataset é responsável tanto pela leitura da base quanto por prover uma representação lógica dela para os demais componentes do sistema. Ele mantém esta representação e reflete nela as mudanças provenientes de interações como Filtragem.

O Gerenciador de Mão é uma das entradas do sistema e é responsável por capturar continuamente os frames detectados pelo Leap Motion. Ele avalia estes frames para verificar se algum dos gestos pré-definidos na ferramenta foi realizado e identificar em que área da visualização está, ou não, ocorrendo as interações.

O Controlador de Widgets é responsável pela lógica (posição, estado, etc.) de componentes que não pertencem intrinsecamente à visualização. Estes componentes são auxiliares e podem ser empregados tanto em tarefas de Exploração, Reconfiguração, Codificação e Filtragem, quando projetados para interação ou no processo de Abstração, quando empregados com funções informativas (como as legendas de estado na interface). As mudanças realizadas sobre os widgets da interface são repassadas ao Gerenciador de Configurações bem como este informa o Gerenciador de Widgets de mudanças feitas diretamente sobre a visualização que precisem ser refletidas nos widgets, permitindo assim a implementação de widgets mais informativos futuramente, a exemplo dos componentes encontrados em [2].

O Motor de Visualização tem a função de representar graficamente a base de dados e os estados a ela atribuídos pelas interações do usuário (zoom, filtros, rotação, *pan*, etc.), ou seja, refletindo todos os processos de interação apontados por [1], além disso operações de Seleção, Reconfiguração, Abstração e Realce podem ser efetuadas diretamente neste componente, e assim, diretamente sobre a visualização. O Motor de Visualização implementado no protótipo utiliza apenas a técnica de dispersão de dados, no entanto as definições do script permitem que técnicas futuramente adicionadas utilizando os moldes já definidos.

O Gerenciador de Configurações guarda as configurações realizadas e propaga essas configurações para componentes que precisem ser notificados da alteração realizada. Sendo que estas configurações podem originar-se tanto da visualização quanto dos widgets interativos.

Ambiente Virtual é onde os estados definidos nos demais componentes são agregados para gerar uma representação coesa ao usuário, além dos estados da visualização e dos widgets. O Ambiente Virtual também agrupa os elementos que

o usuário vê, para o auxílio da percepção do ambiente, como iluminação e sombras.

C. Funcionalidades

Esta seção apresenta as funcionalidades do protótipo proposto para a interação baseada em movimentos das mãos e gestos. Para fins de exemplificação, foi utilizada uma base de dados de carros simplificada, com 14 atributos e 191 itens.

A visualização apresentada com o protótipo é o Dispersão de dados 3D, com as cores configuradas para o atributo “Origem do veículo” da base de dados, a forma configura da para o atributo “Combustível”, o eixo X configurado para o atributo “Potência do motor”, o eixo Y configurado para o atributo “RPM” e o eixo Z configurado para o atributo “Número de cilindros”.

A Figura 2 apresenta o menu inicial do protótipo sendo iniciado a partir de um gesto predefinido, que neste caso é colocar a mão esquerda aberta sobre o Leap Motion, sendo que o polegar deve estar aberto a aproximadamente 90 graus. O menu aparece entre o polegar e indicador da mão esquerda e se movimenta (translação e rotação) de acordo com a movimentação da mão esquerda. Este menu apresenta quatro opções que podem ser acessadas tocando com a mão direita sobre uma das opções. As opções disponíveis são: Zoom/Pan, Rotacionar, Filtrar e Selecionar a Visualização.

Ao selecionar uma das opções apresentadas pelo menu inicial, é apresentado ao usuário a ferramenta correspondente ao que foi selecionado. Quando o usuário seleciona a primeira opção (Zoom/Pan) ele pode ampliar, diminuir e transladar a visualização para ter um melhor ponto de vista da visualização.

A operação de Pan é realizada ao colocar a mão direita apenas com o dedo indicador esticado sobre o Leap Motion, de acordo com que o usuário movimenta a mão para cima, baixo e lados a visualização acompanha seu movimento. A operação de Pan é parada quando o usuário retira a mão para fora do alcance do Leap Motion ou quando estica outro dedo além do indicador.

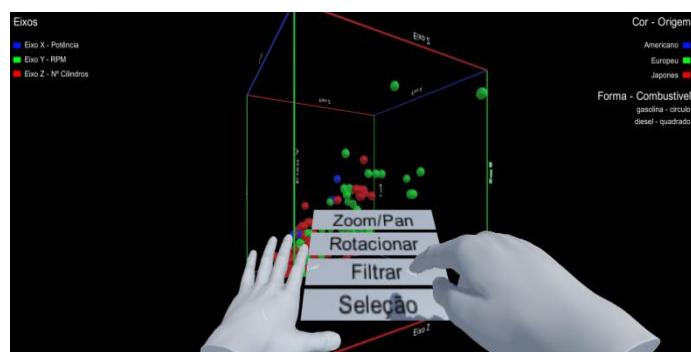


Fig. 2. Menu inicial do protótipo.

A operação de Zoom é semelhante à operação de Pan, entretanto esta é realizada quando o usuário apresenta abas as mãos com os dedos indicadores esticados. À medida que as duas mãos se afastam uma da outra a visualização é ampliada proporcionalmente, e a medida que as mãos se aproximam uma da outra a visualização é encolhida.

As operações de Zoom e Pan são importantes para a realização das tarefas de visualização, pois viabilizam ao usuário uma visão geral do conjunto de dados que está sendo explorado e também permite que o usuário navegue neste conjunto de dados em busca de padrões, pontos discrepantes ou certa particularidade que julgue importante.

A segunda opção apresentada no menu inicial do protótipo é a rotação, que permite ao usuário rotacionar a visualização em nos três eixos. A operação de rotação é realizada através de uma esfera que possui a indicação dos eixos na visualização.

A Figura 3 (a) mostra a esfera responsável pela interação de rotação da visualização. A interação inicia a partir da colisão do modelo da mão do usuário com a esfera, e dependendo da movimentação da mão, quando em contato com a esfera, a rotação é aplicada tanto na esfera quanto na visualização. Sendo assim, o usuário terá uma melhor noção de qual rotação o usuário está aplicando sobre a visualização.

A operação de rotação é importante em visualização da informação, especialmente quando se trata de uma visualização 3D, pois neste tipo de visualização podem ocorrer muitas oclusões, que podem ser mitigadas quando o conjunto de dados é rotacionado. Outro fator importante da rotação é proporcionar um melhor ponto de vista para o usuário, que pode configurar qual o melhor ângulo de visão dos dados.

As operações de Zoom, Pan e rotação são primordiais em um ambiente tridimensional, pois ajudam o usuário a navegar neste ambiente, explorando inclusive pontos que antes estavam oclusos. A realização dessas operações através de interfaces tradicionais como mouse e teclado, pode ser difícil de aprender e assimilar no ambiente 3D, o que torna o estudo de novas interfaces e interações promissor, no sentido de que estes estudos podem revelar métodos mais eficientes e intuitivos de interação com ambientes tridimensionais. Este protótipo viabiliza que a interação ocorra no próprio ambiente tridimensional o que pode ser mais intuitivo e fácil de aprender.

A Figura 3 (b) mostra o menu inicial com a opção de rotacionar selecionada, indicando para o usuário qual tarefa ele está realizando no momento. Este menu pode ser aberto a qualquer momento quando o usuário deseja realizar outra operação. O último estado da visualização é mantido para que o usuário possa realizar várias tarefas consecutivas.

A partir da Figura 3 (b) foi selecionado a opção filtro, que permite ao usuário realizar filtros no conjunto de dados apresentado refinando este conjunto e permitindo que o usuário remova itens que não está interessado deixando apenas os dados de relevantes para ele no momento.

Para realizar a opção de filtro pelo protótipo proposto, o usuário deve interagir com *widgets* tridimensionais que ficam dentro do ambiente virtual. O primeiro *widget* apresentado permite que o usuário escolha qual dos atributos da base podem ser filtrados, outros dois possíveis *widgets* são apresentados após a seleção de um atributo da base. Esses *widget* variam de acordo com a natureza do atributo selecionada no primeiro *widget*.

Este protótipo possui dois tipos de filtro: categórico e contínuo. Quando os dados de um atributo são contínuos, geralmente com dados numéricos de ponto flutuante é apresentado para o usuário um *widget* tridimensional de *slider* no qual o usuário pode diminuir ou aumentar um limite superior e inferior para os dados.

Quando os dados de um atributo são categóricos, geralmente dados que podem ser listados, é apresentado ao usuário uma lista de valores que podem ser habilitados ou desabilitados, fazendo com que todos os itens da base que contenham os valores desabilitados para um determinado atributo seja removido da visualização. O usuário pode habilitar um item desabilitado quando quiser, fazendo com que os itens filtrados apareçam novamente.

A Figura 4 mostra o primeiro *widget* mais a esquerda com o atributo Origem selecionado. Como o atributo Origem nesta base de dados pode ser enumerado e não é uma atributo com valores ponto flutuante é apresentado uma lista de valores, que

podem ser habilitados e desabilitados para a realização do filtro.

No caso da Figura 4 todos os itens que possuem o valor Europeu no atributo Origem foram removidos da visualização, como os itens de Origem Europeu estavam configurados com a cor verde, todos os itens de cor verde foram removidos da visualização.

A operação de filtro em visualização da informação é importante para que o usuário concentre-se apenas em uma parcela do conjunto de dados, especialmente em visualizações tridimensionais, em que alguns itens menos importantes para o usuário podem cobrir itens mais importantes. A possibilidade de o usuário realizar essas interações dentro do ambiente tridimensional faz com que ele não precise recorrer ao mouse e teclado, podendo interagir completamente através do dispositivo de rastreio das mãos, no caso o Leap Motion.

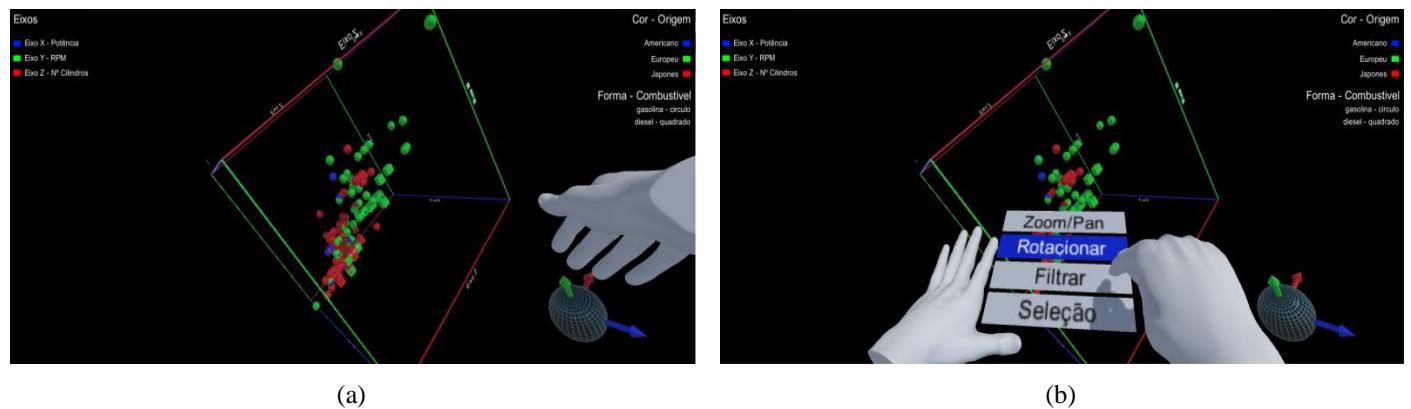


Fig. 3. Mecanismos de rotação e seleção de interação.

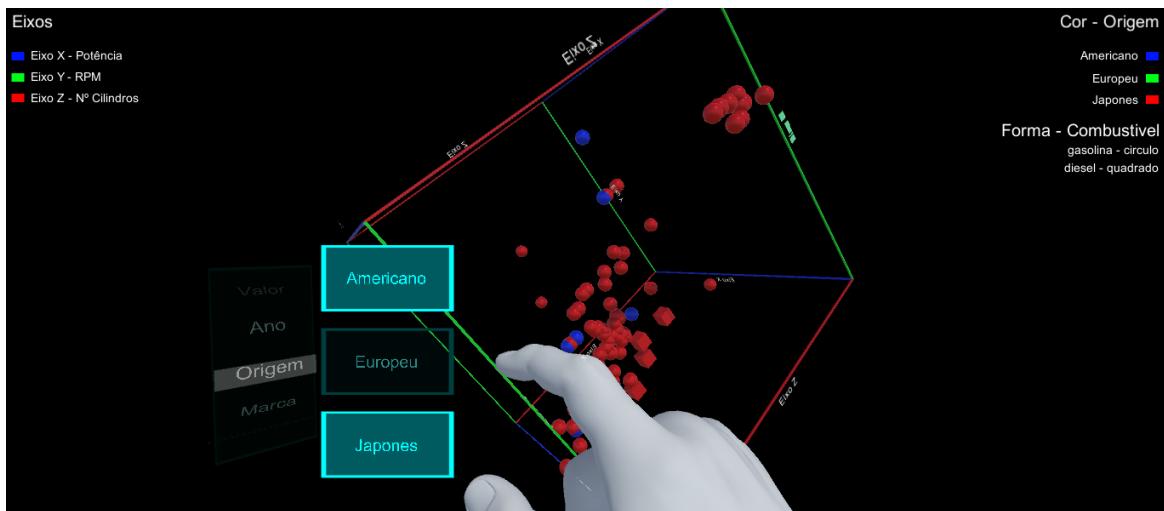


Fig. 4. Filtro categórico.

A última opção apresentada pelo menu inicial é a seleção. Neste o usuário é livre para interagir (tocando) com os itens da base de dados que estão na visualização. Ao tocar em um item da visualização este item é selecionado e é apresentado ao usuário três opções, que são: Selecionar itens de mesma cor, selecionar itens de mesma forma e visualizar os detalhes dos itens selecionados.

A Figura 5 mostra as opções dadas ao usuário quando o mesmo interage com um item na visualização. Para se chegar na visão da Figura 5 é necessário realizar um conjunto de operações de Zoom e Pan. Logo após as operações de Zoom e Pan o usuário deve acessar o menu inicial e escolher a última opção (Selecionar). Nesta opção o modelo tridimensional da mão é transladado mais para frente para que tenha mais contato com a visualização, permitindo que o alcance das mãos seja mais amplo.

Quando o usuário seleciona apenas um item na visualização e escolhe visualizar os seus detalhes é apresentado ao usuário todos os atributos deste item, já quando o usuário seleciona mais de um item na visualização e escolhe visualizar os detalhes desses itens é apresentado ao usuário algumas estatísticas sobre o conjunto de dados selecionado, como por exemplo média, desvio padrão, moda, mediana, para cada um dos atributos desses itens selecionados.

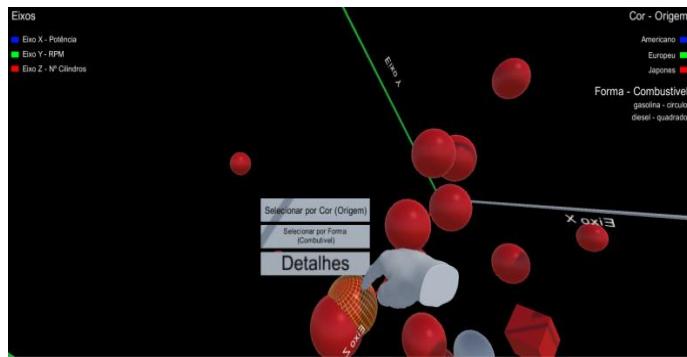


Fig. 5. Opções de seleção dos itens.

A ferramenta de seleção é importante para visualização da informação, pois permite que o usuário possa tanto ver detalhes de um determinado item, quanto relacionar um grupo de itens e obter informações deste grupo.

V. CONCLUSÃO E TRABALHOS FUTUROS

Este artigo apresentou uma ferramenta de InfoVis em ambiente totalmente 3D com interação através das mãos na forma de um protótipo inicial para avaliação futuras investigações de melhores técnicas para interação com as mãos em Visualização de Informação.

Durante a elaboração e desenvolvimento do protótipo foram identificados os seguintes pontos fortes: o ganho no espaço de desenho da visualização, controles que tem modelos mentais mais condizentes com um ambiente 3D do que *widgets* 2D, um protótipo para futuras investigações na área.

Foram identificados também alguns pontos fracos, que são: forte presença de efeito *jitter* na utilização do protótipo, a necessidade de utilizar gestos pré-definidos para algumas tarefas com possível detrimento da naturalidade, e ausência de configurações e manipulação da base de dados em si.

A visualização da informação é uma importante área de pesquisa e o estudo das interfaces de interação com as mesmas tem destaque por permitir avanços em relação a um melhor envolvimento, experiência de utilização e uma melhor performance de utilização.

Como trabalhos futuros serão realizadas melhorias neste protótipo aliado a testes de desenvolvimentos para prevenir eventuais bugs e desvios de funcionalidades. Após, as melhorias o protótipo será utilizado para a realização de inspeções com especialista de usabilidade e finalmente teste tanto de usabilidade quanto de performance com usuários.

REFERÊNCIAS

- [1] J. S. Yi, Y. Kang, J.T. Stasko, J.A. Jacko. "Toward a Deeper Understanding of the Role of Interaction in Information Visualization". IEEE Transactions on Visualization and Computer Graphics. vol. 13, pp. 1224-1231. Novembro de 2007.
- [2] W. Willet, J. Heer, M. Agrawala. "Scented Widgets: Improving Navigation Cues with Embedded Visualizations". IEEE Transactions on Visualization and Computer Graphics. vol. 13, pp. 1129-1136, Novembro de 2007.
- [3] M. O. Ward, G. Grinstein, D. Keim. "Interactive Data Visualization: Foundations, Techniques, and Applications". A K Peters/CRC Press, Junho de 2010.
- [4] C. Ware. "Information Visualization: Perception for Design" (Third Edition). Morgan Kaufmann, Junho de 2012.
- [5] S. Card, J. Mackinlay, B. Shneiderman. "Readings in Information Visualization: Using Vision to Think". Morgan Kaufmann, Fevereiro de 1999.
- [6] L. Vlaming, C. Collins, M. Hancock, M. Nacenta, T. Isenberg, S. Carpendale. "Integrating 2D mouse emulation with 3D manipulation for visualizations on a multi-touch table". ACM International Conference on Interactive Tabletops and Surfaces. pp. 221-230, ACM New York, 2010.
- [7] L. Jiang, A. Nandi. "SnapToQuery: providing interactive feedback during exploratory query specification". Proceedings of the VLDB Endowment Journal. vol. 8, pp. 1250-1261, Julho de 2015.
- [8] Leap Motion. Disponível em: <https://www.leapmotion.com/>. Último acesso: 03/10/2015. Science, 1989.
- [9] Unity 3D. Disponível em: <https://unity3d.com/pt>. Último acesso: 03/10/2015.

Muito Além do Papel

Uma Experiência de Design e Avaliação de Realidade Aumentada para Cartões Impressos

Tatiana Aires Tavares

Centro de Informática

Universidade Federal da Paraíba

João Pessoa, Brasil

tatiana@lavid.ufpb.br

Herson Barreto, Lívia Barreto, Júlio Carvalho

Centro de Informática

Universidade Federal da Paraíba

João Pessoa, Brasil

Leandro de Assis, Mayrton Dias, Rennan Temoteo

Centro de Informática

Universidade Federal da Paraíba

João Pessoa, Brasil

Abstract—In the past few years, developers around the world have started launching applications that use Augmented Reality (AR) and aim to make the virtual world an inherent part of our daily lives. This paper issues the interaction by using user interfaces based in Augmented Reality (AR). To do that, an app was used, which involved interactive printed cards and Aurasma platform. The cards were displayed on the wall and the users were asked to interact with them through their smartphones or tablets. This app was used during the Webmedia 2014 and appreciated by 48 users. The current paper also presents the design experience of utilizing an app for Augmented Reality (AR). Afterwards, the obtained results are discussed.

Keywords—Augmented Reality, User Experience, Interaction

Resumo—Nos últimos anos, os desenvolvedores de todo o mundo começaram a lançar aplicações que utilizam a Realidade Aumentada (RA) e visam tornar o mundo virtual parte integrante de nosso cotidiano. Este artigo busca discutir a interação usando interfaces de usuário em Realidade Aumentada (RA). Para tanto, uma camada virtual foi desenvolvida, que envolveu cartões impressos e a plataforma Aurasma. Os cartões foram exibidos na parede e os usuários foram convidados a interagir com eles através de seus smartphones ou tablets. Este experimento foi realizado durante a Webmedia 2014 e apreciado por 48 usuários. Por fim, são apresentados e discutidos os resultados obtidos.

Palavras-chave—Realidade Aumentada, Experiência de Usuário, Intereração.

I. INTRODUÇÃO

Realidade Aumentada (do inglês: *Augmented Reality*) é um conceito que está ganhando espaço nos últimos anos. Esta tecnologia de interatividade pode ser utilizada em ambientes diversos, desde aplicações voltadas a ambientes corporativos, até aplicações voltadas ao entretenimento. Para Kirner (2007) a Realidade Aumentada (RA) é definida pelo

enriquecimento do ambiente real com objetos virtuais, usando algum dispositivo tecnológico, funcionando em tempo real.



Fig. 1. Exemplo de RA no cotidiano de arquitetos e construtores

Quando esse conceito é aplicado à mídia impressa (papel) temos a *paper-based augmented reality* (ou Realidade Aumentada baseada em papel). Este conceito foi introduzido em 2007 por Jonathan Hull et al., para eles: a Realidade Aumentada baseada em papel usa marcações de texto para dados eletrônicos e através de uma câmera de telefone como dispositivo de reconhecimento exibe uma camada virtual para os usuários [9].

Este artigo apresenta uma experiência de design e implementação de uma camada virtual em RA. A camada desenvolvida combina real e virtual para proporcionar um retrospecto aos 20 anos de um evento científico chamado Webmedia 2014¹. Para tanto, o papel ganhou vida através de uma plataforma de RA chamada Aurasma [1]. Também são discutidos resultados de experiências reais de uso desse app realizadas durante o evento. A experiência incluiu também a concepção de cartões impressos que atuavam como pequenos pôsteres, através dos quais o usuário poderia acessar o modo interativo em RA. Os cartões foram exibidos em um estande do

¹ <http://webmedia2014.lavid.ufpb.br/>

evento e os participantes foram convidados a interagir com eles através da aplicação escolhida. Além disso, a atividade desenvolvida permitiu incluir aspectos de acessibilidade na interface voltados para deficientes visuais e auditivos. Dessa forma, alguns cartões apresentavam símbolos em Braille e traduções em LIBRAS geradas pelo aplicativo Handtalk [7].

II. RA E RA BASEADA EM PAPEL

Realidade Aumentada é a noção de que uma visão indireta de um mundo real tem os seus elementos aumentados via entradas geradas por computador, tais como, imagens, vídeos, ícones. Além disso, tecnologias de RA podem promover melhorias do mundo real com textos, imagens e objetos virtuais, gerados por computador [3]. Para Milgram (1994) a Realidade Aumentada é a mistura de mundos reais e virtuais em algum ponto da realidade/virtualidade contínua, que conecta ambientes completamente reais a ambientes completamente virtuais.

Em [2] são citadas as seguintes características como propriedades para sistemas em AR:

- Combinar objetos reais e virtuais no ambiente real;
- Executar interativamente em tempo real;
- Alinhar objetos reais e virtuais entre si;
- Requerer a aplicação de todos os sentidos, incluindo audição, tato e força e cheiro.

Um aspecto indispensável para o desenvolvimento de qualquer interface de usuário é a preocupação em torná-la o mais agradável e transparente possível. Esse aspecto não é diferente quando se trata do desenvolvimento de interfaces de usuário que se utilizam de Realidade Aumentada. O autor é ainda mais crítico, uma vez que a noção de real e virtual se misturam causando uma só impressão na recepção do usuário final. Carvalho et al. (2012) acrescentam que quando se considera a utilização de RA, é importante ponderar as seguintes questões:

- Utilizar dispositivos de interação que se adequem à tarefa;
- Facilitar uso e aprendizagem;
- Promover o desempenho do usuário e sua satisfação.

No caso específico da RA baseada em papel, é possível reutilizar essa lista de considerações acrescida da preocupação com o design do produto impresso, com a escolha da marcação (gatilho para a camada a virtual) e com o posicionamento adequado ao dispositivo de reconhecimento. Segundo Hull (2007) a RA baseada em papel traz um novo nível de interatividade a documentos em papel e permite que estes sejam atualizadas dinamicamente. Novos links podem ser facilmente criados e o conteúdo dos links pode ser alterado sem modificar o documento original impresso.

Um modelo para criação e uso de RA baseada em papel é apresentado na Fig.2 [9]. Este modelo foi utilizado como guia para delinear aspectos do desenvolvimento do estudo de caso,

detalhado na próxima seção. Este modelo sugere etapas para criação da camada virtual: indexação, reconhecimento e exibição. Em (a) a indexação: definir gatilhos associados à marcação da imagem. Em (b) o reconhecimento: algoritmo capaz de reconhecer a marcação do texto/imagem que identifica o gatilho, ou seja, local em que a câmera está apontando e onde será disparada a exibição da camada virtual. E em (c) o dispositivo de reconhecimento que será capaz de exibir a camada virtual. Como este modelo é fundamentado em links percebemos que a indexação assume o papel de mapear marcações no texto/imagem (x,y) em hiperlinks.

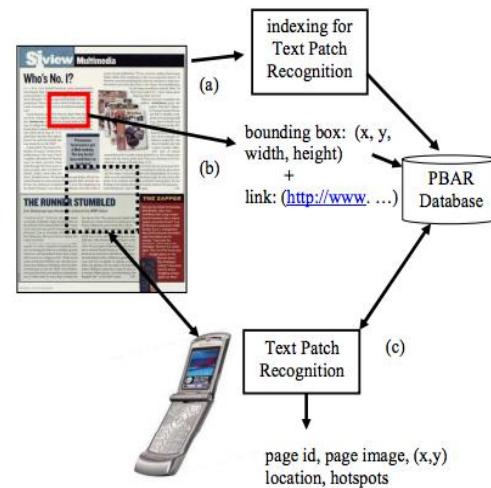


Fig. 2. Modelo de Criação de RA Baseada em Papel. Fonte: [9] pág. 205.

III. ESTUDO DE CASO: WEBMEDIA 20 ANOS

Como dito anteriormente, para a elaboração deste estudo tomou-se por base o modelo de criação de RA em papel apresentado na Fig.2. A Fig.3 apresenta a adaptação deste modelo de criação de RA Baseada em Papel para o presente estudo. Como se pode observar, foi mantida a divisão das etapas principais: indexação, reconhecimento e exibição. Para endereçar novas necessidades foram adicionadas atividades, tais como a elaboração de um modelo de navegação para o conteúdo impresso como descrito a seguir.

A etapa de **indexação** é marcada por duas atividades essenciais: a marcação e o conteúdo impresso. Dessa forma, os cartões impressos foram projetados para promover uma fácil interação e oferecer apelo visual para atrair a atenção dos usuários (ou visitantes). Cores vibrantes e imagens que caracterizavam de forma representativa a linha do tempo do evento foram aspectos observados neste projeto como pode ser visto na Fig.4. Além do aspecto visual, também foi adicionado um item para toque que é a legenda em Braille. Esta *feature* foi projetada para suportar a acessibilidade como será discutido mais tarde.



Fig. 3. Modelo de Navegação Temporal dos Cartões.

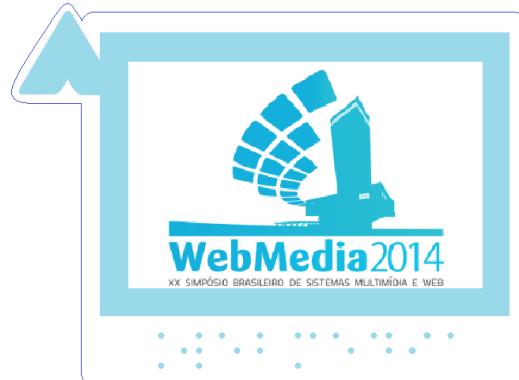


Fig. 4. Cartão Impresso.

Como o conteúdo impresso deste estudo de caso possui vários elementos visuais, foi adicionado um modelo de navegação análogo ao modelo de navegação de páginas Web. No entanto, dessa vez ao invés de páginas Web são descritos cartões impressos. O conteúdo foi distribuído pelos cartões em ordem cronológica, sendo utilizado o logotipo referente a cada ano de evento para identificá-lo (segundo o exemplo da Fig.4), conforme a ilustração da Fig.5.

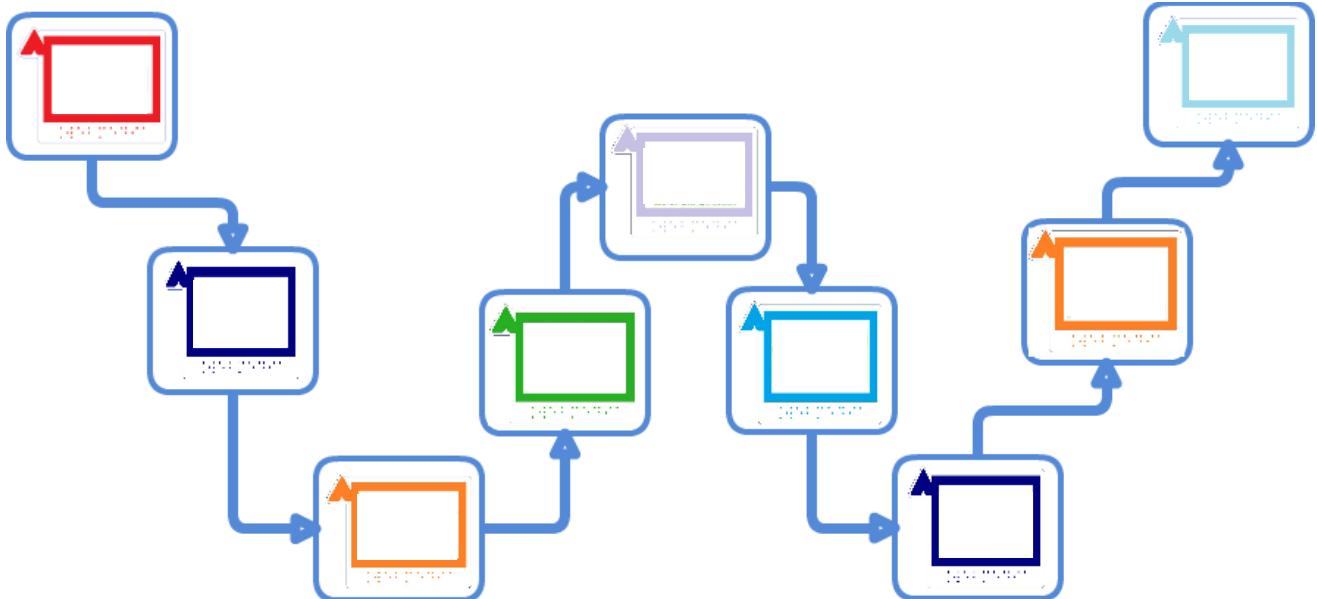


Fig. 5. Modelo de Navegação dos Cartões.

Para cada cartão, segundo a ordem cronológica (ano de realização dos eventos) do mesmo, o conteúdo digital acessível pela sua imagem gatilho correspondente foi projetado (aura). Este conteúdo foi definido e produzido de forma incremental, de forma a contar a história das edições passadas do evento através de vídeos e fotografias retiradas dos sites oficiais dos mesmos e de material fornecido por voluntários.

Uma vez definido o design dos cartões, partiu-se para o projeto das camadas virtuais que se sobrepõe aos cartões e proporcionam ao visitante a experiência de RA. Em outras palavras, trabalhar na identificação dos gatilhos e da exibição

das mídias que compõem a camada virtual. Para que este objetivo fosse alcançado utilizou-se a ferramenta Aurasma (AURASMA, 2015), disponível de forma gratuita para iOS e Android. A ferramenta Aurasma é uma plataforma de desenvolvimento mobile para RA.

Para a plataforma Aurasma as camadas virtuais são definidas como *auras*, que permitem a identificação de gatilhos. Através dessa plataforma é feito um mapeamento de qualquer imagem, vídeo ou objeto criando o reconhecimento. Desta maneira, o aplicativo mistura o real com o virtual e gera um conteúdo interativo. Ainda se podem criar conjuntos de

auras, ou canais, que podem ser compartilhados através de redes sociais como Facebook e Twitter.

Adicionalmente foi criado um canal Webmedia onde as auras projetadas para os cartões impressos foram disponibilizadas. Dessa forma, após a instalação e criação da conta de usuário, o usuário deve selecionar na tela inicial (vide Fig.6-a) o canal desejado. Neste caso, o canal Webmedia conforme Fig.6-b.

Após estar seguindo o canal, o usuário já está apto a navegar pelas camadas virtuais do Webmedia 2014. Sendo necessária uma conexão de internet (provista pelo evento). A Fig.7-a ilustra as informações do canal e possibilita a visualização de todas as auras. A Fig 7-b apresenta a visualização de uma camada virtual para um dos gatilhos de um dos cartões impressos. Ao total foram desenvolvidos onze cartões impressos e onze camadas virtuais que foram apreciadas por aproximadamente 50 visitantes durante a realização do evento. O experimento de avaliação é detalhado nas subseções seguintes.



Fig. 6. Telas do Aurasma. Em (a) tela de exploração e em (b) canal Webmedia.

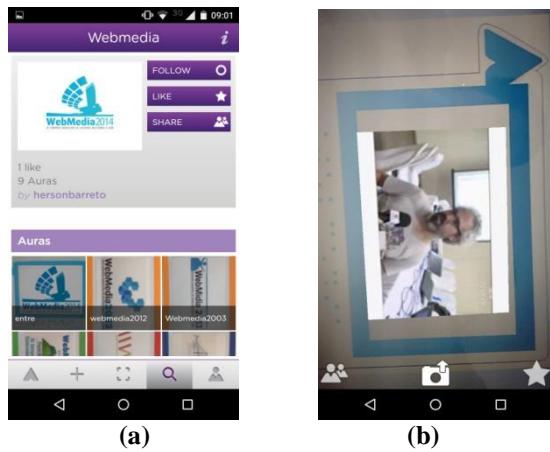


Fig. 7. Telas do Aurasma. Em (a) tela do canal e em (b) visualização da aura

IV. EXPERIMENTO

A importância da avaliação de interface está em conhecer um pouco mais sobre como o usuário interage com a aplicação, conhecer os pontos fortes e fragilidades das soluções de design

adotadas. Em produtos inovadores, o feedback do usuário é de extrema relevância, pois permite também que o produto seja conhecido. Nesse caso, desejou-se avaliar a receptividade dos usuários à Realidade Aumentada como interface de usuário. Então, optou-se pela avaliação através da observação da experiência real dos usuários com o aplicativo.

O objetivo principal deste experimento é avaliar a aceitação dos usuários a interação com RA. A metodologia do experimento foi bastante simples: foram elaborados cartões interativos para onze edições do evento, e para cada cartão um vídeo com fatos e fotos daquele evento foi produzido. E, então, usuários foram convidados a experimentar a aplicação e depois responder três perguntas rápidas, como detalhado nas subseções seguintes.

A. Cenário de Uso

As experiências com usuários foram realizada nos dias 18, 19, 20 e 21 de novembro de 2014 durante o Simpósio Brasileiro de Sistemas Multimídia e Web (ou Webmedia). O Webmedia é um congresso promovido anualmente pela Sociedade Brasileira de Computação que visa o intercâmbio científico e técnico entre alunos, pesquisadores, e profissionais das áreas de multimídia, hipermídia e Web. O congresso celebrou em 2014 a sua vigésima edição, ocorrendo na cidade de João Pessoa – PB e reuniu aproximadamente 200 participantes.

Durante o evento os cartões interativos foram distribuídos em um estande e a equipe de desenvolvimento do projeto ficou nesse espaço convidando os visitantes para interagir. A interação foi feita utilizando tablets oferecidos pela equipe ou através dos smartphones dos próprios voluntários que se disponibilizaram a instalar a Aurasma e acessar o canal do Webmedia.

B. Público-Alvo

O público-alvo foram os participantes do evento formado por palestrantes, estudantes e professores da área de Computação e Mídias Digitais. Como entusiastas das tecnologias atuais, a maioria dos participantes já possuía experiência com a utilização de smartphones e tablets, o que tornou a aplicação de uso simples, porém grande parte não conhecia o aplicativo Aurasma e o conceito da Realidade Aumentada, gerando grande expectativa. Em especial, a equipe teve consulta com um usuário surdo durante o experimento e um usuário cego durante a concepção dos cartões impressos.

C. Instrumento de Avaliação

Os visitantes do estande foram convidados a responder um questionário com três perguntas simples, com o objetivo de analisar a qualidade da interação. As perguntas são:

- Gostou da Realidade Aumentada?
- Utilizaria a tecnologia em outro contexto?
- Você já conhecia a Realidade Aumentada?

Além das perguntas os visitantes foram observados durante a experiência de uso e alguns registros e apontamentos foram feitos pelos avaliadores.

D. Resultados Obtidos

Os participantes foram convidados a interagir utilizando smartphones e tablets como pode ser visto na Fig.8 e na Fig.9, respectivamente. É importante salientar que o conteúdo (camada virtual) disponibilizado pelo smartphone ou pelo tablet é o mesmo. A intenção não foi avaliar o dispositivo e sim a interação. No entanto, é sabido que o dispositivo influencia na interação e isso foi verificado durante os testes realizados.



Fig. 8. Experiência de Usuário com Smartphone.



Fig. 9. Experiência de Usuário com Dispositivo Móvel.

Primeiramente, foi verificado se os visitantes tinham conhecimento prévio sobre Realidade Aumentada através da pergunta “Você sabia o que era Realidade Aumentada?”. Para esta pergunta 77% assinalou que já conhecia enquanto 23% não. Mesmo em um nicho fortemente tecnológico, quase ¼ dos entrevistados não tinha conhecimento de RA. A Fig.10 e Fig.11 mostram os resultados obtidos com a aplicação do instrumento de avaliação (questionário).



Fig. 10. Resultados da pergunta “Você sabia o que era Realidade Aumentada?”

Na Fig.11 pode-se observar que o resultado foi bem positivo em todas as questões e que quase 80% dos usuários entrevistados gostaram da experiência que tiveram.



Fig. 11. Sumário dos Resultados Obtidos.

Para a pergunta “Gostou da Realidade Aumentada?” as respostas dos visitantes demonstraram que a aceitabilidade foi excelente. Foi obtido o seguinte *score* em uma escala de 1-10 (onde 10 representa a satisfação máxima):

- 79% pontuaram com 10 (score máximo);
- 19% pontuaram com 9 e,
- 2% pontuaram com 8.

Acredita-se que esse resultado pode ter sido causado, pelo fato de se tratar de uma tecnologia recente, que causava certa surpresa nos visitantes. E se tratando de uma nova tecnologia, o público pôde reagir de forma positiva ou negativa. Do ponto de vista de interface de usuário, foi observado que o fato do

aplicativo ser reativo, de maneira que, o usuário não precise clicar em nenhum botão apenas posicionar o dispositivo, sem precisar de um manual de instruções, ele acaba se tornando um convite para a interação com o usuário. Os visitantes demonstraram através de suas reações um estado de entusiasmo e também controle. O fato de ser surpreendido pela tecnologia contribuiu para a autoconfiança na interação, pois rapidamente os visitantes queriam ir para o próximo cartão e muitos solicitaram informações de como instalar o aplicativo.

O resultado da próxima pergunta: “*Utilizaria a tecnologia em outro contexto?*” foi bem surpreendente, uma vez que 100% dos voluntários assinalaram SIM. Como os testes foram realizados em um ambiente acadêmico, e que a maioria do público-alvo conhece ou trabalha com tecnologia, esse resultado mostra a boa aplicabilidade dessa tecnologia. Esse score é bastante animador, pois denota a boa impressão que a experiência em Realidade Aumentada causou nos visitantes. Saúde, jogos e turismo foram contextos de uso citados pelos visitantes.

Em particular, foi notado que as *features* para acessibilidade foram bem aceitas pela comunidade em geral. Mas em especial, a observação da experiência de um visitante surdo foi bem estimulante, primeiro pela satisfação demonstrada por ter conteúdo em LIBRAS. E depois, pela fácil adaptação que foi verificada ao dispositivo e à interação.

Como a RA baseada em papel não necessita de qualquer informação prévia para disparo do gatilho, o visitante surdo apenas posicionou o dispositivo na imagem. O vídeo do conteúdo e o vídeo do intérprete foram apresentados para ele. É importante também destacar que a utilização do aplicativo Handtalk facilitou bastante a geração dos vídeos utilizados como mídia adicional para cada conteúdo. Durante o experimento um voluntário surdo fez uso do app e manifestou sua satisfação e até surpresa ao ter o conteúdo “preparado” para ele. Já a sinalização em Braille para usuários cegos apenas foi apreciada por um usuário cego fora do evento. O usuário que nos auxiliou na confecção da sinalização validou os cartões impressos em Braille. No entanto, como essa experiência não ocorreu durante o evento, isto é, com a mesma infraestrutura do evento, não foi utilizada. Questões de ergonomia e posicionamento não puderam ser discutidas.

V. DISCUSSÃO

Através dessa avaliação alcançada por questionários respondidos pelos usuários, foi observado que a maioria deles expressou satisfação e entusiasmo, por se tratar de uma abordagem nova, que deve ser mais explorada no futuro para ser aplicada também no cotidiano, para facilitar o compartilhamento de informações em ambientes diversos e ao mesmo tempo estimular as pessoas a buscarem saber do que se trata o conteúdo ali disposto a ser revelado.

Com o conceito de Realidade Aumentada, pôde-se promover uma informação interativa e funcional com a utilização do aplicativo Aurasma para os usuários. Como foi visto, poucos conheciam o conceito da tecnologia, mas todos (100%) voltariam a utilizar a mesma em outros contextos.

A avaliação dos resultados obtidos por observação da experiência realizada também revela resultados satisfatórios e uma boa receptividade do trabalho realizado.

Como comentado anteriormente foram inseridos em algumas camadas virtuais a legenda em LIBRAS como também a marcação em Braille nos cartões impressos. Embora este não fosse o objetivo deste estudo, a adição dessa *feature* na camada virtual, demonstrou o grande potencial da utilização desse tipo de ferramenta em contextos de acessibilidade. A versatilidade no acesso e na exibição do conteúdo das camadas virtuais são pontos chave para uma interação mais transparente. Por sua vez, a interação mais transparente que quase surpreende o usuário, é um passo em direção à acessibilidade. Uma vez carregado o canal, o usuário não necessita mais selecionar, digitar ou dizer alguma coisa para iniciar a interação que ocorrerá à medida que as marcações são reconhecidas pelo dispositivo. Esse fato é um forte indício do potencial da RA para promoção da acessibilidade.

Por fim, seja para divertir, informar ou incluir, percebemos através deste estudo que a RA é uma realidade cada vez mais comum no cotidiano. Além disso, a RA é uma forma generosa de interação que abre muitas portas e muitas oportunidades. A RA permeia diferentes estilos de interação e permite trabalhar com elementos que variam do papel à projeção holográfica. Essa abrangência de elementos é muito bem vinda para desmistificar a tecnologia e popularizar seu uso em diferenciados contextos.

Este estudo é apenas um ponto inicial na discussão da Realidade Aumentada no dia-a-dia, tendo como intenção incorporar novos contextos, novas ferramentas e novos instrumentos de avaliação em trabalhos futuros.

REFERÊNCIAS

- [1] AURASMA (2015) – <http://www.aurasma.com/explore/>
- [2] Azuma, R. et al. (2001) “Recent Advances in Augmented Reality.” IEEE Computer Graphics and Applications, v.21, n.6, p. 34-47. 19
- [3] Insley, S. (2003) ”Obstacles to General Purpose Augmented Reality <http://islab.oregonstate.edu/koc/ece399/f03/final/insley2.pdf>
- [4] KIRNER, Claudio; SISCOUTTO, Robson. Realidade virtual e aumentada: conceitos, projeto e aplicações. In: Livro do IX Symposium on Virtual and Augmented Reality, Petrópolis (RJ), Porto Alegre: SBC. 2007.
- [5] Milgram, P. et. al. (1994) “Augmented Reality: A Class of Displays on the Reality-Virtuality Continuum”. Telemanipulator and Telepresence Technologies, SPIE, V.2351, p. 282-292.
- [6] de CARVALHO, J. O. F. (2012). O papel da interação humano-computador na inclusão digital. *Transinformação*, 15(3). Acessível em: <http://www.scielo.br/pdf/tin/v15nspe/04.pdf>
- [7] HandTalk (2015) - www.handtalk.me/
- [8] de Castro, Rafael Henrique Assis, Anna Carolina Soares Medeiros, and Tatiana Aires Tavares. "Natural Interaction for 3D Manipulation in Telemedicine: A Study Case developed for Arthon Video Collaboration Tool." In *Workshop de Realidade Virtual e Aumentada-WRVA 2012*. 2012.
- [9] Hull, Jonathan J., Berna Erol, Jamey Graham, Qifa Ke, Hidenobu Kishi, Jorge Moraleda, and Daniel G. Van Olst. "Paper-based augmented reality." In *Artificial Reality and Telexistence, 17th International Conference on*, pp. 205-209. IEEE, 2007.

Risks MultiProjection Use in Virtual Reality as Social Technology

Luciene Bulhões Mattos

DG – Campus Salvador

Instituto Federal da Bahia – IFBA

Salvador-Bahia

lueta32@hotmail.com, lueta@ifba.edu.br

Abstract—In Brazil and in the world are being developed to focus research on simulations with intelligent virtual reality systems from the areas of health to the arts. The possibility of realization of what is thought or felt, once considered an abstraction or imagination of the event, enables a collective sense of transcendence. The Virtual Reality - RV, Augmented Reality - RA and systems structured by neural networks such as multiprojection systems (or distributed) consolidate this supposed union between abstraction and imagination. However, the applications of these systems as social technologies can generate expectations and conflicts or misleading ways of perceiving the change. This article attempts to build a dialogue about it watching the constructs of these systems when used as applicable sciences to and for certain social groups to and from certain social interest groups.

Keywords—*Simulation; RVI systems; imagination; systems multiprojection; social technologies.*

The simulation, however, corresponding to the stages of the previous intellectual activity to rational exposition: imagination, mental bricolage, attempts and errors.

Lèvy, 1993

The resources of virtual and augmented reality makes it possible to generate simulations. Observing a technical definition of simulation, "The simulation is used to make virtual objects have appropriate behaviors such as movement, collision, reaction, physical simulation, etc. Behaviors need not imitate reality, but should give the synthetic elements, specific properties (Kirner; Tori , 2006, p. 25). ", you can see the possibilities of approximation to reality mediane cert. To

achieve an acceptable degree of realism, the graphic realism (the forms and images) were measured in frames per second that determined the performance. The level of performance determined the usability of such resources.



Figure 1. Cave -
https://upload.wikimedia.org/wikipedia/commons/6/6d/CAVE_Crayoland.jpg

The multiprojection systems in RV are staged through CAVEs (Figure 1) and its applications propose to the desired quality of performance. They require a graphic multiprocessor. The tightly coupled systems are expensive and slightly coupled systems require "[...] two or more computer systems connected through a communication line. Each system has its own processor, memory, devices input / output and operating system and can function independently (Tanenbaum in Guimarães et

al, 2003). These features allow systems tightly coupled to the management and processing applied to desenvolvimento dos multiprojeção systems.

II. POINT OF VIEW AND MULTIPROJECTION

A. Multiprojection systems depend on us

Parallel images are generated using some techniques. They are: a) Sort-first, the screen is divided into different screens, processors are responsible for rendering calculations, which affect the region pixel; b) Sort-middle, arbitrary subsets of primitives are associated with a geometric processor and c) Last-Sort, each processor is associated with a set of primitives and solve the problem of visibility of the pixels, calculating pixel values and sending processors composition.

The technological system multiprojection enabled the creation of a language adapted to everyday life. We are systems (machines) managed by people (intelligence). The techniques used for imaging are well leveraged, imagined and adapted to coercive strategies for different purposes, take the lollipop from the mouth of a child to generate a confusion between aesthetic subject by manipulating images of their physical or audible (voice). A true transmutation.

Thus, it is understood as screen display collective processors like people in the process of rendering resolution / decision, pixel or each change, modeling adapted to the daily life of a person or a group. The objects are ordinary people or victims of speculation among us. Here is the instrumentalization of human beings through Social Technology and relativism in multiprojection.

In countries in which technological development occurs in parallel to human development initiatives, guided by investments in education, this transmutation occurs less abusive. There is a separation of wheat and tares. In developing countries, which are marked by strong conflicts between class, notoriously, no indiscriminate use of these technologies in the social environment with very serious consequences. One is the weakening of the communication rather than the strengthening of information systems, which renew neocapitalist sociotechnical assemblages.

The man-machine hybridization dreamed, body-technology is already weakening man's need while developing human. Man and humans are not the same thing? With the advent of Man-Machine, Body-Technology in VR hybridizations as a Social Technology there seems to be a new species such as the movies, the EX MAM, which should be interpreted as "Ex man or one who was once a human being." Would not it be better to tell technologies for technology education?

III . THE LAST STRAW TO NEED OF THE CAGE (...) CAVE

About the ethical and philosophical problems, a picture attributed to the painter Pablo Picasso called "Las Demoiselles d'Avignon" (Figure 2) should be can help as example. Certain groups enjoy casual sex and extramarital relationship and like to attend certain environments specific to the subject. Women are seen in the picture with her naked body in what seems a tent,

studio and workshop. What is done there will not be important to disclose here.

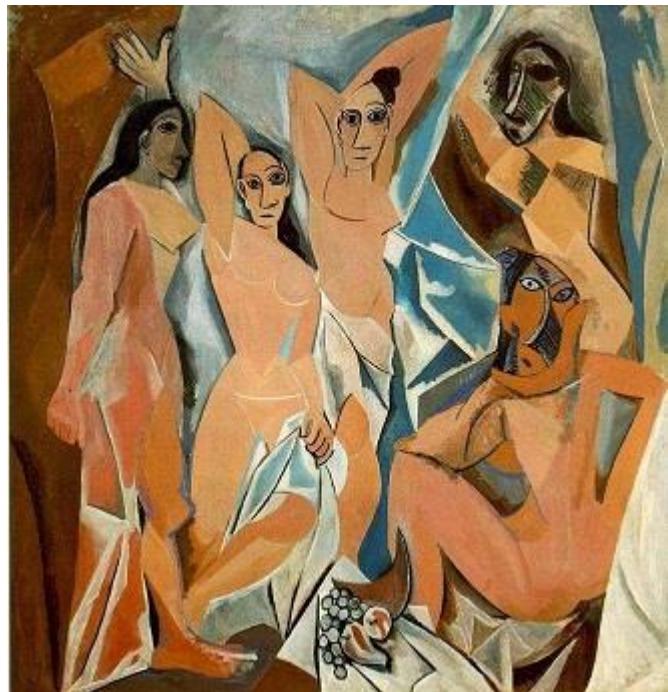


Figure 2. Les Demoiselles D'Avignon, Pablo Picasso - <http://www.google.com>

In case of an overlap between virtual and real environments culminating in Virtual Reality Intelligent - RVI or in Virtual Augmented - RA, as would occur a mixture in a concrete real situation and Intelligent Virtual Reality systems? Between the experience in a CAVE and a scene as in "Las Demoiselles d'Avignon".

This text leaves it to the reader to examine. Maybe the picture becomes a cage, the cage for commercial prostitution mass due to lack of regulation on the use of living laboratories attached to RVI systems. Seeing the artwork above and think of possible interpretations ranging between metonymy and metaphor, there is no doubt that the systems in RV, RA and RVI are inadvisable outside the academic environment.

Like, any connection that includes such systems in mobility systems (cell) or software interaction (userZoom), since they streamline the information in mass. Open source technologies have enabled the copy of this software and its full commercial distribution parallel without even paying attention to public health problems. And even those who have no use systems like Moodle, which allows interaction cataloging some thought into action.

The irresponsibility of the meaning given to the CAVE and Living Labs transforms the fruit of scientific and academic research in a mere projection area of sensations and pleasures. What was to encourage the full development in human evolution, turns into its opposite. A constant anarchy in the XXI

century. Here is the stronghold of pseudoscience. As a second example, imagine a residential condominium with 208 units or apartments.

Each apartment is that one or more "Las Demoiselles d'Avignon" or a family residence. Suddenly, people start to behave differently. Finds that there is a CAVE somewhere. It monitors and represents the characters of "Las Demoiselles d'Avignon" with the condo residents. Many do not know what's going on. The CAVE has become the "cage-node" network systems that main body multiprojeções and "Las Demoiselles d'Avignon" a scene commonplace everyday. Explodes then a set of contradictions. Who is actor or character? Where are moral elements? Behold the exit Laboratory Engineering, Medicine, Arts Workshop, etc., Systems in RV, RA and RVI tend to corrupt themselves.

The use of Assistive Technologies Invasive - TAI at Systems Reality, Augmented and Mixed as a form of social control and domination. Would be good if this new form of innovation provoke another focused on the development and growth with citizenship. But it is worrying. These are assistive technologies Invasive - TAI that, being structured by neural networks, stimulate seizure tendency of the subject to be taken to the summit of the devaluation of factors relevant to the human condition such as welfare, good living, health, education, etc., and maximize human passions.

Simulation in Virtual Reality Systems sums to assistive technologies Invasive - TAI when exposed to excessively folksy even manage to sort new ways to meet or know. In this sense, they promote a disorder almost silent as an examination of bees, a set of postmodern images that result from an action model meaningless or multisentido. Are the effects of a guided multiprojection a collective sensibility. Is dominated by technocrats associated with the logic of neoliberal capitalism, Systems RV, RA and TAI consolidate the hegemonic model of domination.

In both there is a set, an apparatus for fishing mental, expanding the class struggle and reducing social development, creating socio-technical structures that are not even their own authors or perhaps new content. Behold, this is creating an illusion multiplicidade Hyper Surreal and grotesque nicely packaged by the label postmodernist. In this context, systems RV, RA and TAI as applied to social technologies are becoming a joke dangerous public that entered the twenty-first century learning society.

REFERENCES

ALEXANDRE CARDOSO AND CESAR AUGUSTO CAMILO AND EDGARD LAMOUNIER, PROGRAMAÇÃO DISTRIBUÍDA APLICADA À

REALIDADE VIRTUAL, IN: AMBIENTES VIRTUAIS: PROJETO E IMPLEMENTAÇÃO, PORTO ALEGRE: SBC, 2003.

Claudio Kirner, Evolução da Realidade Virtual no Brasil, In: X Symposium on Virtual and Augmented Reality, João Pessoa, 2008.

FRITJOF CAPRA, AS CONEXÕES OCULTAS: CIÊNCIA PARA UMA VIDA SUSTENTÁVEL, SÃO PAULO: CULTRIX, 2005.

Leandro de Lima Castro, Computação natural: uma jornada ilustrada, São Paulo: Editora Livraria da Física, 2010.

Luciene Bulhões Mattos, A relação homem-máquina, corpo-tecnologia em realidade virtual, In: Anais da Semana de Mobilização Científica - SEMOC, 2003, Salvador.

_____. Cognitive Ergonomics and Production Information Systems under the gaze of interdisciplinary Design or Industrial Design, In: Anais do Congresso, Recife: UFPE, 2012.

Disponível em:
http://iospress.metapress.com/content/d83wl46876028652/full_text.pdf; Acesso em: 05.2012.

_____. A relação homem-máquina, corpo-tecnologia em realidade virtual, In: Anais da VI SEMOC, Salvador: UCSAL, 2003.

_____. Transmutações do objeto em realidade virtual, In: Anais do 5º Encontro Regional de Expressão Gráfica - Educação Gráfica - perspectiva histórica e evolução, 2006, Salvador.

Pierre Lévy, O que é o Virtual?, São Paulo: Ed. 34, 1996.

_____. As tecnologias da inteligência, São Paulo: Ed. 34, 1993, Disponível em: <http://portugues.free-ebooks.net/ebook/As-Tecnologias-da-Inteligencia/html/76/>; Acesso em: 07/2013.

Renato Peterman and Wyllian Fressati, Um estudo sobre a aplicação de interface cérebro-computador na Computação, Unipar, PR. Disponível em: <http://www.scribd.com/doc/92707795/Uma-Abordagem-sobre-Interface-Cerebro-Computador-e-suas-Aplicacoes-na-Computacao>; Acesso em: jul/2013.

Robson Moreira Tenório, Cérebros e Computadores: a complementaridade analógico-digital na informática e na educação, 4. ed. São Paulo: Escrituras, p. 139-197.

Romero Tori and Claudio Kirner and Robson Siscoutto, Fundamentos e Tecnologia de Realidade Virtual e Aumentada, In: Book of Pre-Simposium, VIII Symposium on Virtual Reality, Belém- PA, 2006.

Standford University, Design Thinking Action Lab Course, Disponível em: <https://venture-lab.org/>; Acesso em: 07.2013.

Ferramentas de Autoria de Realidade Aumentada para Educação: revisão sistemática

Augmented Reality Authoring Tools for Education: Sistematic Revision

André Barone Rodrigues

Ciência da Computação

Faculdade Campo Limpo Paulista (FACCAMP)
Campo Limpo Paulista, Brasil
andre.barone@gmail.com

Valéria Farinazzo Martins

Faculdade de Computação e Informática
Universidade Presbiteriana Mackenzie
São Paulo, Brasil
valfarinazzo@hotmail.com

Diego Roberto Colombo Dias

Departamento de Computação, LaVIIC
Universidade Federal de São Carlos
São Carlos, Brasil
diegocolombo.dias@gmail.com

Marcelo de Paiva Guimarães

Universidade Federal de São Paulo (UAB/UNIFESP)/
Programa de Pós-Graduação da Faculdade Campo Limpo
Paulista (FACCAMP), São Paulo, Brasil
marcelodepaiva@gmail.com

Resumo— A realidade aumentada é uma tecnologia capaz de auxiliar a transmissão de conhecimento, mas para tanto é preciso facilitar o desenvolvimento das aplicações. O objetivo desse artigo é identificar na literatura científica trabalhos sobre ferramentas de autoria de realidade aumentada voltadas para educação que possibilitem a geração de objetos de aprendizagem. Para isso, foi realizada uma revisão sistemática de artigos publicados entre 2010 e 2015 nas bases da ACM e do IEEE. Essa pesquisa resultou em 4 artigos para análise, conforme os critérios estabelecidos para inclusão. Os artigos mostraram que os pesquisadores buscam desenvolver ferramentas de fácil uso que permitam a autoria de conteúdo em realidade aumentada. Contudo, ainda existem barreiras a serem ultrapassadas para que os educadores utilizem a realidade aumentada de maneira efetiva no processo de ensino-aprendizagem.

Palavras-chaves— *realidade aumentada, educação, ferramentas de autoria*

Abstract— Augmented reality is a technology capable to help on knowledge transmission, but for that is necessary to ease the applications development. The objective of this article is to identify scientific literature works related to authoring tools for augmented reality with educational purposes that generates learning objects. For this, a systematic revision has been conducted among articles published between 2010 and 2015 on ACM and IEEE databases. This research resulted on 4 articles for analysis, according to the criteria for inclusion. These articles showed that researchers are looking for the development of user-friendly tools that enable the content authoring on augmented reality. However, there are some barriers that need to be overcome for teachers to use augmented reality effectively in the teaching-learning process.

Keywords— *augmented reality, education, authoring tools*

I. INTRODUÇÃO

A Realidade Aumentada (RA) é uma tecnologia que permite observar o ambiente real enriquecido com objetos virtuais gerados digitalmente e apresentados por meio da tela de um computador ou dispositivo móvel [1]. Assim, permite formas inovadoras de interação com o meio real, o que cria diversas oportunidades, dentre elas, a de auxiliar a transmissão de conhecimento, independentemente do contexto educacional – por exemplo, ensino fundamental ou superior. Para que os ambientes de aprendizagem adotem de maneira efetiva a RA no processo de ensino-aprendizagem é necessário que as aplicações de RA sejam geradas por usuários não especialistas, ou seja, que não possuem conhecimentos aprofundados em computação, como, por exemplo, de programação e de detecção de padrões. Além disso, o processo de geração de criação delas deve demandar um tempo razoável, pois é desejável que o educador gaste um tempo mínimo com questões tecnológicas. Além disso, o custo deve ser acessível.

Segundo Martins & Guimarães [2], nos últimos anos surgiram diversas soluções que têm facilitado o desenvolvimento das aplicações de RA. Porém, ainda exigem conhecimento técnico e/ou requerem tempo considerável para a geração de tais soluções. Segundo eles, o uso das aplicações de RA no contexto educacional exige que as seguintes barreiras sejam superadas: (1) geração de conteúdo - o processo de criação ou aquisição de conteúdo (3D, imagens, sons) deve ser facilitado; (2) processo de desenvolvimento - os educadores devem ser capazes de criar as aplicações sem conhecimentos de programação; (3) montagem de laboratórios - as aplicações devem ser de fácil uso, sem exigir requisitos especiais e com viabilidade financeira. Esses pontos ainda são barreiras porque

existe uma disparidade entre o estado da arte da RA e o tempo necessário de amadurecimento para ser implantada de maneira efetiva no contexto educacional, ou seja, para ser utilizada de maneira fácil e com custo acessível.

Este artigo tem como objetivo realizar um mapeamento das soluções atuais que permitem o desenvolvimento de aplicações de RA de maneira fácil, que é feito por meio de ferramentas de autoria. Então, tais ferramentas devem ser voltadas para usuários finais, cujo foco de uso é educacional. Para isso, realizou-se uma revisão sistemática. Com o objetivo de ampliar as possibilidades de reuso das aplicações de RA pelos educadores, buscou-se na revisão identificar também as ferramentas de RA capazes de gerar aplicações no formato de objetos de aprendizagem [3].

O restante do artigo está organizado da seguinte forma: a seção II apresenta o método aplicado para realizar a revisão sistemática; a seção III mostra os artigos resultantes da busca; a seção IV realiza a discussão sobre as ferramentas encontradas; e, por fim, a seção V, as conclusões.

II. MÉTODO

Estudo da revisão de literatura científica em artigos publicados no período de 2010 a 2015 e coletados nas seguintes bases de dados: ACM (<http://dl.acm.org>) e IEEE (<http://ieeexplore.ieee.org/Xplore>). Em ambas as bases utilizou-se a combinação dos termos *Authoring* e *Augmented Reality* para o *abstract* e *Learning Object* ou *Education* para as *Keywords*. Na base do IEEE foi usada a busca por comando, enquanto na ACM os termos foram aplicados na ferramenta de busca avançada.

Após executadas as buscas, foi feita a análise e leitura dos resumos dos artigos encontrados, conforme os seguintes critérios de inclusão: artigos publicados entre 2010 e 2015; publicados em inglês ou português. Como critério de exclusão foi adotado: artigos que não tratavam exclusivamente de ferramentas de autoria.

III. RESULTADOS

O total de artigos encontrados foi de 43, dos quais foram selecionados quatro para leitura e análise completas, conforme os critérios de inclusão e exclusão. Todos os artigos escolhidos são da base do IEEE e nenhum artigo em português foi encontrado. A Tabela 1 apresenta a descrição dos artigos selecionados para leitura e análise, que têm seus métodos e resultados relatados a seguir.

Jee, Lim, Youn & Lee [4] mostraram uma ferramenta de autoria denominada *An Immersive Authoring Tool for Augmented Reality-based E-Learning Applications* para aplicações E-Learning baseadas em RA. Nesse estudo é citado que o modelo de ensino por meio da Internet evoluiu de modelos sem interação para os que usam lista de atividades e perguntas e respostas como forma de interação e de captura da atenção do aluno. Entretanto, eles consideram essas atividades maçantes e pouco intuitivas e, assim, acreditam que a RA possa ser mais eficiente nesses aspectos. O estudo também identificou a existência de bibliotecas que permitem a construção de aplicações para autoria de RA, como ARToolKit

[14], que usa recursos de visão computacional para identificação de marcadores para sobreposição de modelos 3D. Outra biblioteca citada é a Studierstube [5], que é de mais alto nível do que o ARToolKit, pois permite o gerenciamento de janelas, renderização de telas e suporte a dispositivos de entrada. Também foram citadas ferramentas que não exigem conhecimentos de linguagens programação, como o *plugin* para Macromedia Director Designers AR Toolkit (DART) [6], e a AMIRE [7], que permitem a criação e modificação de aplicações de RA. Não foi apontada nessas ferramentas a capacidade de gerar aplicações de RA para dispositivos móveis. Esses pesquisadores optaram por construir uma ferramenta sobre um software para edição de 3D chamado Maya [11], assim foi possível desenvolver um ambiente integrado de construção e modelagem 3D para autoria de RA. Para avaliar a eficiência do sistema os autores conduziram um experimento comparativo com alunos do ensino primário em aulas de inglês e ciências, no qual metade deles teve aula com material audiovisual tradicional e outra metade foi submetida ao sistema imersivo de E-Learning. Professores e alunos relataram que o modelo imersivo trouxe mais benefícios, pois os alunos se concentraram e participaram mais ativamente das aulas.

TABELA 1. DESCRIÇÃO DOS ARTIGOS SELECIONADOS NA BASE DE DADOS IEEE E ACM, SEUS OBJETIVOS E RESULTADOS

Artigos	Ferramenta	Objetivo	Resultados
Jee, Lim, Youn & Lee, 2011 [4]	An Immersive Authoring Tool for Augmented Reality-based E-Learning Applications	Projetar uma ferramenta de autoria que requer pouco esforço para autoria de conteúdo em RA.	A ferramenta ajuda os desenvolvedores a projetar aplicações, permitindo que eles fiquem focados apenas nos aspectos formais dessas aplicações, como marcadores e os modelos geométricos associados a estes marcadores.
Farias, Dantas & Burlamaqui, 2011 [9]	Educ-AR	Apresentar o software Educ-AR, usado para criar apresentações em RA de maneira fácil, sem a ajuda de um programador.	Desenvolveram uma API (<i>Application Programming Interface</i>), cujo componente principal é a biblioteca JARToolKit3D, que permite a criação de camadas em RA sobre o Educ-AR. Apontaram um caso de estudo onde alunos indicaram que o uso de RA torna a aula mais atrativa e interessante e professores indicaram ser viável seu uso.
Moralejo, Sanz, Pesado & Baldassari, 2013 [10]	AuthorAR	Apresentação da ferramenta AuthorAR, uma ferramenta de autoria para desenvolvimento de atividades educacionais.	É um trabalho em andamento, que produziu um protótipo.

Cubillo, Martín, Castro, Díaz, Colmenar & Boticki et al., 2014 [13]	ARLE	<p>Prover uma ferramenta de autoria, cujo uso seja fácil pelos professores, suplantando assim as duas razões pelas quais os autores acreditam que a RA não seja uma tecnologia de uso recorrente no ensino: a dificuldade em desenvolver experiências de aprendizado em RA e a falta de ferramentas desenvolvidas especificamente para educação.</p>	<p>A ferramenta mostrou-se eficiente na transmissão do conhecimento e houve retorno positivo por parte dos autores de conteúdo (professores), baseado pelas médias das notas entre o grupo do experimento que fez uso da ferramenta (7,59) e do grupo de controle (5,82), que não usou.</p>
---	------	--	---

Farias, Dantas & Burlamaqui [9] publicaram um estudo onde é apresentada a criação da API (Application Programming Interface) Educ-AR, que é baseada em ARToolKit. O principal aspecto apontado pelos autores é o fato dela não exigir que o instrutor tenha conhecimentos de programação, aplicações de RA ou modelagem 3D. No trabalho apresentado, foi criada uma ferramenta web, que os educadores se registram, criam classes e submetem marcadores e objetos 3D. Os artefatos em RA são gerados dentro de uma classe pela associação de objetos 3D com marcadores. A aplicação permite também o gerenciamento e criação de marcadores personalizados.

Eles realizaram um experimento com 20 pessoas, divididas em 2 grupos de 10, sendo um em ambiente Linux e outro em Windows, que executaram tarefas que começaram com o registro no portal, passando a inserção e associação de marcadores e modelos 3D, execução do Educ-AR e observação dos modelos quando os marcadores eram manipulados. 60% destas pessoas precisaram de ajuda na associação entre modelos e marcadores e 20 % não conseguiram executar o Educ-AR por ausência do Java na máquina. Os pesquisadores concluíram que o uso da tecnologia de RA não pode ser considerado uma solução eficiente por si só, são necessários outros apoios também, como o pedagógico. Os autores apresentaram um estudo de caso, executado no SENAI, em que os alunos aprovaram o uso da RA como ferramenta de apoio à aprendizagem, expondo que o uso dessa tecnologia torna a aula mais interessante e atrativa.

Outro estudo, realizado por Moralejo, Sanz, Pesado & Baldassari [10], apresentou a RA como uma tecnologia de assistência para alunos com necessidades especiais, particularmente com dificuldades de comunicação. Esse trabalho identificou a possibilidade do uso de um método híbrido, ou seja, que envolva não só o reconhecimento por visão computacional, mas também o uso de GPS (*Global Positioning System*) para adição dos objetos virtuais à exibição do conteúdo real. Os pesquisadores apresentaram um levantamento de algumas aplicações de RA para pessoas com necessidades especiais, sendo eles: o PictogramRoom [8], que ajuda na comunicação de pessoas com autismo; EyeRing [12],

que é um dispositivo tipo anel com conexão Bluetooth e câmera, para o auxílio de pessoas com deficiência visual, respondendo às imagens captadas com comandos de áudio. Percebendo que poucos projetos permitem a criação de conteúdo de RA voltado para educação, os autores apresentam a ferramenta AuthorAR com objetivo de auxiliar professores na criação de material em RA. O funcionamento do AuthorAR, na versão apresentada no estudo, é baseado em dois paradigmas de atividade: exploratórias, quando a atividade apresenta algo novo, como um conceito, fórmula, regra etc.; estruturação de frases, através de *templates* que permitem ao professor criar atividades para desenvolver nos alunos a composição de frases na forma sujeito-verbo-objeto. Ambos os tipos de atividades permitem ao professor adicionar alguma forma de capturar o retorno do aprendizado por parte dos alunos, como questionários ou um feedback para frases incorretas. O trabalho apresentado é parte de um projeto em desenvolvimento e ainda está sendo avaliado.

Um estudo, realizado por Cubillo, Martín, Castro, Díaz, Colmenar & Boticki [13], cita que novas tecnologias como Web 2.0, dispositivos móveis, ambientes virtuais e RA têm potencial para transmitir informação e adquirir conhecimento. Segundo os autores, a RA se mostra oportuna para o ensino de temas que envolvam treino prático e fora da sala de aula, onde os alunos necessitam de interação com o mundo real. No entanto, muitas aplicações educacionais de RA são específicas para um assunto, com conteúdos fixos, disponibilizados pelos desenvolvedores. Por outro lado, há ferramentas de RA voltadas para o uso geral, ou seja, sem direcionamento para o ensino. Assim, eles criaram um ambiente de aprendizado chamado de ARLE (*Augmented Reality Learning Environment*) que é capaz de integrar RA com o conteúdo teórico, sem que o autor tenha habilidades em programação. O sistema é baseado na arquitetura cliente-servidor, onde o conteúdo é enviado ao dispositivo móvel assim que a aplicação cliente é lançada. Para autoria de conteúdo, foi construída a ferramenta web ARLE Web-based Authoring Tool, que permite ao autor designar o formato do recurso gerado, seja imagem, vídeo, áudio ou objetos em formato 3D. A ferramenta possibilita também a inclusão de uma descrição do conteúdo e a criação de um questionário no formato de múltiplas escolhas. Por fim, o autor deve selecionar uma imagem que a câmera do dispositivo móvel reconhecerá como padrão para exibição do recurso. Quando o conteúdo é gerado, ele fica armazenado na biblioteca ARLE, permitindo seu reuso por outros autores. O ARLE possui também uma aplicação de visualização para dispositivos Android, através da qual o estudante faz uso do conteúdo, apontado a câmera do dispositivo para notas ou textos que foram usadas como referência na criação.

A avaliação do sistema foi feita por professores e estudantes, divididos em 2 grupos, sendo que um usou o sistema ARLE, chamado de grupo experimental, que poderia acessar o conteúdo tanto em sala de aula como em casa. O outro grupo, que não fez uso da tecnologia de RA, foi chamado de grupo de controle, podendo apenas acessar o vídeo em sala de aula. Foi então conduzido um exame de avaliação, com nota máxima de 8 pontos, para ambos os grupos, e os que fizeram uso do ARLE obtiveram uma média de 7,59 enquanto o grupo de controle obteve uma média de 5,82, mostrando assim um

sensível aumento no aprendizado do assunto. Os professores também deram um retorno positivo, considerando a velocidade do sistema e a facilidade de adicionar novos recursos virtuais e descrições.

IV. DISCUSSÃO

Conforme a análise dos trabalhos encontrados, a RA é objeto de estudo de alguns pesquisadores com o intuito de aliar ao processo de aprendizagem, por meio de ferramentas eletrônicas, como sistemas de E-Learning. O objetivo dos trabalhos analisados neste artigo é alcançar ferramentas que tornem o processo de autoria de RA confortável para qualquer pessoa, sem que haja a necessidade do envolvimento de especialistas em RA. Os trabalhos encontrados apontaram desde ferramentas em alto-nível para geração do conteúdo, inclusive com associação de atividades didáticas, como APIs que permitem o desenvolvimento de aplicações em RA de maneira mais fácil.

Um ponto importante levantado nas pesquisas é quanto à flexibilidade das ferramentas. Existem aplicações voltadas para o uso de RA na educação, porém, estes programas ficam restritos a uma disciplina e, com a flexibilização das ferramentas, torna-se possível criar conteúdo aplicável a diversas áreas do ensino e, em especial, disciplinas que requerem interação com o meio real. É apontado ainda que o uso da RA pode auxiliar no aprendizado de pessoas com necessidades especiais, apontando então uma característica importante à inclusão social.

Outro ponto de destaque nessa revisão é que apesar das ferramentas encontradas serem, geralmente, capazes de gerar aplicações de RA educacionais para qualquer contexto, elas não buscam promover o reuso das mesmas por diversos professores. Tanto que nenhuma das ferramentas encontradas tratam, por exemplo, de empacotar as aplicações no formato de objetos de aprendizagem. Então, abre-se a oportunidade e desafio de criar ferramentas de autoria que sejam de fácil uso e que também promovam o reuso deles pelos educadores.

V. CONCLUSÃO

O uso da RA como ferramenta de aprendizagem pode apoiar diversas disciplinas da educação. Para isso, os educadores devem contar com ferramentas de autoria capazes de facilitar todo o processo de desenvolvimento das aplicações de RA, simplificando desde a criação de conteúdos (objetos 3D, imagens e sons) até o uso nos laboratórios das instituições de ensino.

Nos trabalhos analisados ficou evidente que existe uma preocupação em criar ferramentas de autoria de fácil uso, porém que ainda existem desafios a serem superados. Além disso, nenhum desses artigos tratou da questão de reuso de tais aplicações por outros professores.

Assim, os próximos passos deste trabalho visam o desenvolvimento de uma ferramenta de autoria capaz de gerar

conteúdos de RA de fácil uso e voltadas para o reuso. Para atender ao reuso, a proposta é empacotar essas aplicações no formato de objetos de aprendizagem, no caso o padrão SCORM (*Sharable Content Object Reference Model*) [15].

REFERÊNCIAS

- [1] Irshad, S., Ramblí, D. R. A.; “User Experience Evaluation of Mobile AR services”. In: International Conference in Mobile Computing and Multimedia (MoMM), 2014. p. 119-126.
- [2] Martins, V. F. ; Guimarães, M. P. “Desafios para o uso de Realidade Virtual e Aumentada de maneira efetiva no ensino”. In: CSBC 2012 - XXXII Congresso da Sociedade Brasileira de Computação, 2012, Curitiba. DEsaIE! - I Workshop de Desafios da Computação Aplicada à Educação, 2012. p. 1-1.
- [3] Santos, M. E. C., Yamamoto, G., Taketomi T., Myiazaki J., Kato H; “Authoring Augmented Reality Learning Experiences as Learning Objects”. In: International Conference on Advanced Learning Technologies, 2013. p. 506-507.
- [4] Jee, H. K., Lim, S., Youn, J. Y.; “An Immersive Authoring Tool for Augmented Reality-based E-Learng Applications”. In: International Conference in Information Science and Applications (ICISA), 2011. p. 1-5.
- [5] Schmalstieg, D., Fuhrmann A., Szalavári, G. H. Z., Encarnaçao, L. M., Gervautz, M., Purgathofer, W.; “The Studierstube Augmented Reality Project”. In: Teleoperators and Virtual Environments February, 2002. p. 33-54.
- [6] MacIntyre, B., Gandy, M., Dow, S., Bolter, J. D.; “DART: A Toolkit for Rapid Design Exploration of Augmented Reality Experiences”. In: ACM symposium on User interface software and technology (UIST), 2004. p. 932-932.
- [7] Grimm, P., Haller, M., Paelke, V., Rheindold, S., Reimann, C., Zauner, J.; “AMIRE – Authoring Mixed Reality”. In: Augmented Reality Toolkit, The First IEEE Workshop, 2002. p. 2-2.
- [8] MAYA <http://www.autodesk.com.br/products/maya/overview>. Acessado em 17/09/2015.
- [9] Farias, L., Dantas, R., Burlamaqui A.; Educ-AR: A tool for assist the creation of augmented reality content for education”. In: International Conference Virtual Environments Human-Computer Interfaces and Measurement Systems(VECIMS), 2011. p. 1-5.
- [10] Moralejo L., Sanz, C., Pesado, P., Baldassari, S.; “AuthorAR: Authoring Tool For Building Educational Activities Based On Augmented Reality”. In: International Conference in Collaboration Technologies and Systems (CTS), 2013. p. 503-507.
- [11] PICTOGRAMAROOM
<http://www.pictogramas.org/proom/init.do?method=whatIsTab>. Acessado em 17/09/2015.
- [12] Nanayakkara, S., Schilkrot, R., Maes, P.; “EyeRing: An Eye on a Finger”. In: Conference on Human Factors in Computing Systems (CHI), 2012. p. 1047-1050.
- [13] Cubillo, J., Martín, S., Castro, M., Díaz, G., Colmenar, A., Boticki, I.; “A Learning Environment for Augmented Reality Mobile Learning”. In: Frontiers in Education Conference (FIE), 2014. p. 1-8.
- [14] Santin, R., Kirner, C., Garbin, T. R., Dainese C. A.; “Ações Interativas em Ambientes de Realidade Aumentada com ARToolKit”. <http://www.lbd.dcc.ufmg.br/colecoes/svr/2004/aumentada5.pdf>. Acessado em 17/09/2015.
- [15] Bohl, O., Schellhase, J., Sengler, R., Winand, U.; “The Sharable Content Object Reference Model (SCORM) – A Critical Review”. In: International Conference in Computers in Education (ICCE), 2002. p. 950-951.

Automação Residencial utilizando Dispositivos Móveis e Wearable Myo

Home Automation Application using Mobile Devices and Wearable Myo

Flávia Gonçalves Fernandes, João Ludovico Maximiano Barbosa, Milton Miranda Neto, Alexandre Cardoso

Faculdade de Engenharia Elétrica – FEELT

Universidade Federal de Uberlândia - UFU

Uberlândia – MG – Brasil

flavia.fernandes92@gmail.com, joao.lmbarbosa@gmail.com, voidmmn@gmail.com, alexandre@ufu.br

Abstract — The Home automation is increasingly facilitating the lives of residents. With the technological evolution of both electronics as Virtual Reality, this practice becomes more accessible, providing comfort for a larger amount of people. This work aims to present the development of a Home Automation system and virtual reality using Myo device, and also a mobile version, uniting the areas of health, engineering and technology. Therefore, the application can provide its users more comfort, convenience, time optimization due to the decrease in routine tasks and mainly safety, convenience, accessibility and economy.

Keywords - home automation; wearable device; Myo.

Resumo — A Automação Residencial está facilitando cada vez mais a vida dos moradores. Com a evolução tecnológica tanto da eletrônica como da Realidade Virtual, esta prática torna-se mais acessível, possibilitando conforto para uma maior quantidade de pessoas. Assim, este trabalho tem como objetivo apresentar o desenvolvimento de um sistema de Automação Residencial e Realidade Virtual utilizando o dispositivo Myo, e também uma versão mobile, unindo as áreas da saúde, engenharia e tecnologia. Logo, a aplicação pode proporcionar aos seus usuários maior conforto, comodidade, otimização do tempo devido à diminuição das tarefas rotineiras e principalmente segurança, praticidade, acessibilidade e economia.

Palavras-chave - automação residencial; dispositivo vestível; Myo.

I. INTRODUÇÃO

O aumento da expectativa de vida do brasileiro, o crescente número de idosos, o desejo e a necessidade de independência das pessoas e os problemas que enfrentam no dia a dia devido à idade avançada ou a alguma deficiência física, são fatores que contribuem para uma evolução significativa da automação residencial e de inovações tecnológicas, como a Realidade Virtual, possibilitando maior conforto e praticidade às pessoas que possuem alguma limitação física [1].

A tecnologia evolui para melhorar a vida das pessoas, em geral, tornando-a mais fácil e agradável. Para os portadores de deficiência, ela chegou para tornar as coisas possíveis, transformando a incapacidade em ação.

No contexto da Automação Residencial, as pessoas têm ao seu alcance uma gama de possibilidades práticas e econômicas que utilizam a automação, desde a básica até a mais abrangente, com sistemas adaptáveis para diversos espaços. O resultado pode ser um ambiente prático, confortável, agradável, mais bonito, valorizado e seguro, de acordo com o interesse do usuário [1].

A automação residencial relaciona-se com a aplicação de sistemas de controle automáticos para as funções encontradas no ambiente, integrando seus acionamentos e visando sempre a praticidade, simplicidade e objetividade dos comandos. Mas, também consideram a estética da aplicação, o conforto do usuário e a valorização do ambiente [2].

Com a evolução tecnológica, os dispositivos móveis, como smartphones e tablets, tornaram-se boas plataformas portáteis para Realidade Virtual (RV) e facilitam a comunicação via rede sem fio. Este nicho de oportunidades amplia a utilização da RV, notadamente, no contexto de melhoria das condições de vivência [3].

Dentre as vantagens de utilização desses dispositivos móveis, encontram-se o acesso à informação de forma rápida, flexível, em tempo real e de qualquer lugar. Essa praticidade motiva o desenvolvimento de aplicações nas áreas da saúde e da educação, por exemplo, visto que são setores cruciais para o desenvolvimento humano e ainda necessitam de inovações para garantir melhor qualidade de vida e formação profissional à população [4].

Este trabalho tem como objetivo principal apresentar uma ferramenta de Realidade Virtual para Automação Residencial com a finalidade de comandar ações, como ligar/desligar cargas elétricas por meio do dispositivo Myo para auxiliar as pessoas a executar atividades do cotidiano. Neste sistema, efetua-se o controle de botões por meio de gestos realizados pelo usuário que esteja utilizando o wearable Myo em um ambiente de RV e acoplado a um sistema de automação residencial, proporcionando maior interatividade e motivação durante o seu uso.

II. FUNDAMENTOS SOBRE WEARABLES

Dispositivos portáteis ou vestíveis, também conhecidos como *wearables*, são pequenos dispositivos eletrônicos que

podem ser facilmente acoplados ao corpo do usuário. Eles possibilitam a presença da Realidade Virtual em diversas aplicações, visto que é uma tecnologia de interface avançada entre um usuário e um sistema computacional [5].

O Myo é um *wearable* com formato de bracelete, também chamado de braçadeira. Este dispositivo pode controlar aplicações por meio de gestos e interagir com computadores e outros meios digitais similares, reconhecendo impulsos elétricos nos músculos do usuário. Além disso, não exige câmeras para rastrear os movimentos da mão ou braço e possuem baixo custo [6].

O Myo é utilizado no antebraço do usuário, conforme pode ser observado na Figura 1. É necessário realizar uma calibração da braçadeira para cada usuário específico, pois cada um possui atividades e contrações musculares específicas. Após efetuada a sua calibração, o *wearable* possibilita controlar softwares e outras aplicações por meio de gestos e movimentos. Seu propósito é controlar computadores, telefones e outros dispositivos, enviando os dados capturados por ele via Bluetooth [6].



Figura 1. Wearable Myo.

III. DESENVOLVIMENTO

Em relação aos aspectos metodológicos e tecnológicos, para a construção do módulo de automação residencial, utilizou-se o Arduino [7], que é uma plataforma de prototipagem eletrônica *open-source*, projetada com um microcontrolador Atmel AVR de placa única, com suporte de entrada/saída embutido, uma linguagem de programação padrão, a qual tem origem em *Wiring*, e é essencialmente C/C++.

O Arduino é o responsável pela parte de controle do sistema de automação, recebendo os comandos e os encaminhando para as cargas elétricas a serem controladas, para efeito demonstrativo, foram utilizados led's para representar o acionamento e desligamento das cargas elétricas [7].

Primeiramente, para o desenvolvimento da aplicação, foi construída uma maquete para representar a sua parte física e efetuar os testes, conforme apresentado na Figura 2.

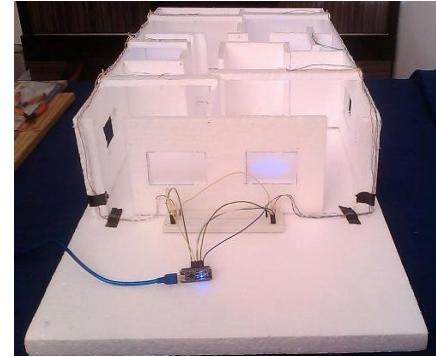


Figura 2. Maquete da aplicação.

A arquitetura de hardware utilizada neste trabalho é a de um Arduino com comunicação serial no servidor que hospeda o *WebService*. Assim, o servidor que expõe o *WebService* deve estar obrigatoriamente conectado direto ao Arduino para facilitar o envio de comandos.

Os módulos de acionamento também estão ligados diretamente ao Arduino para evitar perdas de potência relacionado à distância.

Uma das desvantagens desta arquitetura é a necessidade de ter na residência o servidor de hospedagem do *WebService* e também fazer grandes mudanças estruturais na residência, já que é preciso centralizar os módulos de acionamento e refazer a fiação das cargas elétricas a serem comandadas. Outra desvantagem é que o número de cargas elétricas que se pode comandar dependerá do número de portas de entrada/saída disponíveis no Arduino.

Na Figura 3, pode-se visualizar o esquema de ligação da arquitetura Arduino com comunicação serial.

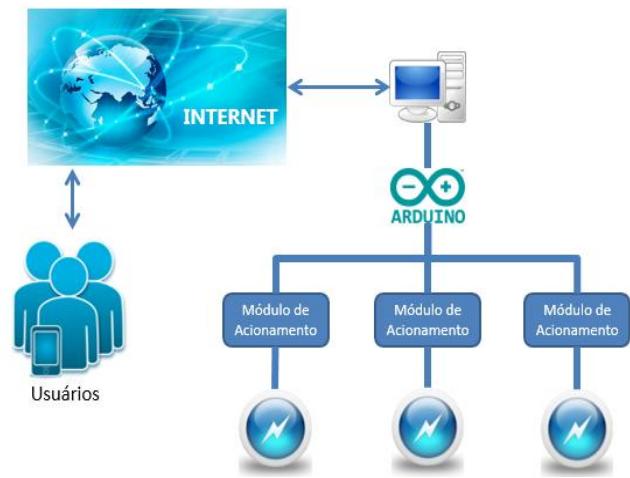


Figura 3. Arquitetura Arduino com comunicação serial.

Mas, comumente, em sistemas de automação residencial, preferem-se aqueles que não necessitem de grandes mudanças estruturais na residência para a sua instalação. Dessa forma, os sistemas que operam sem fio (*wireless*) ganham destaque, pois não necessitam fazer grandes

mudanças estruturais na residência para adaptar o sistema elétrico para sua instalação.

Desse modo, na Figura 4, é mostrado um exemplo de arquitetura genérica que trabalha sem fio.

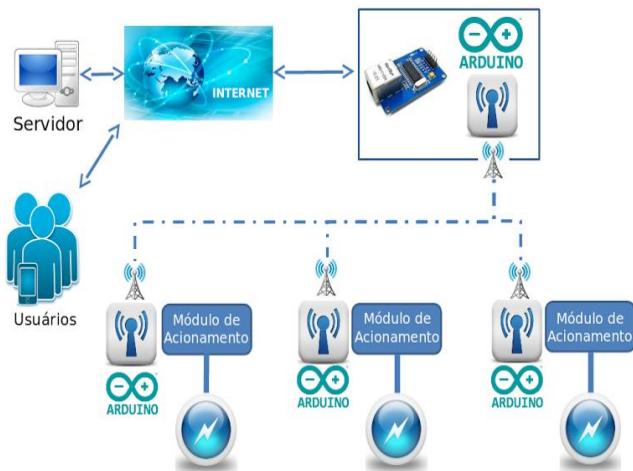


Figura 4. Arquitetura Arduino com dispositivo Wireless.

Visto que esta arquitetura trabalha sem fio, torna-se necessário utilizar dispositivos *wireless* que comuniquem entre si e com o Arduino para a troca de informações e comandos entre os módulos de acionamento e a central de comandos, exemplos desse tipo de dispositivo são: NRF24L01, módulo XBee, *WifiShield*.

Nesta arquitetura, o servidor que hospeda o *WebService* não precisa estar conectado diretamente ao Arduino devido à criação de uma central de comandos que se conecta à Internet, a qual é composta por um Arduino, um *EthernetShield* e um dispositivo de *wireless*. Isso é possível pois o *EthernetShield* pode ser utilizado como um servidor web para disponibilizar páginas simples contendo dados de sensores. O dispositivo *wireless* também é conectado ao Arduino para encaminhar e receber os comandos entre a central e os módulos de acionamento. No módulo de acionamento, o Arduino é utilizado para interpretar os comandos recebidos e ligar ou desligar as cargas elétricas.

Seguindo esta arquitetura, o número de cargas elétricas que podem ser controladas não depende mais do número de portas de entrada e saída do Arduino, mas sim do protocolo de comunicação desenvolvido para a troca de dados entre a central de comandos e os módulos acionadores.

Como os componentes do módulo de acionamento junto com o Arduino e o dispositivo *wireless* são relativamente pequenos, pode-se montar um pequeno circuito, o qual poderia ser colocado dentro da caixinha, onde fica o interruptor da carga elétrica, reduzindo ainda mais as mudanças estruturais na residência para a sua instalação.

Na Figura 5, é exibida a arquitetura dos módulos do sistema desenvolvido.

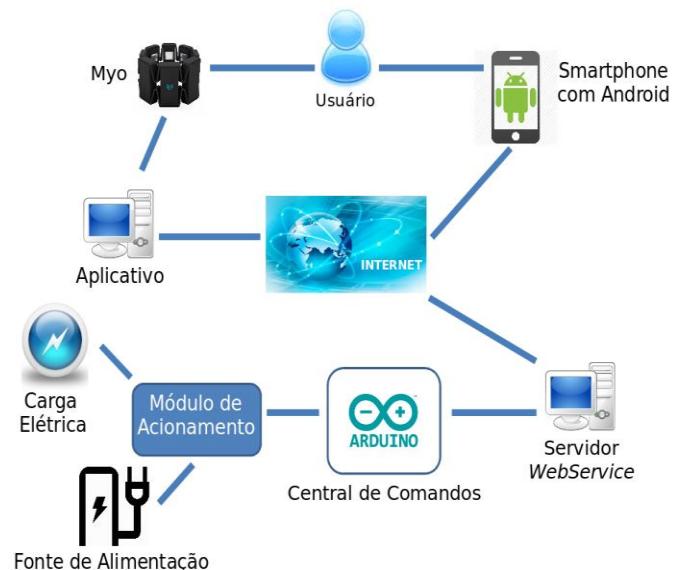


Figura 5. Arquitetura da aplicação.

O sistema apresentado funciona da seguinte maneira: o usuário pode manipular o sistema de automação residencial através de gestos utilizando o dispositivo vestível *Myo*, onde seleciona-se o cômodo da casa que deseja executar a ação para, então, acionar os comandos ligar/desligar. Estes comandos, por sua vez, são enviados para um *WebService*, feito em linguagem de programação C# utilizando o software *Microsoft Visual Studio* [8]. Posteriormente, o servidor que hospeda este *Web Service* encaminha o comando para o sistema de automação residencial através de comunicação serial.

Na Figura 6, é mostrada a interface da aplicação desenvolvida no *Visual Studio*, após o usuário ligar a lâmpada da sala 1 do protótipo. O símbolo de estalar os dedos no canto inferior direito da tela é o responsável por ligar ou desligar as cargas elétricas do sistema utilizando o *Myo*.



Figura 6. Lâmpada da sala 1 ligada.

A Figura 7 representa a interface onde todas as seis lâmpadas da residência estão ligadas.

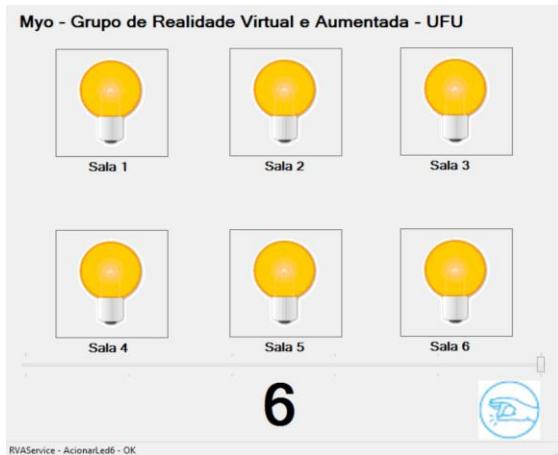


Figura 7. Todas as lâmpadas da residência ligadas.

A central de comandos do *WebService* ainda se comunica com um sistema de RV disponibilizado em dispositivo móvel (celular) com plataforma Android, o qual realiza as mesmas ações, permitindo selecionar o cômodo da residência e acionar os comandos ligar/desligar das cargas elétricas que estão sendo controladas, conforme pode ser observado na Figura 8.



Figura 8. Tela da aplicação na plataforma Android.

IV. CONCLUSÕES

Portanto, este trabalho apresentou a utilização da tecnologia de Realidade Virtual e do dispositivo *wearable* Myo combinada a Automação Residencial para aumentar a praticidade de controle de ambientes, permitindo ampliar as possibilidades de acessibilidade e comodidade.

No desenvolvimento da parte eletrônica do sistema de automação, optou-se em desenvolver uma arquitetura de comunicação serial do Arduino com o servidor que hospeda o *WebService*. No entanto, como explicitado no desenvolvimento deste trabalho, esta arquitetura possui algumas limitações que inviabilizam o desenvolvimento comercial do mesmo. Porém, devido ao custo e à complexidade optou-se em desenvolvê-la como protótipo.

Logo, o sistema desenvolvido pode proporcionar aos seus usuários maior conforto, otimização do tempo devido à diminuição das tarefas rotineiras e principalmente segurança, praticidade e economia.

Como trabalhos futuros, cita-se a implementação da arquitetura com dispositivos *wireless*, que permite desacoplar os módulos do sistema de automação residencial, fazendo com que o servidor que hospeda o *WebService* possa ficar em qualquer lugar, desde que se tenha conexão com a internet. Além disso, as mudanças estruturais para a instalação do sistema de automação residencial tornam-se praticamente nulos, pois não é preciso fazer grandes alterações na fiação elétrica da residência.

Outro trabalho futuro é o desenvolvimento do aplicativo em dispositivos móveis para outros sistemas operacionais, além do Android. Além disso, também pretende-se implementar o sistema de comunicação com o Myo para plataformas móveis, pois normalmente estes dispositivos estão mais próximos do usuário, e podem ser levados para outros lugares com facilidade. Ainda pode-se aprimorar os gestos utilizados para comandar as cargas elétricas a fim de que usuários que possuam determinadas deficiências físicas possam utilizar o sistema.

AGRADECIMENTOS

Os autores agradecem à CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) pelo apoio financeiro que permitiu a realização deste trabalho, e também ao Grupo de Realidade Virtual e Aumentada (GRVA) da Universidade Federal de Uberlândia (UFU) pelo incentivo.

REFERÊNCIAS

- [1] R. Marques. Automação residencial aquece o mercado em 2011. Disponível em: <<http://planetech.uol.com.br/2011/03/14/automacao-residencial-aquece-o-mercado-em-2011/>>. Acesso em: 30 abr. 2015.
- [2] J. R. Muratori; P. H. D. Bó. Automação residencial: histórico, definições e conceitos. Disponível em: <http://www.osetorelectrico.com.br/web/documentos/fasciculos/Ed62_fasc_automacao_capI.pdf>. Acesso em: 30 abr. 2015.
- [3] C. Kirner ; R. Siscoutto. Realidade virtual e aumentada: conceitos, projeto e aplicações. Livro do IX Symposium on Virtual and Augmented Reality. Porto Alegre: SBC, 2007. 202 p.
- [4] R. M. Costa; M. W. Ribeiro. Aplicações de realidade virtual e aumentada. Porto Alegre: SBC, 2009. 146 p.
- [5] Microsoft. Vestir-se para o Futuro: Microsoft Duo rompe com Wearable Conceito da tecnologia, Microsoft News Center, 2013.
- [6] Thalmic Labs - Myo Gesture Control Armband. Disponível em: <<https://www.thalmic.com/myo/>>. Acesso em: 30 jun. 2015.
- [7] Arduino Build Process. Arduino. Disponível em: <<http://arduino.cc/en/Hacking/BuildProcess>>. Acesso em: 30 abr. 2015.
- [8] Microsoft Visual Studio. Disponível em: <<https://www.visualstudio.com/>>. Acesso em: 30 jun. 2015.

Um Sistema de Captura de Movimento a baixo custo para Monitoramento do Idoso Institucionalizado

Luciano H. Quintino

Universidade Tecnológica Federal do Paraná
Cornélio Procópio, Paraná

Eduardo F. Damasceno

Universidade Tecnológica Federal do Paraná
Cornélio Procópio, Paraná

Adriane C. A. Silva

Universidade Tecnológica Federal do Paraná
Cornélio Procópio, Paraná

Eduarda Almeida

Universidade Tecnológica Federal do Paraná
Cornélio Procópio, Paraná

Resumo—Este trabalho apresenta um sistema de captura de movimento baseado em captura de vídeo em tempo real com visão computacional aplicado ao monitoramento de idosos institucionalizados. Tal sistema consiste em basicamente monitorar idosos institucionalizados em um ambiente doméstico, e alertar o cuidador na possibilidade de ocorrência de situações que coloquem em risco à sua integridade física, como por exemplo, quedas. Atualmente existe uma lacuna na utilização de tecnologias mais modernas em áreas da saúde que não pertencem ao ambiente hospitalar e às práticas cirúrgicas. É relatado aqui um estudo feito na intenção de melhorar a velocidade e a precisão na prestação de serviços de emergência para idosos institucionalizados que sofrem quedas, as quais podem causar uma quantidade significativa de sequelas caso não sejam rapidamente socorridas.

Abstract—This paper presents a motion capture system based on video capture in real time with computer vision applied to the monitoring of elderly patients. This system consists of track body of institutionalized elderly in a home environment, and alert the caregiver in the possibility of situations that will endanger their physical integrity, such as falls. Today yet exist a gap in the use of latest technologies in health areas that do not belong to the hospital and surgical practices. It is reported here a study on the intention to improve the speed and accuracy in providing emergency services for institutionalized elderly people suffering falls, which can cause a significant amount of sequels if not quickly rescued.

I. INTRODUÇÃO

O envelhecimento populacional é considerado um fenômeno mundial. No passado, ter uma idade avançada era um privilégio de poucos, no entanto passou a ser uma experiência cada vez mais crescente entre pessoas de todo o mundo [1]. De acordo com o IBGE (Instituto Brasileiro de Geografia e Estatística), no Brasil, é estimado que a população idosa dobra no período de 2000-2025, chegando a 30 milhões de pessoas com mais de 60 anos [2]. Além disso, haverá seis idosos para cada cinco crianças com menos de cinco anos de idade e possivelmente uma deficiência no número de médicos geriatras para atender a essa população.

Grande parte destas pessoas já sofreram algum acidente doméstico, mas dependendo do acidente, pode ser fatal. Este acidente normalmente é uma queda [4]. As quedas, geralmente, são responsáveis pelas perdas da autonomia e da independência do idoso, suas consequências mais comuns são: as fraturas, a

imobilidade, a restrição de atividades e declínio geral da saúde [3].

Muitos idosos no Brasil, não possuem condições para privatizar alguém para dar assistência 24 horas por dia, e por isso recorrem as casas ou clínicas de saúde especializadas em tratamento geriátrico. Muitas delas recebem o simpático nome de “Lar do Vovó”, ou “Casa de Descanso”, mas ainda persiste o nome no consciente da população de “Asilo”. Esses lugares especializados possuem um grupo de profissionais adequado para atendimento ao volume de idosos, todavia, no período de maior pico de movimento ou nos períodos noturnos este volume de profissionais tendem a diminuir. E é nestes períodos em que se concentram os horários de maior proporção de atendimento à acidentes domésticos como a queda. Por causa disso, ao ocorrer um acidente, não tem ninguém naquele momento para socorrer a vítima.

Atualmente a tecnologia tem avançado muito, sendo estes avanços em sua maioria benéficos à população em geral. Em contrapartida, em frente à vasta quantidade de possíveis aplicações destas novas tecnologias, terminam por existir algumas áreas de pesquisa onde a aplicação de novas tecnologias são pouco desenvolvidas.

A assistência domiciliar monitorada, representa uma solução interessante tanto para evitar um colapso no sistema hospitalar, quanto para promover a saúde em casa. No entanto, os ganhos potenciais com essa mesma, exigem infraestruturas de hardware e software nas residências para atender a essa demanda, e melhorar a qualidade no gerenciamento das informações produzidas. Este trabalho se dedica ao estudo de um novo meio de monitoramento de idosos com foco no pronto atendimento em caso de quedas ou acidentes corriqueiros de modo não invasivo ao idoso e muito menos intrusivo em sua privacidade.

II. DEFINIÇÃO DO PROBLEMA

O crescimento exponencial de desenvolvimento tecnológico que vem ocorrendo nas últimas décadas tem um efeito amplo nas mais diversas áreas de aplicação, possibilitando realizar experimentos que antes não se imaginavam possíveis. Estes avanços têm um impacto direto na qualidade de vida da população e consequentemente na expectativa de vida média desta população.

Quando comparado com épocas passadas, atualmente é muito maior a chance de uma pessoa chegar a uma idade mais avançada, sendo isto refletido no aumento acelerado de pessoas na faixa etária acima dos sessenta anos em países desenvolvidos [4].

Este é fato muito encorajador em geral, mas deve-se levar em conta os fatores negativos relacionados a esse assunto, como a degeneração em geral da saúde, causada pelo avanço da idade de uma pessoa.

Algumas das debilidades que vem com o avanço da idade são: a fraqueza muscular, problemas de visão, perda de equilíbrio, menor capacidade de reação e desgaste ósseo [5]. Estas características do envelhecimento dificultam a vida dos idosos e faz com que muitos procurem instituições especializadas para dar suporte a esses idosos, para assim evitar os riscos relacionados a estas fraquezas.

Podemos definir a queda como sendo "o deslocamento não intencional do corpo para um nível inferior à posição inicial com incapacidade de correção em tempo hábil, determinado por circunstâncias multifatoriais, comprometendo a estabilidade" [6].

Para os idosos, as quedas se constituem em um importante problema de saúde pública em decorrência da mortalidade e do elevado custo social e econômico resultante das lesões provocadas [7]. Sabe-se que na grande maioria dos casos de lesões causadas por quedas foram justificados por escorregões e tropeços [6], muitas vezes levando cerca de dez minutos para que ocorra o atendimento, mesmo na maioria dos casos a queda causando fraturas e até traumatismos cranianos [8].

As consequências das quedas para os idosos podem ser bastante limitadoras, em alguns casos, fatais [9]. A incapacidade de se levantar sozinho após uma queda leva a um período conhecido por "long lie", ou seja, a longa permanência no chão após a queda [7][8]. Isto pode trazer complicações como desidratação, pneumonia, escaras e medo de outras quedas [10].

Sabe-se que os índices de recuperação destas lesões são maiores quando o atendimento é mais rápido, e em caso que se demora mais de 30 minutos de atendimento, não se observa recuperação [6].

Com índices significativos de sequelas e em menor número, mortes em idosos que sofrem quedas, surge a necessidade de se desenvolver uma ferramenta para realizar o monitoramento de pessoas idosas, que consiga avisar que a mesma tenha sofrido uma queda para que assim possa ser socorrida o mais breve possível, visando assim evitar essas possíveis consequências.

Com este objetivo em foco, foi realizada a busca na literatura científica sobre sistemas de monitoramento de quedas automatizado.

III. TRABALHOS ANTERIORES

Em nossa pesquisa, foram traçados parâmetros que orientassem o estudo inicial para otimizar nosso busca por resultados tanto em produtos de software quanto em arquiteturas para dar suporte a solução de nosso problema. Destacam-se basicamente três métodos de monitoramento: i) monitoramento do ambiente; ii) monitoramento do corpo; iii) monitoramento

combinado (sinais biomédicos acompanhados de dados do ambiente); e; iv) monitoramento por acionamento de dispositivo.

A. Monitoramento do Ambiente

Quando se aplica esta técnica de monitoramento é necessário equipar o local com uma infraestrutura que possibilite a coleta de dados. Este tipo de monitoramento foi encontrado nos trabalhos [11][12][13]. A técnica aplicada é o monitoramento por meio de visão computacional e reconhecimentos de padrões de movimento ou variação de movimento em áreas de risco de queda [13]. A principal vantagem deste tipo de monitoramento é a de não ser necessário a fixação de qualquer equipamento no corpo do indivíduo. Uma desvantagem no sensoriamento do ambiente é a necessidade de uma área coberta pelo sistema de rastreamento, o que torna essa abordagem difícil de ser aplicada em ambientes externos. Ademais uma outra desvantagem é o processamento necessário para o reconhecimento de padrão a cada nova imagem adquirida pelas câmeras.

B. Monitoramento do Corpo

Nesta abordagem são agrupados os trabalhos que medem o movimento do corpo por meio de dispositivos conectados a ele. Estes dispositivos são denominados de IMU, (do inglês Inertial Measurement Units) são os sensores habilitados a capturar o movimento a 6 ou 9 graus de liberdade. Para isso possuem acelerômetro, giroscópio, magnetômetro triaxiais, e coletam simultaneamente os dados do movimento.

Nos trabalhos [14][15][16] relatam esta abordagem de maneira muito eficiente, evidenciando a principal vantagem desta prática: não pressupor a existência de infraestrutura no ambiente. Isto possibilita ao indivíduo realizar suas atividades em qualquer ambiente. Já as principais desvantagens são a necessidade de vestir equipamentos, o que exige um bom projeto do wearable [14], a possibilidade de descalibragem dos sensores [15], além do desconforto em anexar ao corpo equipamentos e dispositivos como eletrodos [16].

C. Monitoramento Combinado

Como o próprio nome diz é uma mistura fluída entre a captação dos dados biomédicos adquiridos por dispositivos ligados ao corpo [17][18] e interpretado os dados do monitoramento do ambiente para identificar o status de saúde no momento do acidente [19][20][28]. As vantagens da união das duas abordagens anteriores são muito relevantes, todavia o custo e a calibração de vários dispositivos são uma razão desmotivadora.

D. Monitoramento por acionamento de dispositivo.

Este sistema é o mais conhecido atualmente, sendo produzido por algumas empresas [21][22][23], o qual é composto por três elementos: O botão transmissor, o console e a central de atendimento. Por meio de um console instalado no local e que, com um único toque, solicita ajuda no caso de alguma emergência. Além disso, também oferece um botão de emergência pessoal, sem fio e à prova d'água, que pode ser utilizado em forma de colar ou pulseira para que quem o utiliza possa circular por todos os cômodos da casa.

Caso o idoso necessite de ajuda, basta apertar qualquer um dos botões para avisar a central de atendimento que recebe o sinal dos usuários, conversa com o usuário (através do console) e chama ajuda, que pode ser um vizinho, parente, médico ou serviço de ambulância.

IV. NOSSA PROPOSTA

Nossa abordagem utiliza o sensor de movimento MS-Kinect, que nos últimos anos se tornou uma ferramenta amplamente utilizada em aplicações de captura de movimento a baixo custo. Ademais a escolha deste dispositivo permeou algumas de suas características tais como componentes robustos que permitem a captura de imagens com informação de profundidade mesmo em ambientes escuros, por fazer uso de infravermelho [27].

Diversos estudos fizeram uso dessa ferramenta com o objetivo de detectar quedas. As técnicas mais utilizadas foram a da análise da distância em relação ao chão[24], o uso do cálculo da velocidade do corpo em direção ao chão [25] ou ambas as técnicas [26].

A. Arquitetura de Software

O sistema é composto por várias estações remotas, nas quais são compostas por um computador e um sensor Kinect. O sistema utiliza o conjunto de câmeras fornecidas pelo Kinect para detectar e rastrear movimentos dos idosos no quarto. Os dados são capturados pelo Kinect na resolução de 640x480 pixels, com uma taxa de 30 quadros por segundo, esses dados são enviados para o computador via conexão USB, que são analisados e então enviados a central de monitoramento via conexão de rede. A Figura 1 ilustra a interação entre os dispositivos utilizados.

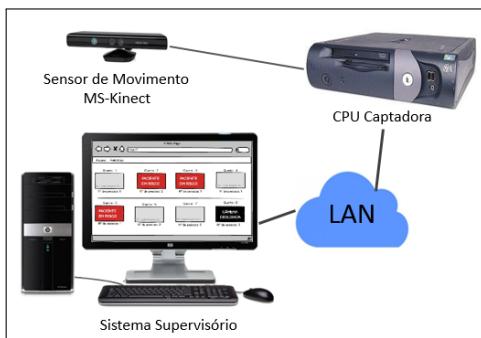


Figura 1. Intereração entre os dispositivos

A central de monitoramento é composta por um servidor web e um banco de dados, no qual é utilizado para armazenar os dados recebidos das estações remotas. O cuidador acessará central de monitoramento por meio de um navegador web. A Figura 2 ilustra a arquitetura do sistema.

B. Sistema de Supervisão

O sistema de supervisão foi desenvolvimento em ambiente web de forma a atender um maior número de dispositivos. Para a comunicação em tempo real com as estações de monitoramento foi utilizado a biblioteca ASP.NET SignalR que

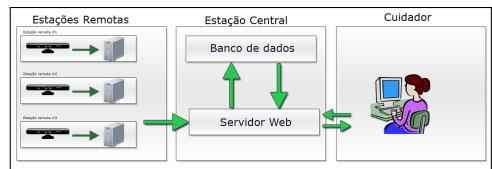


Figura 2. Arquitetura do Sistema

possui suporte a web sockets e a outras técnicas compatíveis com navegadores mais antigos.

O sistema dispõe de uma área de monitoramento, relatórios e gráficos de quedas, cadastros de quartos e gerenciamento de usuários. A Figura 3 e 4 mostram uma visualização dos gráficos de alertas e quedas gerados pelo sistema.

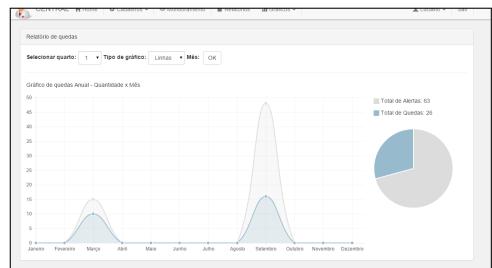


Figura 3. Gráfico de Alerta e Quedas Anual

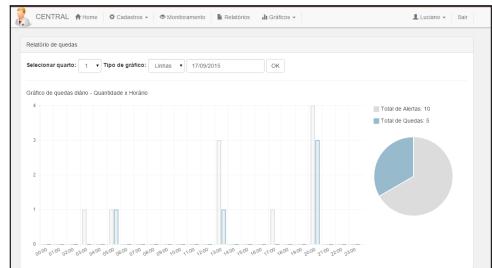


Figura 4. Gráfico de Alerta e Quedas Diário

A área de monitoramento do sistema permite ao cuidador visualizar informações das pessoas rastreadas e a sua posição em coordenadas cartesianas 2D. Na Figura 5 observamos o monitoramento de uma pessoa em estado normal e outra em alerta.



Figura 5. Área de Monitoramento

C. Regras de Avaliação de Queda

Para a avaliação das quedas foram utilizados dois algoritmos que verificam a movimentação de um esqueleto em um certo espaço de tempo. O primeiro algoritmo é responsável por armazenar os dados dos esqueletos em um vetor e analisar a altura das suas articulações superiores do momento atual em relação às articulações armazenadas nos momentos anteriores. Caso haja uma variação negativa (em relação ao chão) superior a 50% entre os dois momentos, o esqueleto é situado como estado de alerta.

O segundo algoritmo é responsável por fazer a análise dos esqueletos que se encontram em estado de alerta, de modo a verificar se o movimento que ocorreu é uma queda ou um falso-positivo. Para tal, foi verificado durante um instante 3 segundos se houve uma variação positiva superior a 10% da altura das suas articulações atual em relação à altura das suas articulações no momento em que entrou em estado de alerta. Caso não haja uma variação superior ao percentual determinado, o esqueleto analisado entra em estado de queda.

A Figura 6 exemplifica o armazenamento dos quadros em vetores e a verificação entre os dados do quatro atual e os dados dos quadros já armazenados em momentos anteriores.

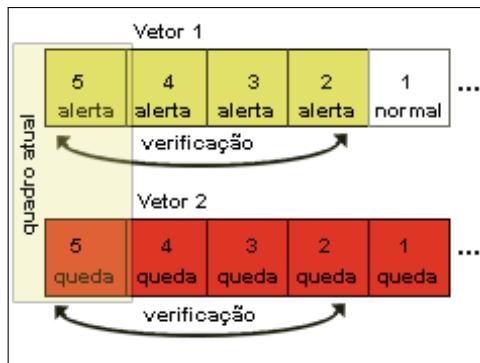


Figura 6. Armazenamento e verificação dos dados

D. Avisos ao Cuidador

Há central de monitoramento possui duas formas de alertar o cuidador, por cores e áudio. As cores alertam sobre o estado em que se encontra o quarto, sendo branco normal, amarelo alerta e vermelho queda. Sempre que uma queda é detectada um alerta sonoro é disparado na central de monitoramento. O alerta sonoro é essencial, visto que o cuidador às vezes tenha que sair da frente da tela do computador, desse modo é possível saber de uma queda no mesmo instante, possibilitando o socorro de forma rápida.

E. Testes de Regras de Queda

Para verificar a eficiência do sistema foram feitas simulações em diferentes situações e ambientes, de forma a conseguir resultados mais precisos e possibilitar uma melhor análise dos dados.

A primeira simulação foi feita com um indivíduo posicionado em frente ao sensor Kinect de forma a simular quedas frontais, traseiras e laterais em ambas as direções. O Kinect foi posicionado a 2 metros do usuário e a uma altura de 1 metro

em um ambiente claro. A Figura 7 exibe a simulação de uma queda lateral, os dados obtidos nessa simulação se encontram na Tabela I.

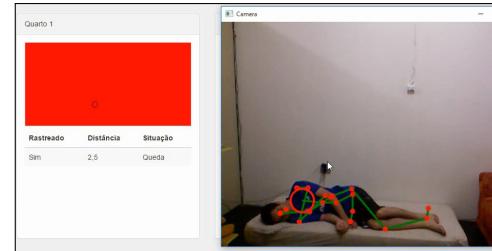


Figura 7. Queda lateral em frente ao Kinect

Tabela I. RESULTADOS - QUEDAS EM FRENTE AO KINECT

Queda	Esquerda	Direita	Frontal	Traseira
1	Detetada	Detetada	Detetada	Detetada
2	Detetada	Detetada	Detetada	Detetada
3	Detetada	Detetada	Detetada	Detetada
4	Detetada	Detetada	Detetada	Detetada
5	Detetada	Detetada	Detetada	Detetada
6	Detetada	Detetada	Detetada	Detetada
7	Detetada	Falha	Detetada	Detetada
8	Detetada	Detetada	Detetada	Detetada
9	Falha	Detetada	Falha	Detetada
10	Falha	Detetada	Falha	Falha

O segundo teste foi feito com uma pessoa em movimentação, simulando uma queda ao entrar em um quarto com a luz apagada. O Kinect foi posicionado a uma distância de três metros da porta e a uma altura de um metro e meio. A Figura 8 mostra o momento de uma queda no escuro. Os dados referentes a este teste se encontram na Tabela II

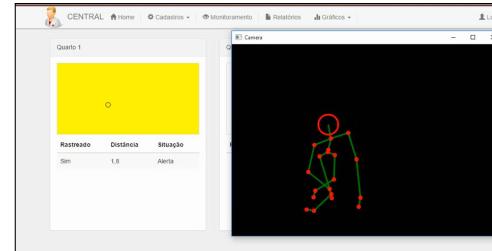


Figura 8. Queda em movimentação na frente do Kinect

Tabela II. RESULTADOS - QUEDA EM MOVIMENTAÇÃO NA FRENTE DO KINECT

Queda	Resultado
1	Detetada
2	Detetada
3	Detetada
4	Detetada
5	Detetada
6	Detetada
7	Detetada
8	Detetada
9	Detetada
10	Falha

No terceiro teste foi simulado quedas ao se deitar e levantar

da cama em um ambiente escuro. Nesse caso o Kinect foi posicionado a um metro de distância da cama e a uma altura de 2 metros para conseguir capturar a pessoa deitada. Os resultados obtidos são mostrados na Tabela III em quanto que a Figura 9 exibe a captura de uma pessoa deitada.

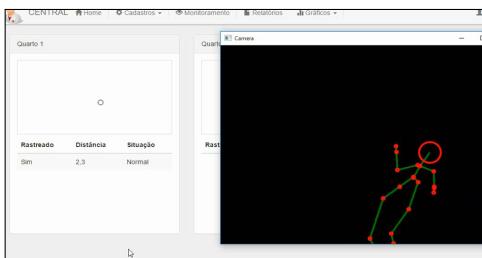


Figura 9. Pessoa deitada capturada pelo Kinect

Tabela III. RESULTADOS - QUEDAS AO DEITAR E LEVAR DA CAMA

Queda	Deitar	Levantar
1	Falha	Falha
2	Detectada	Falha
3	Falha	Falha
4	Falha	Detectada
5	Detectada	Falha
6	Falha	Falha
7	Falha	Falha
8	Detectada	Falha
9	Falha	Falha
10	Falha	Falha

Na última simulação foi criado uma situação em que uma pessoa esta caída e uma segunda tenta fazer o socorro, porém a mesma também cai. O Kinect foi posicionado a uma distância de dois metros e meio a uma altura de um metro e meio. A Figura 10 mostra o momento em que uma segunda pessoa tenta fazer o socorro. Os resultado são apresentados na Tabela IV.

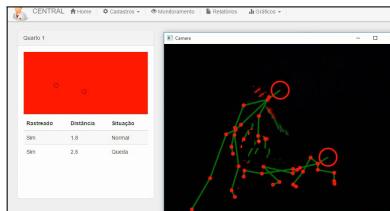


Figura 10. Captura de duas pessoas pelo sensor

Tabela IV. RESULTADOS - UMA SEGUNDA PESSOA CAINDO AO PRESTAR SOCORRO

Queda	Resultado
1	Falha
2	Falha
3	Falha
4	Falha
5	Falha
6	Falha
7	Falha
8	Falha
9	Falha
10	Falha

V. ANÁLISE DOS RESULTADOS

Podemos observar com base nos dados das Tabelas I, II e III que o Kinect comporta-se melhor a uma altura aproximadamente entre um metro e um metro e meio. A melhor posição obtida foi a frontal, com pessoa se movendo em direção as lentes do sensor, conforme a Tabela II. A maioria das falhas ocorreram devido a uma perca aleatória dos pontos do esqueleto e a uma limitação de capacidade do Kinect, no qual o mesmo não consegue plotar corretamente o esqueleto em certas situações devido a diversos fatores externo, como distância da câmera, ângulo, altura, quantidade de objetos encontrado no campo de visão, posicionamento do esqueleto entre outros. Essas falhas ficaram evidentes nos dados das Tabelas III e IV no qual o sensor foi submetido a um teste com algumas das situações mencionadas.

O Kinect também apresentou uma falha ocasional que ocorre quando a pessoa rastreada pelo sensor sai do seu campo de visão, ocasionando na perca da referência do esqueleto e o mesmo acaba sendo plotado em um lugar aleatório, devido as regras definidas o sistema detecta uma sinal de alerta indevidamente, porém uma queda não é acionada devido ao segundo algoritmo de verificação, amenizando um pouco o problema.

Outro dado importante foi a capacidade do Kinect detectar o esqueleto no escuro, o mesmo não mostrou um diferença significante em relação a esse fator, conforme as Tabelas I e II.

O Kinect também conseguiu capturar mais de uma pessoa ao mesmo tempo conforme a Figura 10, porém devido a suas limitações já mencionadas o mesmo não conseguiu plotar duas pessoas juntas e caídas ao mesmo tempo, ocasionando muitas vezes em perca do esqueleto ou esqueletos fantasmas.

Para a realização dos testes do presente trabalho foi utilizado a primeira versão do Sensor Kinect. Acredita-se que nova verão do dispositivo lançado para o console Xbox One, apresentaria um algoritmo mais eficiente, diminuindo das limitações mencionadas.

VI. CONCLUSÃO

Neste trabalho abordamos a utilização do Kinect como um sensor de baixo custo para monitoramento de idosos institucionalizados e concluímos que os resultados obtidos foram positivos, uma vez que o sistema proposto foi capaz de capturar e analisar os dados para a verificação de quedas, entretanto tal sistema está longe de poder ser utilizado como solução única para o problema abordado. O mesmo pode ser aprimorado de várias formas em pesquisas futuras, desenvolvendo regras mais eficientes e extraíndo mais informações de padrões de quedas, além da utilização de uma versão mais aperfeiçoada do sensor, por exemplo a versão do Xbox One, que possui algoritmos e hardware mais avançado.

Este trabalho foi de grande importância, uma vez que foi possível compreender melhor o funcionamento do sensor Kinect e ter permitido a utilização do dispositivo para outros meios afora da área de entretenimento.

Para trabalhos futuros destaca-se a importância da utilização de mais de um sensor de monitoramento. Uma vez que

os dados obtidos neste estudo foram decorrentes da utilização de apenas um sensor. A utilização de vários sensores poderia melhorar quanta a ocorrência de falso-positivo e problemas que envolve o campo de visão do sensor. Outra abordagem que poderia ser interessante, é a utilização do dispositivo de áudio acoplado do Sensor Kinect, tal recurso poderia melhorar o algoritmo quanto as ocorrências de falso-positivo e possibilitar interação com a central de monitoramento por meio de comando por voz. No entanto tais implementações seriam mais complexas, sendo estas propostas adequadas para futuras pesquisas.

REFERÊNCIAS

- [1] L. Sousa, H. Galante, and D. Figueiredo, “Qualidade de vida e bemestar dos idosos: um estudo exploratório na população portuguesa,” RevSaúde Pública, vol. 37, no. 3, pp. 364–71, 2003.
- [2] IBGE (2008). “Projeção da População do Brasil: População brasileira envelhece em ritmo acelerado”, http://www.ibge.gov.br/home/presidencia/noticias/noticia_impressao.php?id_noticia=1272, [mar / 2013].
- [3] Fabrício, S.C.C.; Rodrigues, R.A.P.; Costa Júnior, M.L. Causas e consequências de quedas de idosos atendidos em hospital público. Revista de Saúde Pública 2004 fev; 38(1): 93-9.
- [4] F. Júnior and O. da Silva, “Queda de idosos que motiva atendimento hospitalar de emergência pelo sus em überlândia-mg: epidemiologia e consequências para a saúde,” 2006.
- [5] E. d. S. F. Coutinho and S. D. da Silva, “Uso de medicamentos como fator de risco para fratura grave decorrente de queda em idosos medication as a risk factor for falls resulting in severe fractures in the elderly,” Cad. Saúde Pública, vol. 18, no. 5, pp. 1359–1366, 2002.
- [6] S. Buksman, A. Vilela, S. Pereira, V. Lino, and V. Santos, “Quedas em idosos: Prevenção,” PROJETO Diretrizes. Associação Médica Brasileira/Conselho, 2008.
- [7] L. Z. Rubenstein and K. R. Josephson, “The epidemiology of falls and syncope,” Clinics in geriatric medicine, vol. 18, no. 2, pp. 141–158, 2002.
- [8] R. W. Sattin, “Falls among older persons: a public health perspective,” Annual review of public health, vol. 13, no. 1, pp. 489–508, 1992.
- [9] M. C. Nevitt, S. R. Cummings, and E. S. Hudes, “Risk factors for injurious falls: a prospective study,” Journal of gerontology, vol. 46, no. 5, pp. M164–M170, 1991.
- [10] M. E. Tinetti, D. I. Baker, G. McAvay, E. B. Claus, P. Garrett, M. Gottschalk, M. L. Koch, K. Trainor, and R. I. Horwitz, “A multifactorial intervention to reduce the risk of falling among elderly people living in the community,” New England Journal of Medicine, vol. 331, no. 13, pp. 821–827, 1994.
- [11] Lee, M. S., Trajkovic, M., Dagtas, S., Gutta, S., Brodsky, T., Philomin, V., ... & Cohen-solal, E. U.S. Patent No. 20,030,058,111. Washington, DC: U.S. Patent and Trademark Office.
- [12] Mihailidis, A., Carmichael, B., & Boger, J. The use of computer vision in an intelligent environment to support aging-in-place, safety, and independence in the home. Information Technology in Biomedicine, IEEE Transactions on, 8(3), 238-247.
- [13] DAMASCENO, E.F. ; Goulart E. ; Moraes, R de . Um Sistema de Monitoramento para Idosos Utilizando Técnicas de Visão Computacional. In: Congresso da Sociedade Brasileira de Computação, 2013, Maceió - AL. Anais do Encontro Nacional do Estudantes de Computação das Escolas Técnicas Federais, 2013. v. 1. p. 23-25.
- [14] Ugulino, W., Ferreira, M., Velloso, E., & Fuks, H. (2012, October). Virtual Caregiver: A System for Supporting Collaboration in Elderly Monitoring. In Collaborative Systems (SBSC), 2012 Brazilian Symposium on (pp. 43-48). IEEE.
- [15] Carvalho, S. T., Erthal, M., Mareli, D., Sztajnberg, A., Copetti, A., Loques, O., & de Janeiro-RJ-Brasil, R. (2010). Monitoramento remoto de pacientes em ambiente domiciliar. XXVIII Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos-Sala de Ferramentas, Gramado, RS, Brasil, 1005-1012.
- [16] MICHAELLES, F. & SCHIELE, B. “Sensing and monitoring professional skiers”. IEEE Pervasive Computing, 4(3):40–46, 2005.
- [17] Cerny, M. Movement Monitoring in the HomeCare System. In World Congress on Medical Physics and Biomedical Engineering, September 7-12, 2009, Munich, Germany (pp. 356-357). Springer Berlin Heidelberg.
- [18] De Moraes, A. H.F.; De Medeiros V R.A.; Brandão, G.B.. Esclerose Lateral Amiotrófica: Uma Revisão Do Estado Da Arte Para Sistemas De Monitoramento Em Homecare. Revista Brasileira de Inovação Tecnológica em Saúde ISSN: 2236-1103, v. 2, n. 4, 2013.
- [19] Culhane, K. M., O'Connor, M., Lyons, D., & Lyons, G. M. Accelerometers in rehabilitation medicine for older adults. Age and ageing, 34(6), 556-560.
- [20] Costin, H., Rotariu, C., Morancea, O., Andtuseac, G., Cehan, V., Felea, V., ... & Costin, C. (2008). Complex telemonitoring of patients and elderly people for telemedical and homecare services. New Aspects of Biomedical Electronics and Biomedical Informatics, 183-187.
- [21] TeleHelp. “telehelp - sempre com você”. [Online]. Available:<http://www.telehelp.com.br/>
- [22] IrisSenior. “Iris Senior”. [Online]. Available:<http://www.irissenior.com.br/>
- [23] Philips. “philips lifeline”. [Online]. Available:<http://www.lifelinesys.com/>
- [24] G. Mastorakis and D. Makris, “Fall detection system using kinects infrared sensor,” Journal of Real-Time Image Processing, pp. 1–12, 2012.
- [25] C. Rougier, E. Auvinet, J. Rousseau, M. Mignotte, and J. Meunier, “Fall detection from depth map video sequences,” in Toward Useful Services for Elderly and People with Disabilities. Springer, 2011, pp. 121–128.
- [26] C. Kawatsu, J. Li, and C. Chung, “Development of a fall detection system with microsoft kinect,” in Robot Intelligence Technology and Applications 2012. Springer, 2013, pp. 623–630.
- [27] ZHANG, Zhengyou. Microsoft kinect sensor and its effect. MultiMedia, IEEE, v. 19, n. 2, p. 4-10, 2012.
- [28] Oliveira, M., Mesquita, P. A., Andrade, L. O. M., Figueiredo, V., Bringel Filho, J., & Hairon, C. Pimentel-L: o Computador 24 horas Sensível ao Contexto para Aplicações em Saúde

ARFoodMenu: Augmented Reality in Food Establishments

ARFoodMenu: Realidade Aumentada em Estabelecimentos Alimentícios

João Otávio Brandão Antunes de Lucena
Statistic and Computer Science Department
Federal Rural University of Pernambuco
Recife, Brazil
joao_otavio93@live.com

João Paulo Silva do Monte Lima
Statistic and Computer Science Department
Federal Rural University of Pernambuco
Recife, Brazil
jpsml@cin.ufpe.br

Daniel Lima Nogueira
Statistic and Computer Science Department
Federal Rural University of Pernambuco
Recife, Brazil
daniellnog@gmail.com

Abstract— This paper aims to present the ARFoodMenu application, a mobile device system which was developed in order to increase the level of entertainment and interactivity of customers in food establishments such as bars and restaurants by using Augmented Reality technology. As main functionalities, the app brings the display of promotion or event in the current period, the interactive virtual menu and the display of a dish in 3D format. Throughout the paper it will be discussed how these features are able to perform that.

Key-words—augmented reality; mobile device; food establishments; promotion or event; interactive virtual menu; 3D format.

Resumo— Este artigo tem por finalidade apresentar o aplicativo ARFoodMenu, um sistema para dispositivos móveis desenvolvido com o propósito de aumentar o nível de entretenimento e interatividade de clientes em estabelecimentos alimentícios como bares e restaurantes através do uso da tecnologia de Realidade Aumentada. Como principais funcionalidades, o aplicativo traz a visualização da promoção ou evento do período atual, o cardápio virtual interativo e a visualização do prato em 3D. Ao longo do artigo será discutido como essas funcionalidades são capazes de realizar essa melhoria.

Palavras-chave—realidade aumentada; dispositivos móveis; estabelecimentos alimentícios; promoção ou evento; cardápio virtual interativo; prato em 3D.

I. INTRODUÇÃO

Não é difícil perceber o quanto é complicada a relação entre fornecedor e cliente nos dias atuais. No ramo alimentício, não é diferente. Um restaurante não pode limitar-se apenas a oferecer bons pratos, é preciso fornecer aos clientes ferramentas que tornem a sua estadia no estabelecimento

prática, simples, interativa e segura, desde o pedido até o consumo do prato.

Atualmente, o número de estabelecimentos alimentícios existentes é elevado e cada um dispõe de um formato de cardápio específico. Contudo, a organização e a disposição dos pratos nesses artefatos muitas vezes são confusas. A desordem da lista de pratos, a ausência de nitidez em cores e letras e a má distribuição das categorias dos pratos são fatores que influenciam negativamente o manuseio do cliente. Outro aspecto negativo é a quantidade limitada de informações expostas sobre o prato. É importante que o cliente possa visualizar dados nutricionais e ingredientes, o que não se vê comumente nesses artefatos, a fim de evitar consumir componentes prejudiciais à saúde ou díspares de sua alimentação regular. Partindo dos problemas supracitados, o aplicativo ARFoodMenu, tem por objetivos contribuir com a solução de algumas dessas questões e aumentar o nível de interatividade e entretenimento dos clientes com funções baseadas na tecnologia de Realidade Aumentada (RA). São três funções: visualização da promoção ou evento, visualização do prato em 3D e cardápio virtual interativo.

Seguidamente (Seção II), este artigo relaciona o trabalho exposto com outros semelhantes e ressalta as contribuições do presente trabalho no âmbito do estado da arte. A Seção III desenvolve um breve resumo sobre o aplicativo ARFoodMenu, a metodologia utilizada para o desenvolvimento das funções de RA e descreve o funcionamento de cada uma das funcionalidades. Na Seção IV, levanta os principais aspectos que conferem utilidade à aplicação. A Seção V apresenta as conclusões obtidas, assim como possíveis trabalhos futuros.

II. TRABALHOS RELACIONADOS

Várias aplicações existentes que fazem uso da tecnologia de RA em bares e restaurantes já circulam no mercado. Acrescentaremos, aqui, alguns aspectos em que o ARFoodMenu avança relativamente a elas.

O Auggy é um aplicativo para dispositivos móveis que utiliza RA rastreando um cardápio real e adicionando conteúdo virtual sobre o mesmo [1]. O modelo 3D do prato é exibido sobre o cardápio real, além de informações nutricionais e uma avaliação do prato em formato de estrelas. Já o ARFoodMenu possui um cardápio com um nível de interatividade maior, onde todos os pratos são exibidos com a detecção de apenas um marcador baseado em imagem.

O Ogle é um aplicativo para dispositivos móveis também usado em restaurantes [2]. A partir do rastreamento de uma imagem, o Ogle fornece uma interação com o usuário que utiliza uma tela sensível ao toque por meio da qual é possível interagir com o conteúdo virtual exibido. Diferentemente do ARFoodMenu, o Ogle é um aplicativo voltado para o entretenimento do cliente.

O WatchAR é um aplicativo de RA para dispositivos móveis que adiciona animações no mundo real partindo também de um rastreamento baseado em imagem [3]. A imagem utilizada é um cardápio promocional de um restaurante e é possível interagir com os itens registrados. O WatchAR tem um foco maior na diversão e interatividade no usuário. No ARFoodMenu, é possível visualizar o cardápio por categoria dos pratos em formato de paginação e pratos em 3D, além de informações adicionais acerca do prato selecionado.

A mesa de menu interativo baseado em projeção [4] encontrada no restaurante Inamo, em Londres, também é um exemplo de trabalho relacionado. Projetores são utilizados para exibir o cardápio do restaurante em cima de uma mesa, e o usuário pode interagir com o mesmo a partir do toque, escolhendo qual prato deseja e realizando o pedido. Mesmo sendo mais completo em relação ao ARFoodMenu, a mesa de menu interativo necessita de mais recursos, como projetores e mesas especiais para a sobreposição das projeções do cardápio. O ARFoodMenu é mais simples e financeiramente mais acessível ao estabelecimento.

O Ezootaa [5] utiliza o rastreamento baseado em imagem, rastreando o cardápio e desenhando o conteúdo virtual sobre o mesmo. O usuário interage com o cardápio, visualizando os pratos e os ingredientes de cada prato, além do preço. Neste aplicativo, os pratos exibidos são imagens 2D. Diferentemente do Ezootaa, com ARFoodMenu o usuário poderá visualizar também a promoção ou evento de um certo período, visualizar pratos em 3D e informações do prato.

III. ARFOODMENU

O acrônimo ARFoodMenu significa “Cardápio de pratos de RA”. Desenvolvido para dispositivos móveis como *smartphones* e *tablets*, inicialmente o ARFoodMenu é compatível apenas com o sistema operacional Android. O intuito deste aplicativo é fornecer um cardápio virtual com interatividade para que o cliente do estabelecimento alimentício possa navegar, escolher pratos e realizar pedidos de

forma prática, sem a necessidade de chamar o garçom. Outra funcionalidade do aplicativo é a possibilidade de visualizar promoções “surpresas” do dia. Além das duas funcionalidades citadas nos parágrafos anteriores, o cliente também poderá visualizar o prato selecionado em 3D, além de informações adicionais, como ingredientes e informações nutricionais sobre o prato. As solicitações de pedidos serão recebidas pelo servidor do aplicativo administrado pelo gerente do estabelecimento como forma de auxiliar o trabalho dos funcionários e estreitar a comunicação entre as partes (funcionários e clientes), o que pode acarretar um melhor desempenho no trabalho por parte dos funcionários e aumentar os níveis de segurança na escolha do prato e a praticidade do cliente.

As funcionalidades de RA desenvolvidas foram promoção ou evento de um certo período, na qual é possível visualizar uma promoção ou evento oferecido pelo estabelecimento em períodos específicos; cardápio virtual, o qual fornece as opções do cardápio real dispostos por categoria, com fotos ilustrativas de todos os pratos; e visualização do prato em 3D, que habilita o cliente a visualizar qualquer prato do cardápio em seu tamanho real.

Para o desenvolvimento das funcionalidades de RA utilizou-se a linguagem de programação Java juntamente com o SDK da Metaio [6]. Como recursos para testes foram utilizados os *smartphones* modelos LG Optimus L5 E610 e Samsung Galaxy Grand Duos GT-I9082, com sistemas operacionais Android Jelly-Bean 4.1.2 e Android Kitkat 4.2, respectivamente, além do *tablet* Samsung Galaxy Note 2 SM-P601, com sistema operacional Android Kitkat 4.4.2. O aplicativo poderá ser baixado inicialmente pela Amazon Appstore.

A. Promoção ou Evento

O usuário do ARFoodMenu está habilitado a visualizar a promoção ou evento de um período específico – horário, dia, etc. – oferecido pelo bar e/ou restaurante. O usuário poderá rastrear o marcador baseado em ID (Figura 1) com a câmera do dispositivo móvel e uma imagem referente à promoção ou evento do período atual será exibida sobre o marcador. A função detecta a imagem do marcador usando a câmera do dispositivo. Posteriormente, o marcador é rastreado e sua pose é usada para que uma imagem seja desenhada sobre o mesmo (Figura 1).



Figura 1. Marcador baseado em ID usado para rastreamento da promoção ou evento no período predeterminado (esquerda) e anúncio virtual da promoção ou evento do dia exibido sobre o marcador rastreado (direita).

No canto inferior esquerdo da Figura 1 é exibido o botão “Pedir”. O usuário tem a opção de realizar o pedido da promoção diretamente no aplicativo. A identificação da mesa em a qual o cliente está sentado é feita através de marcadores baseados em ID que se encontram nas mesas.

B. Cardápio Virtual Interativo

No cardápio virtual interativo, mostrado na Figura 2, uma imagem é rastreada e são exibidas sobre ela imagens 2D dos pratos fornecidos pelo estabelecimento. Os pratos podem ser visualizados por categoria, como carnes, massas e etc. O título da categoria atual dos pratos pode ser visto no topo superior central da Figura 2.

O usuário pode interagir com o cardápio virtual selecionando a página virtual do cardápio clicando nos botões “Anterior” e “Próxima”, nos cantos inferiores esquerdo e direito, respectivamente, como mostra a Figura 2. Ao escolher o prato (Figura 3) desejado, o usuário poderá dar um clique único no prato escolhido e realizar interação por toques e gestos, como aumentar ou diminuir o tamanho da imagem do prato e/ou movimentá-la usando a função de arrastar. Ainda com o prato selecionado, o usuário tem a possibilidade de visualizar as informações nutricionais (Figura 4) clicando no botão com o nome do mesmo, assim como os ingredientes que compõem o prato (Figura 5).



Figura 2. Cardápio virtual interativo sendo exibido sobre a imagem rastreada.



Figura 3. Prato selecionado do cardápio virtual interativo.

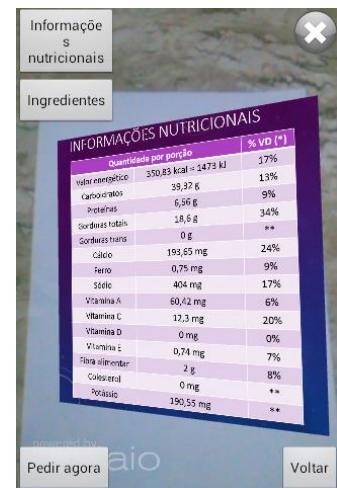


Figura 4. Informações nutricionais do prato selecionado no cardápio virtual.



Figura 5. Ingredientes que compõem o prato selecionado no cardápio virtual.

C. Visualização do Prato em Tamanho Real

O aplicativo também oferece ao usuário a opção de visualizar o prato escolhido em 3D, ou seja, um formato que se assemelhe à refeição no tamanho real inserido dentro de um prato físico na mesa.

A partir do rastreamento de um marcador posicionado próximo ao prato (Figura 6), no qual a posição e a escala relativa ao prato na mesa são previamente conhecidas, o conteúdo virtual é posicionado sobre o prato (Figura 6). Sendo assim, é possível visualizar o prato de vários ângulos, fornecendo ao cliente uma visão das dimensões do prato, semelhante ao prato real oferecido pelo estabelecimento. O aplicativo suporta modelos 3D nos formatos MD2, OBJ e FBX.



Figura 6. Conteúdo 3D (pizza) sendo exibido sobre o prato. O marcador rastreado encontra-se ao lado do prato.

IV. DISCUSSÃO

A principal finalidade do ARFoodMenu é facilitar o trabalho dos funcionários e fornecer praticidade no momento de escolher um prato ao cliente, com entretenimento e interatividade em alto nível providos pelas funcionalidades de RA.

O cliente que esteja interessado nas promoções ou eventos não precisa direcionar-se ao garçom ou gerente, que porventura podem estar indisponíveis no momento. Basta acionar o aplicativo, selecionar a função de promoções do dia, rastrear o marcador oferecido pelo estabelecimento e visualizar o conteúdo. Sem RA, o estabelecimento também pode divulgar promoções por impressos distribuídos dentro e fora da área do estabelecimento. Contudo, caso a administração decida encerrar ou adicionar novas promoções ou eventos, será necessário imprimir novas folhas ou panfletos com conteúdos novos.

O cardápio virtual interativo oferece vantagens ao cliente como a visualização de pratos por categoria, auxiliando-o na identificação rápida dos pratos. Também é possível visualizar imagens 2D dos pratos. Outra vantagem é a visualização de ingredientes e informações nutricionais de cada prato. Além de auxiliar o cliente, o cardápio virtual também facilita a inserção de conteúdos no aplicativo. Com o cardápio virtual, é preciso ter apenas uma nova imagem e as informações do prato em formato digital, sem ser necessária a utilização de outros materiais como papel, tinta de impressora, impressoras, mão de

obra etc. para cardápios físicos, melhorando o custo-benefício do estabelecimento. Por fim, com a visualização da refeição em 3D, o cliente terá a possibilidade de analisar as dimensões da mesma, como o tamanho (largura, comprimento e altura) e a disposição real dos ingredientes no prato em vários ângulos, podendo assim fazer uma comparação com o prato real e uma análise detalhada da refeição. Ter um *preview* do prato desejado obriga, ainda, o estabelecimento a fornecer aquilo que o prato em 3D oferece de conteúdo.

V. CONCLUSÃO E TRABALHOS FUTUROS

Tendo em vista a inovação no âmbito de bares e restaurantes, o ARFoodMenu incorporou algumas funções de RA. Foram vistas as vantagens que o aplicativo de RA pode trazer para os funcionários do estabelecimento alimentício e para o cliente, o qual é o público alvo desta aplicação.

De fato, incorporar RA nesse ambiente é uma ótima estratégia para atrair a atenção do público e auxiliar na economia do estabelecimento. Isso pode ser deduzido pelo fato de que alguns aplicativos introduzidos nessa área – como os mostrados na Seção II – já são bem aceitos. Não implica dizer que outras aplicações mais simples também não possam solucionar tais problemas, mas o intuito de ter RA no ARFoodMenu é o de que, além de resolver todos os problemas expostos nesse artigo, ele também possa entreter o cliente e mudar a forma de interação entre usuário e aplicativo.

A principal mudança prevista é a implementação do cardápio virtual interativo com a proposta de fazer um mapeamento de uma mesa do bar e/ou restaurante a partir de um marcador baseado em ID. O mapeamento seria feito usando rastreamento de imagem estendido [7] e com isso não seria necessário manter o marcador na mesa após o rastreamento do cardápio, funcionando como uma mesa interativa semelhante à do restaurante Inamo [4].

REFERÊNCIAS

- [1] Auggy. Augmented Reality Smart Menu. GT IEEE Innovation Team. Disponível em: <https://www.youtube.com/watch?v=PXoTSYxV6dYJ>. Acessado em 22 de Outubro de 2014.
- [2] Ogle. Reality Augmented Entertainment App. Disponível em: <http://www.youtube.com/watch?v=kK4Db2s28SQ>. Acessado em 21 de Outubro de 2014.
- [3] Watchar. Augmented Reality Platform. Chalie Pizza Restaurant, Lituânia. Disponível em: <https://www.youtube.com/watch?v=X88b9MpsSo>. Acessado em 22 de Outubro de 2014.
- [4] Inamo Interactive Table Menu. Restaurant Menu Table Projection. Disponível em: https://www.youtube.com/watch?v=ENFqP7A_0BI&list=PLNqg4C2iu2-onhpZYMPDEICuwJP2r8QsO. Acessado em 22 de Outubro de 2014.
- [5] Ezootaa Interactive. Interactive Augmented Menu. Disponível em: <http://www.youtube.com/watch?v=BAy9H5cEDxw>. Acessado em 21 de Outubro de 2014.
- [6] Metaio. The Augmented Reality Company. Documentation. Disponível em: <https://dev.metaio.com/junaio/documentation/>. Acessado em 05 de Janeiro de 2015.
- [7] SLAM and Extended Image Tracking. 3D Extended tracking. SDK. Documentation. Metaio Developer Portal. Disponível em: <http://dev.metaio.com/sdk/documentation/tracking-config/optical-tracking/slam-and-extended-image-tracking/>. Acessado em 14 de Março de 2015.

Utilização de ambientes virtuais imersivos para o ensino

Vagner Scamati

Programa de Mestrado

FACCAMP

Campo Limpo Paulista, SP, Brasil

vagnerscamati@hotmail.com

Diego Colombo Dias

Departamento de Ciência Computação

UFSCAR

São Carlos, SP, Brasil

diegocolombo.dias@gmail.com

Valéria Farinazzo Martins

Faculdade de Computação e Informática

Universidade Presbiteriana Mackenzie

São Paulo, SP, Brasil

valfarinazzo@hotmail.com

Marcelo de Paiva Guimarães

Universidade Federal de São Paulo (UNIFESP-UAB)/

Programa de Mestrado da Faccamp

São Paulo, SP, Brasil

marcelodepaiva@gmail.com

Resumo — Este artigo tem o objetivo de identificar, na literatura, ambientes de realidade virtual imersivos e interativos utilizados como apoio ao processo de ensino-aprendizagem. Para isso, é apresentada uma revisão sistemática da literatura de artigos publicados entre 2013 e 2015 na base de dados IEEE, ACM e Portal CAPES. O estudo resultou em cinco artigos, conforme os critérios de inclusão e exclusão aplicados. A análise dos artigos aponta que o uso de ambientes virtuais imersivos voltados ao ensino é uma tendência promissora, visto que os ambientes virtuais podem contribuir para a relação ensino e aprendizado.

Palavras-chave: realidade virtual; ensino, aprendizagem, educação.

I. INTRODUÇÃO

Nos últimos anos houve uma crescente demanda por ambientes virtuais para apoiar o processo de ensino-aprendizagem. Isso ocorreu devido à capacidade destes ambientes de gerarem mundos sintéticos, que permitem a imersão, navegação e interação por meio da exploração dos sentidos básicos do corpo.

Destaca-se nestes ambientes de realidade virtual, quando comparados com outras formas de interação homem-computador, a possibilidade de simular situações possíveis e impossíveis no mundo real. Por exemplo, assim como é possível transportar o usuário em um museu com obras de arte, também se pode colocá-lo dentro de um vulcão durante uma erupção, algo que não é possível em uma situação real. Essa capacidade de simular amplia as possibilidades do enriquecimento do processo de ensino com diversas situações. Além disso, tem também como vantagem a possibilidade de repetir a mesma simulação sem colocar em risco os alunos.

Contudo, o uso efetivo de ambiente de realidade virtual no contexto educacional é ainda limitado a apenas algumas instituições. Isso ocorre devido a vários motivos, dentre eles, a

necessidade de equipamentos, professores treinados, simuladores adequados com o conteúdo e de pessoal de apoio. A incorporação dos ambientes de realidade virtual no meio educacional somente acontecerá quando essas barreiras forem superadas. Além disso, é fundamental que estejam atrelados a um propósito educacional bem definido. Somente assim essa tecnologia poderá agregar valores ao processo de ensino-aprendizagem.

Existem na literatura exemplos de uso de ambientes de realidade virtual em diversas áreas do conhecimento e em variados níveis de instrução, como, por exemplo, em simulação de processos químicos e físicos, e na aprendizagem de pilotagem de aviões. Segundo Dede[1], esses ambientes apresentam indícios de contribuição na melhoria e aprimoramento do processo ensino-aprendizagem.

Este artigo apresenta o resultado de uma revisão da literatura sobre ambientes virtuais imersivos voltados ao ensino. Assim, são apresentados resultados que este tipo de ambiente tem mostrado em algumas áreas.

O restante do artigo está organizado da seguinte forma: a seção II apresenta o método utilizado para a construção da revisão; a seção III apresenta os artigos resultantes da revisão realizada; a seção IV mostra as discussões com base nos artigos selecionados, e, por fim, a seção V apresenta as conclusões.

II. MÉTODOS

A pesquisa abrangeu as bases de dados eletrônicas consideradas como sendo as fontes científicas mais relevantes. As bases de dados eletrônicas selecionadas foram: IEEE, ACM e o Banco de Teses da Capes (<http://servicos.capes.gov.br/capesdw/>). Os tipos de trabalhos considerados foram: teses, dissertações, monografias, artigos de revisão e artigos publicados em conferências ou periódicos. Os seguintes descritores foram utilizados para a busca dos

artigos: “virtual immersive environment”, “education” e “virtual learning environments”. As buscas nos sites de referências na Internet foram realizadas por meio de opções avançadas de cada motor de pesquisa, fazendo uso de filtros de busca, em um intervalo de publicação entre 2013 e 2015.

Após executadas as buscas com a combinação dos termos, os resumos dos artigos retornados pelos buscadores foram lidos e analisados, segundo os seguintes critérios de inclusão: ter sido publicado no período 2013 e 2015; estar em língua inglesa, e obedecer aos descritores. Os critérios de exclusão considerados foram: excluir artigo que não apresentasse a versão completa para leitura; excluir artigos que não relacionavam o conceito solicitado com o objetivo do tema da revisão.

III. RESULTADOS

As buscas retornaram um total de 15 artigos. Desse total, apenas cinco artigos foram selecionados conforme os critérios de inclusão e exclusão para a leitura completa [2]. Os artigos encontrados e selecionados para a leitura completa estão sistematizados na Tabela 1.

Em 2013, Fengru et al [3], apresentaram um estudo sobre a construção e visualização de ambientes geológicos complexos baseados no VGE (ambiente virtual geográfico). Esta solução baseou-se na ferramenta OpenSimulator, que é distribuída e colaborativa. Ela permite que alunos interajam com o ambiente virtual. Este estudo apontou que a prática de visitas a campos geológicos, como demonstrado na Figura 1, que os alunos podem interagir com a superfície e profundidade, podendo tornar o aprendizado nas dimensões geográficas mais eficiente e eficaz.



Fig. 1. Simulação com campos geológicos através do VGE [3].

Já no estudo de Wagner et al [4], os autores apresentam um ambiente virtual 3D combinado com outras tecnologias, como WEB 2.0, aos estudantes de um curso baseado em ensino a distância. A solução apresentada por eles tem como objetivo diminuir a distância transacional entre os professores, tutores e alunos, tornando-se então uma ferramenta de apoio a esta modalidade de curso. Este ambiente virtual 3D foi desenvolvido com o penSim. O estudo de caso foi

implementado em um programa técnico brasileiro, o PRONATEC.

Voss [5] apresentou uma proposta de criação de um laboratório virtual imersivo para o ensino de redes de computadores por meio de um cenário de laboratório base denominado ASTERIX. Esse ambiente proporciona aos alunos o conhecimento técnico em redes de computadores e até mesmo a capacidade de aplicar esse conhecimento em mundo real, bem como a noção de buscar a melhoria por meio da aprendizagem e do trabalho colaborativo. Foi utilizado o *OPENSIM 3D* no processo de simulação, focando nas características da realidade virtual. A aplicação permite que o aluno efetue configurações e testes sobre uma rede virtual de computadores por meio de uma interface gráfica do usuário (GUI). A Figura 2 apresenta a execução do ASTERIX.



Fig. 2. Laboratório Virtual de ensino de redes de computadores [5]

Os resultados do trabalho apresentado por Callaghan et al. [6] exploram como os mundos virtuais e técnicas de jogos de vídeo podem ser usadas para criar ambientes altamente imersivo e envolvente para o ensino de material relacionado à engenharia. Este processo descreve como os educadores podem aprender a usar efetivamente as características mais atraentes de jogos de computador, por exemplo: a participação ativa, intrínseca e o retorno imediato. Para os cursos de educação à distância relacionados com a utilização da engenharia é o único método realista de proporcionar o aprendizado prático, permitindo que alunos utilizem remotamente os laboratórios para práticas, sem restrições de tempo ou considerações geográficas.

Este artigo explorou como todos esses componentes individuais podem, com sucesso, ser integrados para criar um ambiente altamente inovador. Além disso, este trabalho demonstra como o ambiente existente é reforçado por meio da adição de uma camada de base de jogo, permitindo que equipes de estudantes trabalhem em conjunto de forma colaborativa e competitiva.

A Tabela 1 a seguir traz os artigos selecionados nesta pesquisa, assim como aponta os objetivos e resultados de cada um.

TABELA 1. ARTIGOS SELECIONADOS

Artigo/Veículo de Publicação	Base de Dados	Objetivo / Resultados
Fengru Huang, et al, 2013; Three-dimensional construction and visualization of complex geologic environments for virtual field practice and virtual education. / Geoinformatics (GEOINFORMATICS), 2013 21st International Conference on.	IEEE	O trabalho apresenta a construção 3D e visualização de ambientes geológicos complexos baseados em Virtual Geographic Ambiente (VGE). A abordagem fornece uma plataforma de interação flexibilidade e sistema de aprendizagem distribuída e colaborativa para a aprendizagem dos alunos e prática virtual
Wagner, R., et al, 2013; Using 3D virtual learning environments in new perspective of education./Information Technology Based Higher Education and Training (ITHET), 2013 International Conference on.	IEEE	O presente trabalho tem como objetivo o desenvolvimento de um ambiente de realidade virtual para os alunos do ensino à distância, a fim aprimorar a forma de integrar mais uma tecnologia de apoio a esta modalidade de cursos.
Voss, G.B., 2013; Context-Aware Virtual Laboratory for Teaching Computer Networks: A Proposal in the 3D OpenSim Environment. / Virtual and Augmented Reality (SVR), 2013 XV Symposium on.	ACM	Este trabalho apresentou uma proposta para a construção de um laboratório virtual envolvente de redes de computadores no ambiente 3D <i>OpenSim</i> , foi focado nas características da realidade virtual, sendo uma aplicação que permite a configuração e testes sobre uma rede virtual de computadores por meio de uma interface gráfica.
Callaghan, M.J., et al; 2013; Using Game-Based Learning in Virtual Worlds to Teach Electronic and Electrical Engineering/ Industrial Informatics,, IEEE Transactions on (Volume:9 , Issue: 1)	IEEE	Este artigo explora como os mundos virtuais e técnicas de jogos de vídeo podem ser usadas para criar ambientes altamente imersivo e envolvente para o ensino de material relacionado com a engenharia.
Banerjee, I., et al., 2013; Introducing immersive technologies for learning: A study evaluating the readiness for using virtual worlds for learning and teaching in an Indian University. / Contemporary Computing (IC3), 2013 Sixth International Conference on.	ACM	Este trabalho apresenta um estudo de caso que demonstra o envolvimento bem sucedido de um mundo virtual 3D online compartilhado como um auxílio educacional para redes de ensino e comunicação nas universidades indianas, promovendo benefícios na tarefa de aprendizagem. O estudo utilizou a ferramenta OpenSim, que é um servidor de hospedagem de mundos virtuais de código aberto para as simulações de ambiente virtual. Os autores deste artigo forneceram dois estudos de caso sobre o mesmo conteúdo educacional. O primeiro teste é um ambiente 2D baseado na web e o segundo um ambiente 3D. Eles concluíram que um ambiente virtual de aprendizagem tem resultados afirmativos e positivos com relação ao envolvimento dos alunos, podendo, inclusive, melhorar suas notas. A usabilidade da ferramenta OpenSim permite que os arquivos podem ser compartilhados e carregados em um laptop padrão e utilizado pelo aluno em qualquer lugar, proporcionando mais flexibilidade.

O trabalho proposto por Banerjee et al. [7] apresenta um estudo de caso que demonstra o envolvimento bem sucedido de um mundo virtual 3D online compartilhado como um auxílio educacional para redes de ensino e comunicação nas universidades indianas, promovendo benefícios na tarefa de aprendizagem. O estudo utilizou a ferramenta OpenSim, que é um servidor de hospedagem de mundos virtuais de código aberto para as simulações de ambiente virtual. Os autores deste artigo forneceram dois estudos de caso sobre o mesmo conteúdo educacional. O primeiro teste é um ambiente 2D baseado na web e o segundo um ambiente 3D. Eles concluíram que um ambiente virtual de aprendizagem tem resultados afirmativos e positivos com relação ao envolvimento dos alunos, podendo, inclusive, melhorar suas notas. A usabilidade da ferramenta OpenSim permite que os arquivos podem ser compartilhados e carregados em um laptop padrão e utilizado pelo aluno em qualquer lugar, proporcionando mais flexibilidade.

As informações apresentadas dentro do mundo virtual ajudam o aluno a explorar o tema a fundo, podendo usar esse conhecimento no cenário do mundo real. A educação imersiva com o OpenSim ajuda os usuários a entender assuntos complexos que não podem aprender no ambiente de sala de aula.

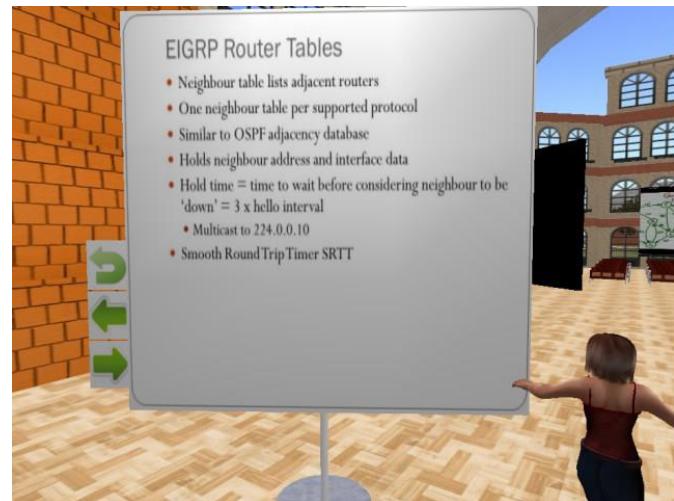


Fig. 3. Exibição do conteúdo usando slide show com botões interativos

IV. DISCUSSÃO

Os avanços tecnológicos experimentados nos últimos tempos, em diversos aspectos, acabaram modificando áreas da sociedade. A área educacional foi uma das que passou por processos e avanços significativos e que possibilitou o avanço de abordagens utilizando-se da tecnologia em seu favor. A criação e utilização de ambientes virtuais que fazem uso intenso ou não de ferramentas que interação por meio da imersão com o aluno ou docente através do aprendizado está consolidando um novo paradigma na relação ensino-aprendizado, incorporando novas experiências tecnológicas, educacionais e sociais.

A adoção da aprendizagem virtual, segundo Banerjee et al.[7], permite que os professores recriem um mundo real, ou que criem mundos completamente novos, provendo experiências que podem ajudar estudantes no entendimento de conceitos ou na execução de tarefas específicas do aprendizado.

Neste tipo de abordagem não existe separação entre ação, reflexão e questionamentos, visto que o aluno aprende a teoria por meio da apresentação de conceitos e definições, vivenciando a prática através das simulações próprias desse tipo de ambiente e realizando os questionamentos utilizando os recursos de comunicação das ferramentas. Por esta razão, a aprendizagem virtual e imersiva, conforme relata Wagner et al [4] é indicada para melhorar a relação ensino e aprendizagem, sobretudo no que diz respeito a atividades práticas, que permitam ao aluno aplicar o conhecimento adquirido, conforme visto nos relatos da literatura.

Contudo, pode-se dizer que aprendizagem virtual imersiva é uma técnica que utiliza conceitos computacionais, notadamente a realidade virtual passa a ser um conceito que decorre da interação, estando diretamente relacionada à questão da motivação do aluno em relação ao ambiente.

O uso de ambientes de realidade virtual imersivos na educação tende a promover esse novo paradigma educacional, onde os alunos não apenas vão receber informações, mas, também, interagir uns com os outros através da manipulação de conteúdo em múltiplas formas, manipulando o conhecimento que deve ser examinado, explorado e estudado. Uma característica fundamental do ambiente virtual é a possibilidade da interação social, fornecendo a capacidade de ações instantâneas e reações em tempo real.

V. CONCLUSÃO

Com base nos resultados encontrados na literatura, foram descritas as potencialidades de se aplicar os ambientes virtuais imersivos para o ensino com o objetivo de melhorar e contribuir com um modelo de educação construtivista através da relação ensino-aprendizagem. Esta abordagem tem se demonstrado eficaz e satisfatória, quando aplicada em um contexto educacional e independe de regras, culturas, países ou grau de escolaridade.

A utilização das propostas de ambientes virtuais para o ensino, envolvendo alunos e docentes na prospecção de soluções em ambientes virtuais de ensino, tem aumentado muito atualmente.

Por fim, os achados científicos apresentados neste artigo sinalizaram com muitos benefícios e com boas perspectivas para a melhoria e utilização deste novo paradigma, que pode ser aplicado em áreas diferentes da sociedade, principalmente a que vimos neste artigo, na educação e ensino. Como trabalho futuro, planeja-se expandir a revisão, ampliando os anos considerados e o uso de novos descritores

REFERÊNCIAS

- [1] Dede, Chris. Immersive Interfaces for Engagement and Learning. Science Magazine. V. 323, p. 66-69, jan/2009.
- [2] Biolchini, J., Mian, P.G., Natali, A.C.C., and Travassos, G.H. (2005). Systematic review in software engineering. Technical report, RT-ES679/05. System Engineering and Computer Science Dept., COOPE/UFRJ.
- [3] Fengru Huang, et. al, 2013. Three-dimensional construction and visualization of complex geologic environments for virtual field practice and virtual education./Geoinformatics (GEOINFORMATICS), 2013 21st International Conference on.
- [4] Wagner, R., et. al, 2013. Using 3D virtual learning environments in new perspective of education./Information Technology Based Higher Education and Training (ITHET), 2013 International Conference on
- [5] Voss, G. B., 2013; Context-Aware Virtual Laboratory for Teaching Computer Networks: A Proposal in the 3D OpenSim Environment. / Virtual and Augmented Reality (SVR), 2013 XV Symposium on.M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 2013..
- [6] Callaghan, M. J., et al; 2013; Using Game-Based Learning in Virtual Worlds to Teach Electronic and Electrical Engineering/Industrial Informatics, IEEE Transactions on (Volume:9 , Issue: 1)
- [7] Banerjee, I., et al., 2013; Introducing immersive technologies for learning: A study evaluating the readiness for using virtual worlds for learning and teaching in an Indian University. Contemporary Computing (IC3), 2013 Sixth International Conference on.

Ensino da Anatomia Humana Por Meio de Ambientes e Modelos Virtuais

Fabrício Quintanilha Baptista e
José Remo Ferreira Brega
Faculdade de Ciências – FC
UNESP – Bauru, SP – Brasil
Email: fabricioqb@gmail.com e
remo@fc.unesp.br

Mário Popolin Neto
Instituto Federal de São Paulo – IFSP
Registro, São Paulo, Brasil
Email: mariopopolin@ifsp.edu.br

Silke Anna Theresa Weber
Faculdade de Medicina de Botucatu – FMB
UNESP – Botucatu, SP – Brasil
Email: silke@fmb.unesp.br

Resumo—The study of anatomy and the human body structure is very important for medical education, although the environment in which we live and work is three dimensional, most forms of human anatomy teaching are offered to students through 2D environments such as books and slides presentation, or through synthetic models and cadaveric dissection techniques, they present a series of problems such as high cost and limited access to synthetic models, or ethical and religious issues in handling corpses. In recent years the teaching methods of human anatomy evolved also, with the emergence of many computer systems that were developed with specific purpose of assisting the teaching/learning of the human anatomy, these solutions feature the most varied techniques to achieve this goal, as 3D visualization, interaction with three-dimensional models and use of techniques augmented reality. These systems have shown great features to help the teaching and learning process of the human anatomy, but the ideal is that the tool is supported by a learning management system to provide real educational support teacher and the student.

I. INTRODUÇÃO

O ensino da anatomia humana tem sido submetido a progressivas alterações nos últimos anos em todo o mundo [1].

Atualmente os responsáveis pela educação médica enfrentam o desafio de melhorar a eficiência do ensino, compensando assim a diminuição de até 55% das horas utilizadas para o ensino da anatomia, tempo esse reduzido em muitos cursos ao redor do mundo [2].

De acordo com [3], os estudantes de hoje são conhecidos por serem nativos da Era Tecnológica e nesse caso uma maior atenção precisa ser desprendida às novas metodologias de ensino, visando tirar o máximo proveito da facilidade e da forma natural com que essa geração consegue se adaptar às tecnologias e aos modernos recursos oferecidos.

Existe ainda a necessidade de atrair a atenção de um público que utiliza dispositivos eletrônicos de comunicação como uma das primeiras fontes de informação. Sendo assim, as novas ferramentas de ensino apoiadas por computador e baseadas em tecnologia da informação são inseridas cada vez mais no ambiente de estudos médicos [3].

Contínuos avanços em tecnologias de computador e de interface permitiram a geração de modelos anatômicos detalhados que oferecem aos usuários além da visualização 3D, a possibilidade de manipular e interagir com esses modelos [4].

Um dos benefícios mais aparentes com a utilização de modelos 3D é a oportunidade de se observar as relações espaciais a partir de praticamente qualquer ângulo desejado, fornecendo assim informações que podem ser difíceis de adquirir com outros métodos de estudo. Além do que a facilidade de se encontrar computadores com acesso a internet nos dias de hoje faz com que esses ambientes virtuais estejam disponíveis praticamente a qualquer hora e lugar a um custo relativamente baixo, oferecendo assim incentivo e estímulo para aprendizagem em grupo ou individual por meio de um estudo adicional em relação às aulas tradicionais [5].

Em alguns casos a eficácia dos métodos de ensino tradicional pode ser prejudicada devido ao tamanho e complexidade das estruturas envolvidas, juntamente com uma incapacidade dos alunos devido à falta de experiência e a ausência de habilidades avançadas de dissecação [6].

Outras dificuldades encontradas nos métodos tradicionais de ensino são: baixa quantidade de peças sintéticas adquiridas pelas instituições devido ao seu custo elevado e problemas relacionados a questões éticas, religiosas e emocionais na manipulação de cadáveres [7].

Esse trabalho tem como objetivos principais elaborar um estudo sobre as formas de ensino da anatomia humana por meio da utilização de ambientes virtuais ou modelos virtuais e propor o desenvolvimento de uma solução para o ensino da anatomia humana por meio da integração de modelos virtuais tridimensionais com o LMS (Learning Management System) Moodle (Modular Object-Oriented Dynamic Learning Environment).

II. REVISÃO DE LITERATURA

Revisão sistemática de literatura, ou simplesmente revisão sistemática, é a identificação, avaliação e interpretação de material científico encontrado para uma questão de pesquisa, tópico ou fenômeno de interesse. A revisão sistemática em si é chamada de estudo secundário, e a produção científica utilizada na revisão sistemática é chamada de estudos primários [8].

Muitas razões fundamentam a realização de uma revisão sistemática, por exemplo, identificar tópicos de uma área de pesquisa que possam sofrer maiores investigações ou sumarizar evidências e aprofundar estudos para conhecer os benefícios e limitações de um método específico [8].

Uma estratégia deve ser seguida para a realização da revisão sistemática, envolvendo uma série de atividades como: especificar as questões de pesquisas, desenvolver um protocolo de revisão, selecionar e avaliar os estudos primários, extrair os dados e avaliar os resultados. Esse protocolo de revisão deve ser capaz de encontrar estudos primários que consigam responder as questões de pesquisa, o processo de extração de dados visa abstrair as informações para a resposta dessas questões e a análise de dados deve summarizar as informações e responder as questões de pesquisa [8].

A fim de realizar um levantamento bibliográfico quanto ao ensino da anatomia humana por meio de ambientes e modelos virtuais, foi executada uma revisão sistemática para responder a seguinte questão de pesquisa:

- Quais são as fundamentações teóricas e as aplicações práticas que viabilizam o ensino da anatomia humana por meio de ambientes virtuais de aprendizagem ?

Foram realizadas buscas por estudos primários em bases científicas, as bases escolhidas foram: IEEE Xplore, ACM Digital Library, Scopus e Science Direct. Quanto aos termos utilizados nas buscas, foram escolhidos: “virtual environment”, “virtual reality”, “human anatomy”, “human anatomy learning”, “virtual human anatomy”, “virtual human organ”, “human anatomy teaching”. A partir desses termos foi montada a seguinte lógica de busca:

- ((“virtual environment”) OR (“virtual reality”)) AND ((“human anatomy”) OR (“human anatomy learning”) OR (“virtual human anatomy”) OR (“virtual human organ”) OR (“human anatomy teaching”))

Essa busca foi realizada nas bases científicas selecionando o intervalo entre os anos de 2010 a 2015 e retornou as quantidades de artigos exibidas na Tabela I.

Tabela I. RESULTADO DAS BUSCAS.

Base Científica	Produção Científica
IEEE Xplore ¹	97
ACM Digital Library ²	33
Science Direct ³	32
Scopus ⁴	157
Total	319
Candidatos	62
Estudos Primários	19

Depois de analisar os candidatos, foram selecionados os artigos que seriam utilizados como estudos primários conforme mostra a Tabela II.

Após a seleção e análise dos estudos primários, notou-se que as soluções não seguiam um padrão em relação às tecnologias empregadas em seu desenvolvimento, utilizando uma grande variedade de linguagens de programação e de técnicas e ferramentas para o desenvolvimento dos modelos virtuais tridimensionais, resultando em soluções que podem ser executadas sobre as mais variadas plataformas, foi então

Tabela II. ESTUDOS PRIMÁRIOS.

Autor	Referência	Tipo de Solução
(Huang, 2011)	[9]	Web
(Stefan et al., 2014)	[10]	Outros
(Ke et al., 2013)	[11]	Outros
(Huang et al., 2010)	[12]	Mobile
(Meng et al., 2013)	[13]	Outros
(Bochicchio & Longo, 2012)	[14]	Desktop
(Melo et al., 2011)	[15]	Web
(Birr et al., 2013)	[16]	Web
(Azkue, 2013)	[17]	Desktop
(Lu et al., 2010)	[18]	Web
(Allen et al., 2015)	[19]	Desktop
(Nogueira et al., 2013)	[20]	Mobile
(Ruisoto Palomera et al., 2014)	[21]	Desktop
(Sander & Golas, 2013)	[22]	Web
(Tworek et al., 2013)	[23]	Desktop
(Codd & Choudhury, 2011)	[24]	Desktop
(Sandrine de Ribaupierre, 2012)	[25]	Desktop
(Adam et al., 2013)	[26]	Desktop
(Cardia da Cruz et al., 2014)	[27]	Web

realizado um agrupamento dos trabalhos referente a sua proposta principal e a plataforma em que essas soluções foram implementadas:

- Soluções web: Sistemas desenvolvidos em alguma plataforma web que dispensa qualquer tipo de instalação extra, é executado no browser do usuário;
- Soluções desktop: Sistemas desenvolvidos para que sua execução ocorra na máquina do usuário, necessitando de alguma instalação prévia para o seu funcionamento;
- Soluções mobile: Sistemas desenvolvidos para que sua execução ocorra em algum dispositivo móvel necessitando de instalação prévia para seu funcionamento; e
- Outras soluções: Foram classificados como outras soluções, os sistemas que necessitam de algum dispositivo ou equipamento extra para seu funcionamento ou visualização, por exemplo, equipamentos para projeção e calibragem em um sistema de realidade aumentada.

A Figura 1 ilustra a distribuição dos estudos primários em relação ao seu tipo de desenvolvimento.

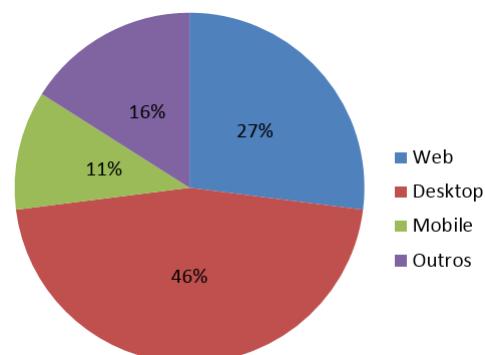


Figura 1. Estudos primários por tipo de desenvolvimento.

¹<http://ieeexplore.ieee.org/Xplore/home.jsp>

²<http://dl.acm.org/dl.cfm>

³<http://www.sciencedirect.com/>

⁴<http://www.scopus.com/>

Deve-se destacar que apesar dessa diversidade de tecnologias, ferramentas e plataformas, dentre os trabalhos que fizeram parte dos estudos primários não foi encontrada nenhuma solução que foi desenvolvida apoiada por um LMS, diferenciando assim a proposta apresentada nesse artigo de todas as outras soluções encontradas.

III. PROPOSTA

Os autores desse artigo propõem o desenvolvimento de uma solução para o ensino da anatomia humana por meio da integração de modelos virtuais tridimensionais com o LMS Moodle.

A solução será desenvolvida integrando os modelos virtuais tridimensionais do corpo humano diretamente na plataforma Moodle, esses modelos poderão ser acessados, manipulados e algumas atividades poderão ser apresentadas de acordo com o modelo selecionado. A integração será feita utilizando o padrão SCORM e a biblioteca javascript Three, pois dessa forma conteúdos padronizados podem ser criados, reutilizados e destruídos facilmente, isso exige do usuário apenas a utilização de navegadores web que suportam o WebGL [28]. A Figura 2 ilustra a arquitetura básica da solução proposta.

O Moodle foi escolhido pelo fato de ser amplamente utilizado pela comunidade acadêmica e por apresentar uma série de vantagens e funcionalidades próprias de um LMS, esses fatores garantem a acessibilidade necessária e o correto apoio pedagógico aos professores e aos alunos.

IV. CONSIDERAÇÕES FINAIS

Por meio dos estudos realizados, pode-se notar que assim como acontece em muitas outras atividades, a evolução da tecnologia tem alterado a forma de ensino e inserido novos paradigmas em suas metodologias. Devido às mudanças sociais, comportamentais e a evolução dos alunos nativos da era tecnológica os métodos tradicionais de ensino sofreram diversas mudanças e uma das áreas que mais tem sentido os efeitos dessas mudanças é atraído a atenção de pesquisadores é a área de estudo das ciências médicas.

Muitos fatores motivam o desenvolvimento de aplicações que por meio da utilização de ambientes virtuais e modelos virtuais tridimensionais tem como principal objetivo proporcionar o ensino da anatomia humana, diversas pesquisas estão sendo desenvolvidas de modo a oferecer recursos e ferramentas para suplementar ou até mesmo para substituir totalmente ou em parte os métodos tradicionais de ensino da anatomia humana.

Por fim pode-se notar que de acordo com os resultados obtidos nos trabalhos selecionados como estudos primários, o uso de tecnologia para o ensino da anatomia humana oferece muitos benefícios aos alunos por meio da visualização de modelos 3D e da interação com esses modelos, fazendo com que o processo de aprendizagem se torne mais contagiente, fornecendo assim uma melhora no processo de aquisição do conhecimento e aumentando o interesse do aluno pelo estudo por meio da utilização desses ambientes interativos, mostrando uma grande aceitação e intenção de uso para soluções desse tipo.

AGRADECIMENTOS

Os autores deste artigo gostariam de expressar seus agradecimentos ao Programa de Pós-Graduação em Ciência da Computação da UNESP.

REFERÊNCIAS

- [1] A. Bleakley, "The curriculum is dead! long live the curriculum! designing an undergraduate medicine and surgery curriculum for the future," *Medical Teacher*, vol. 34, no. 7, pp. 543–547, 2012. [Online]. Available: <http://www.tandfonline.com/doi/abs/10.3109/0142159X.2012.678424>
- [2] R. L. Drake, J. M. McBride, N. Lachman, and W. Pawlina, "Medical education in the anatomical sciences: The winds of change continue to blow," *Anatomical Sciences Education*, vol. 2, no. 6, pp. 253–259, 2009. [Online]. Available: <http://dx.doi.org/10.1002/ase.117>
- [3] A. Platt, "Teaching medicine to millennials," *The Journal of Physician Assistant Education*, vol. 21, no. 2, pp. 42–4, 2010.
- [4] R. B. Trelease, "Anatomical informatics: Millennial perspectives on a newer frontier," *The Anatomical Record*, vol. 269, no. 5, pp. 224–235, 2002. [Online]. Available: <http://dx.doi.org/10.1002/ar.10177>
- [5] J. A. Juanes and P. Ruisoto, "Technological advances and teaching innovation applied to health science education," in *Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality*, ser. TEEM '13. New York, NY, USA: ACM, 2013, pp. 3–7. [Online]. Available: <http://doi.acm.org/10.1145/2536536.2536537>
- [6] C. M. Adams and T. D. Wilson, "Virtual cerebral ventricular system: An mr-based three-dimensional computer model," *Anatomical Sciences Education*, vol. 4, no. 6, pp. 340–347, 2011. [Online]. Available: <http://dx.doi.org/10.1002/ase.256>
- [7] M. A. de Azambuja Montes and C. T. V. de Souza, "Análise da taxa de reprovação na disciplina de anatomia humana em cursos da saúde," 2007.
- [8] B. Kitchenham, O. Pearl Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering - a systematic literature review," *Inf. Softw. Technol.*, vol. 51, no. 1, pp. 7–15, Jan. 2009. [Online]. Available: <http://dx.doi.org/10.1016/j.infsof.2008.09.009>
- [9] H.-M. Huang, "A collaborative virtual learning system for medical education," in *Data Mining and Intelligent Information Technology Applications (ICMIa), 2011 3rd International Conference on*, Oct 2011, pp. 127–130.
- [10] P. Stefan, P. Wucherer, Y. Oyamada, M. Ma, A. Schoch, M. Kanegae, N. Shimizu, T. Kodera, S. Cahier, M. Weigl, M. Sugimoto, P. Fallavollita, H. Saito, and N. Navab, "An ar edutainment system supporting bone anatomy learning," in *Virtual Reality (VR), 2014 iEEE*, March 2014, pp. 113–114.
- [11] T. Ke, P. Xinhua, and G. Yuan, "Interactive experience design of medical simulation training," in *Service Science and Innovation (ICSSI), 2013 Fifth International Conference on*, May 2013, pp. 42–45.
- [12] H.-M. Huang, Y.-L. Chen, and K.-Y. Chen, "Investigation of three-dimensional human anatomy applied in mobile learning," in *Computer Symposium (ICS), 2010 International*, Dec 2010, pp. 358–363.
- [13] M. Meng, P. Fallavollita, T. Blum, U. Eck, C. Sandor, S. Weidert, J. Waschke, and N. Navab, "Kinect for interactive ar anatomy learning," in *Mixed and Augmented Reality (ISMAR), 2013 IEEE International Symposium on*, Oct 2013, pp. 277–278.
- [14] M. Bochicchio and A. Longo, "Learning objects and online labs: The micronet experience," in *Remote Engineering and Virtual Instrumentation (REV), 2012 9th International Conference on*, July 2012, pp. 1–7.
- [15] J. S. S. Melo, L. M. Brasil, R. Balanuki, E. Ferneda, and J. S. Santana, "Medical simulation platform," 2011.
- [16] S. Birr, J. Monch, D. Sommerfeld, U. Preim, and B. Preim, "The liveranatomyexplorer: A webgl-based surgical teaching tool," *Computer Graphics and Applications, IEEE*, vol. 33, no. 5, pp. 48–58, Sept 2013.
- [17] J.-J. Azkue, "A digital tool for three-dimensional visualization and annotation in anatomy and embryology learning," *Eur. J. Anat*, vol. 17, no. 3, pp. 146–154, 2013.

- [18] J. Lu, L. Li, and G. Sun, "A multimodal virtual anatomy e-learning tool for medical education," in *Entertainment for Education. Digital Techniques and Systems*, ser. Lecture Notes in Computer Science, X. Zhang, S. Zhong, Z. Pan, K. Wong, and R. Yun, Eds. Springer Berlin Heidelberg, 2010, vol. 6249, pp. 278–287. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-14533-9_28
- [19] L. K. Allen, S. Bhattacharyya, and T. D. Wilson, "Development of an interactive anatomical three-dimensional eye model," *Anatomical Sciences Education*, vol. 8, no. 3, pp. 275–282, 2015. [Online]. Available: <http://dx.doi.org/10.1002/ase.1487>
- [20] J. M. Noguera, J. J. Jiménez, and M. C. Osuna-Pérez, "Development and evaluation of a 3d mobile application for learning manual therapy in the physiotherapy laboratory," *Computers & Education*, vol. 69, pp. 96 – 108, 2013. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0360131513001747>
- [21] P. R. Palomera, J. A. J. Méndez, and A. P. Galino, "Enhancing neuroanatomy education using computer-based instructional material," *Computers in Human Behavior*, vol. 31, pp. 446 – 452, 2014. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0747563213000940>
- [22] B. Sander and M. M. Golas, "Histoviewer: An interactive e-learning platform facilitating group and peer group learning," *Anatomical Sciences Education*, vol. 6, no. 3, pp. 182–190, 2013. [Online]. Available: <http://dx.doi.org/10.1002/ase.1336>
- [23] J. K. Tworek, H. A. Jamniczky, C. Jacob, B. Hallgrímsson, and B. Wright, "The lindsay virtual human project: An immersive approach to anatomy and physiology," *Anatomical Sciences Education*, vol. 6, no. 1, pp. 19–28, 2013. [Online]. Available: <http://dx.doi.org/10.1002/ase.1301>
- [24] A. M. Codd and B. Choudhury, "Virtual reality anatomy: Is it comparable with traditional methods in the teaching of human forearm musculoskeletal anatomy?" *Anatomical Sciences Education*, vol. 4, no. 3, pp. 119–125, 2011. [Online]. Available: <http://dx.doi.org/10.1002/ase.214>
- [25] S. de Ribaupierre and T. D. Wilson, "Construction of a 3-d anatomical model for teaching temporal lobectomy," *Computers in Biology and Medicine*, vol. 42, no. 6, pp. 692 – 696, 2012. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0010482512000662>
- [26] A. Hamrol, F. Górski, D. Grajewski, and P. Zawadzki, "Virtual 3d atlas of a human body – development of an educational medical software application," *Procedia Computer Science*, vol. 25, pp. 302 – 314, 2013, 2013 International Conference on Virtual and Augmented Reality in Education. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1877050913012416>
- [27] L. Cardia da Cruz, V. De Almeida Thomaz, and J. de Oliveira, "Aicoh 3d: Interactive atlas of human body," in *Virtual and Augmented Reality (SVR), 2014 XVI Symposium on*, May 2014, pp. 24–27.
- [28] M. Popolin Neto, I. A. B. Sossai, F. Baptista, D. Santos, N. N. Braga, S. Weber, and J. R. F. Brega, "Tecnologias na integração de ambientes virtuais tridimensionais e a plataforma de ensino e aprendizagem moodle," *Revista do Simpósio Interdisciplinar de Tecnologias na Educação*, vol. 1, pp. 148–156, 2015. [Online]. Available: <http://sinte.btv.ifsp.edu.br/index.php/SInTE/article/view/0001-0020>

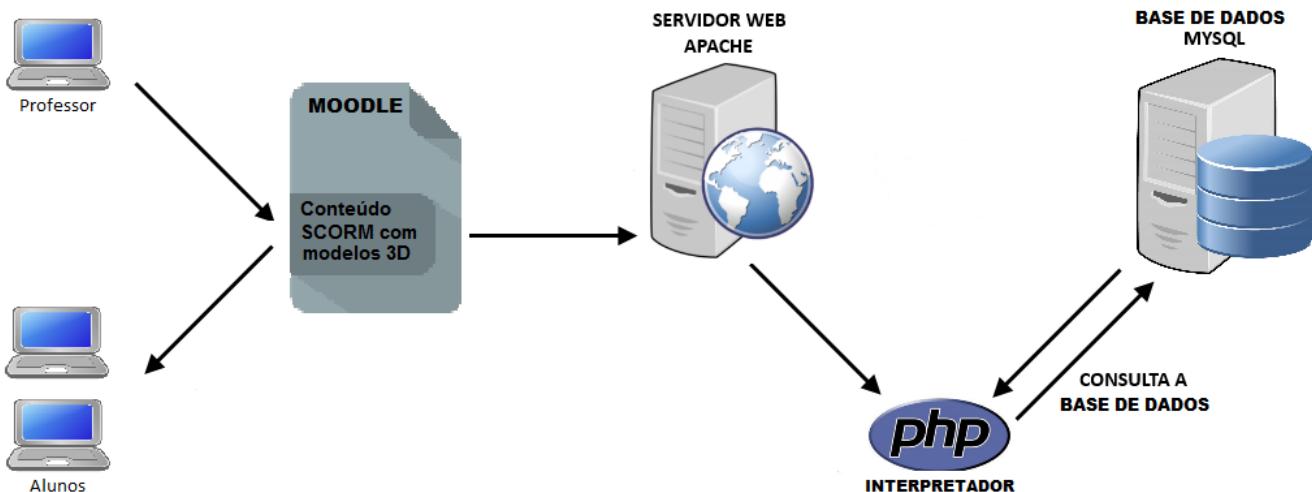


Figura 2. Arquitetura básica da solução.

Ferramenta de empacotamento de conteúdos de realidade aumentada no formato de objetos de aprendizagem

Bruno Alves de Carvalho

Ciência da Computação

Faculdade Campo Limpo Paulista (FACCAMP))

Campo Limpo Paulista, Brasil

tbrunoalves@gmail.com

Valeria Farinazzo Martins

Faculdade de Computação e Informática

Universidade Presbiteriana Mackenzie

São Paulo, Brasil

valfarinazzo@hotmail.com

Marcelo de Paiva Guimarães

Universidade Federal de São Paulo (UAB/UNIFESP)/

Programa de Graduação da Faculdade Campo Limpo

Paulista (FACCAMP), São Paulo, Brasil

marcelodepaiva@gmail.com

Resumo— *Objetos de Aprendizagem são conteúdos educacionais voltados para facilitar o processo de ensino-aprendizagem. Quando empacotados, podem ser compartilhados, por exemplo, em repositórios na internet. Por outro lado, a Realidade Aumentada é uma ferramenta que possui recursos que podem facilitar o processo educacional. Então, torna-se natural o empacotamento de conteúdos educacionais construídos com a Realidade Aumentada no formato de Objetos de Aprendizagem. Este artigo tem como objetivo principal a apresentação de uma ferramenta voltada para esse empacotamento no padrão SCORM. Além disso, apresenta um exemplo de empacotamento.*

Palavras Chave— Realidade Aumentada, SCORM, Empacotamento, Objeto de aprendizagem.

I. INTRODUÇÃO

Existe uma constante busca por soluções tecnológicas para o auxílio do processo de ensino-aprendizagem. Por exemplo, existem propostas que envolvem o uso de dispositivos móveis, lousas digitais e ambientes virtuais de simulação. Porém, somente a adoção de tecnologias não é suficiente para garantir o sucesso do processo de ensino-aprendizagem. Então, torna-se necessário que o uso delas esteja incorporado a estratégias educacionais apropriadas.

A Realidade Aumentada (RA), devido a sua capacidade de adicionar objetos virtuais (por exemplo, imagens e sons 3D) em um ambiente real, é uma tecnologia que vem sendo explorada em diversos contextos educacionais. Ela possui potencial de apoiar no processo cognitivo do aprendiz, propiciando não apenas a parte teórica do ensino, mas também, principalmente a experimentação prática do conteúdo abordado [1]. Contudo, a maioria dos trabalhos que envolvem a RA são desenvolvidos para serem utilizados especificamente por um único professor. Como consequência, a reutilização de conteúdos torna-se limitada.

A ideia de reuso do mesmo conteúdo por diversos professores está alinhada aos Recursos Educacionais

Abertos (REA) [2], que são materiais de ensino, aprendizado e pesquisa em qualquer suporte ou mídia, que podem ser utilizados ou adaptados por terceiros. Este conteúdo deve estar sob domínio público ou estar licenciado de forma permissiva. A possibilidade de uso de qualquer maneira e adaptação é fundamental para preservar a autonomia do professor – a autoria do material deve ser preservada.

Uma forma de criar conteúdo do tipo REA é empacotá-los como Objetos de Aprendizagem (OA), que, segundo Wiley [3], são recursos digitais ou não-digitais que contribuem para o processo de ensino-aprendizagem. O ideal é que eles sejam armazenados em repositórios na internet, o que proporciona então o acesso fácil por professores e alunos. Os OAs seguem padrões, como, por exemplo, o IEEE-LTSC (Learning Technology Standards Committee [4]), o IMS (Global Learning Consortium), ARIADNE (Alliance of Remote Instructional Authoring and Distribution Network for Europe) e o SCORM (Sharable Content Object Reference Model). Além da reusabilidade, fazem parte das características dos OAs a acessibilidade e a interoperabilidade. Atualmente, o Moodle [5] é a ferramenta de destaque no suporte a OA. Ela é capaz de importar esses objetos para o uso nos cursos.

Este trabalho tem como objetivo apresentar uma ferramenta capaz de empacotar conteúdo educacional baseado em RA no formato de OA. Por isso, no contexto deste trabalho os OAs podem ser definidos como materiais baseados em realidade aumentada voltados para o ensino. O padrão de OA escolhido foi o SCORM. Além da ferramenta, é apresentado um estudo de caso de empacotamento.

O restante do artigo está organizado da seguinte forma: a seção II apresenta os OAs e o padrão SCORM; a seção III apresenta a ferramenta de empacotamento; a seção IV mostra um exemplo de empacotamento, no caso foi criado um OA a partir de um conteúdo educacional

desenvolvido com a ferramenta Flaras; e, por fim, são mostradas as conclusões na seção V.

II. OBJETOS DE APRENDIZAGEM E O PADRÃO SCORM

Entre os diferentes conceitos sobre Objetos de Aprendizagem (OA), Wiley[2] os define como qualquer recurso digital que seja reutilizável para assistir à aprendizagem e que sejam distribuídos pela rede. Outro conceito é que eles são granulares, reutilizáveis, independentes de mídia [7].

Além do conteúdo didático, faz parte também dos objetos os seus metadados, que são dados sobre os objetos. Assim, eles contêm informações a respeito do conteúdo dos objetos e sobre a sua utilização. Por isso, os metadados são essenciais para promover a reutilização dos objetos. Os repositórios utilizam os metadados para catalogar e recuperar os OAs [8, 9].

Para que a reutilização dos OAs seja uma realidade é necessário que ferramentas, como os repositórios e os ambientes virtuais de aprendizagem, sejam compatíveis com eles. Para isso acontecer, então é fundamental a adoção de padrões, como o SCORM [10]. Além de permitir que os conteúdos educacionais sejam empacotados no formato de OA, esse padrão sugere quais serviços são essenciais para disponibilizar os objetos on-line via Web.

Assim, a adoção de padrões como o SCORM torna-se fundamental para que os conteúdos educacionais sejam empacotados no formato de OA e, consequentemente, reaproveitados em diversos contextos [11].

III. PROCESSO DE EMPACOTAMENTO

A Figura 1 mostra os passos necessários para que o conteúdo de RA gerado e empacotado no formato SCORM seja utilizado pelo estudante. Inicialmente, é criado o conteúdo de RA em uma ferramenta de autoria, como no Flaras (passo 1). Além do conteúdo relacionado diretamente com a aplicação de RA, pode também fazer parte outros dados, como arquivos HTML, imagens e animações. Logo após, utiliza-se a ferramenta de empacotamento apresentada neste trabalho (passo 2). Esta ferramenta cria os arquivos que o SCORM exige, e os adiciona aos outros arquivos gerados na ferramenta de autoria. O *imsmanifest.xml* é um exemplo de arquivo que é adicionado, nele é descrita a estrutura lógica do conteúdo educacional, ou seja, faz parte dele os *links* para todo o conteúdo do objeto. Em seguida, compacta-se todos os arquivos. Por fim, um ambiente de aprendizagem, como o Moodle, importa o OA (passo 3) e o disponibiliza como conteúdo ao aluno (passo 4). A partir desse momento, o ambiente de aprendizagem torna-se responsável pelo gerenciamento do objeto, ou seja, ele cuida de todas ações que ocorrem, por exemplo, redirecionando o usuário para um outro objeto e registrado todas as ações que ocorrem no ambiente.

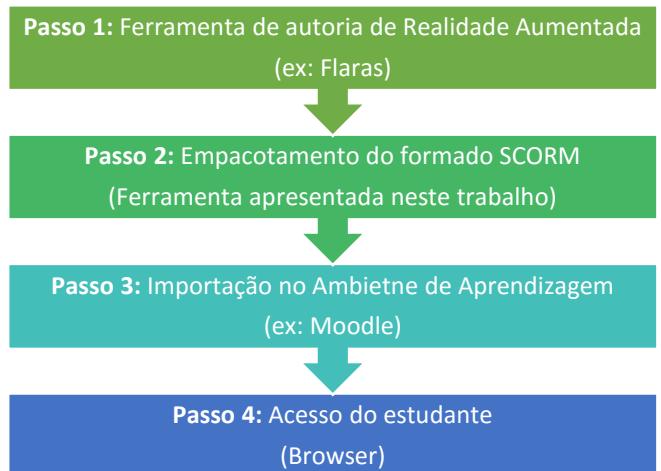


Figura 1: Geração, empacotamento e visualização do conteúdo de RA

O produto final gerado pela ferramenta descrita neste trabalho é um arquivo compactado no formato “pkzip” ou “pif”. O processo de empacotamento realizado não altera o conteúdo do material de RA criado pela ferramenta de autoria – arquivos HTML e flash. A Figura 2 apresenta a interface da ferramenta de empacotamento. Nela o usuário informa a localização (ícone de lupa) dos arquivos da aplicação de RA e o nome destino do OA. Quando acionado o botão de geração (ícone de gravação), o OA no padrão SCORM é gerado.

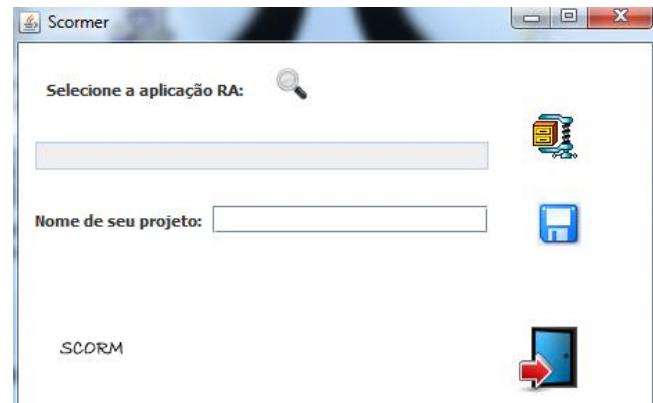


Figura 2: Interface da ferramenta de empacotamento

Essa ferramenta foi desenvolvida na linguagem Java com a ferramenta NetBeans 8.0.2. O diagrama de classes apresentado na Figura 3 mostra as principais classes do empacotador. A classe *Principal* é a que inicializa o aplicativo e constrói a interface gráfica. Ela utiliza os métodos da classe *gerenciaArquivos* para tratar a adição dos arquivos, a configuração dos valores internos dos arquivos, como, por exemplo, nomes dos *links* utilizados no arquivo *imsmanifest.xml*. A classe *Principal* também usa a classe *Compactador* para gerar e validar o OA.

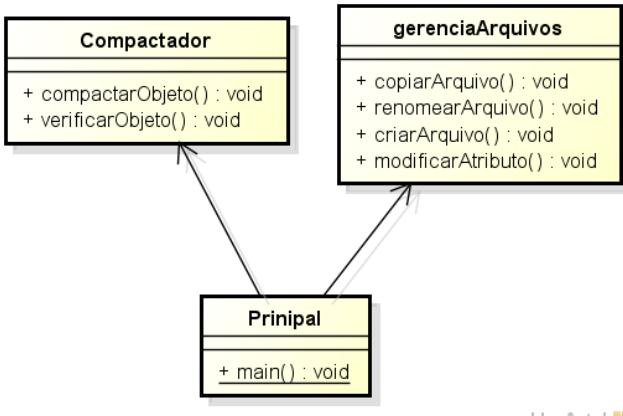


Figura 3: Diagrama de classes

IV. ESTUDO DE CASO

Com o objetivo de validar a ferramenta de empacotamento, criou-se um conteúdo educacional que envolve RA. A Figura 4 mostra uma das páginas desse conteúdo. O objetivo dele é ser utilizado como material de apoio no processo de ensino-aprendizagem de formas geométricas. Então, ele é composto por um menu (lado esquerdo) com diversas atividades de ensino. Os alunos utilizam um marcador para manipular as formas geométricas em 3D dos objetos. As páginas foram construídas utilizando HTML 5. O conteúdo de RA que foi embutido nas páginas foi desenvolvido com a ferramenta de autoria Flaras. Por isso, nesse caso o conteúdo de RA é em Flash.



Figura 4: Conteúdo educacional que utiliza RA

Além de visualizar os objetos 3D, os alunos contam com jogos no qual eles devem descobrir qual é a forma geométrica apresentada. A Figura 5 apresenta um exemplo. Nela são mostrados três objetos 3D e o aluno deve selecionar com o mouse qual é o que representa um cubo.

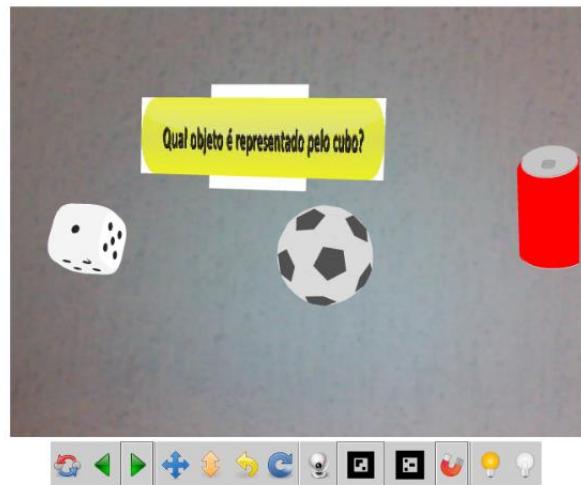


Figura 5: Interface do Objeto de Aprendizagem de Ensino de Geometria

Após a criação do conteúdo, o mesmo foi submetido ao processo de empacotamento na ferramenta desenvolvida nesse trabalho. A validação e execução foi realizada no Scorm cloud (<http://scorm.com/scorm-solved/scorm-cloud-features/>), que é um ambiente de aprendizagem on-line. Ele é voltado para o armazenamento e distribuição de conteúdo e-learning. Nele é possível carregar os objetos SCORM e apresentá-los. Para isso, foi necessário apenas criar uma conta nessa ferramenta, importar o OA e executá-lo. A página resultante foi a mesma apresentada na Figura 4. Não foi percebido nenhum problema de desempenho, pois o conteúdo foi executado diretamente no browser, como é realizado originalmente. Também não foi detectado nenhum problema de incompatibilidade. Os testes foram realizados no sistema operacional Windows 7, Windows 8 e Windows 10 nas versões 32 e 64 bits.

V. CONCLUSÃO

A busca pela inserção de novas tecnologias que possam auxiliar o processo de ensino-aprendizagem é constante. Atualmente, existem diversas ações de uso de RA como ferramenta auxiliadora de ensino em diversas situações, como, na ilustração de formas geométricas. Porém, para que a tecnologia esteja ao alcance de todos e que seja possível o reuso entre os professores, é necessário que a adoção de padrões para a criação dos mesmos, como, por exemplo, o SCORM.

Este trabalho apresentou uma ferramenta capaz de empacotar conteúdos didáticos baseados em RA no formato SCORM. Com isso, torna-se possível devolver materiais diversos e disponibilizados em repositórios, e, consequentemente, promover o reuso dos mesmos. Para os testes, realizou-se o desenvolvimento do conteúdo e validação do mesmo em um ambiente de ensino.

Como trabalhos futuros sugere-se o empacotamento de novos conteúdos e a adição na ferramenta de empacotamento de conteúdos de RA desenvolvidos com outras ferramentas de autoria, além do Flaras.

VI. REFERÊNCIAS

- [1] CARDOSO, A. & LAMOUNIER, E. Aplicações na Educação e Treinamento. In Realidade Virtual e Aumentada-Uma Abordagem Tecnológica, capítulo 9, páginas 343–357. Realidade Virtual e Aumentada Uma abordagem tecnológica, SBC, 2008.
- [2] WILEY, D. A. Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy. In D. A. Wiley (Ed.), The Instructional Use of Learning Objects. [Online]. Disponível em: <http://reusability.org/read/chapters/wiley.doc>. [Acessado em 18 de Agosto de 2015].2000.
- [3] ADL - Advanced Distributed Learning. Sharable Content Object Reference Model (SCORM) 2004 3rd Edition Documentation Suite. Disponível em: [Acesso em: 11 de set. de 2015].
- [5] Site Oficial do Moodle: Moodle www.moodle.org.[Acesso em: 20 de set. 2015].
- [6] MUZIO, J.; HEINS, T.; MUNDELL, R. Experiences with Reusable eLearning Objects: From Theory to Practice. Victoria, Canadá, 2001.
- [7] MACHADO, L. L. e Silva, J. T. Objeto de aprendizagem digital para auxiliar o processo de ensino-aprendizagem no ensino técnico em informática. Centro Interdisciplinar de Novas Tecnologias na Educação, Universidade Federal do Rio Grande do Sul, Porto Alegre. 2005.
- [8] SILVA, Maria da Graça Moreira da. Novas Aprendizagens, 2004. Disponível em:<http://www.abed.org.br/congresso2004/por/htm/146-TC-D2.htm>. Acesso em: 01 de set. 2015.
- [9] DZIEKANIAK, G. V.; KIRINUS J. B. Web Semântica. Revista Eletrônica de Biblioteconomia e Ciência da Informação. 2004. Volume 9 Issue:18 pages/rec.No:20-39. Disponível em:<<http://www.doaj.org/doaj?func=abstract&id=190138>>. [Acessado em: 09 dez. 2014].
- [10] SANTOS, M. E. K. L. Objetos e Ambientes virtuais de aprendizagem no ensino de matemática: um estudo de caso para o estágio supervisionado de docência. 2007. 103p. Dissertação de mestrado – Universidade Cruzeiro do Sul, São Paulo, 2007.
- [11] ARAÚJO, Moysés. Educação a Distância e a Web Semântica: Modelagem Ontológica de Materiais e Objetos de Aprendizagem para a Plataforma COL. São Paulo, 2003.
- [12] KIRNER, C. Disponível em: <http://ckirner.com/flaras2/introducao/Primeira>. [Acessado em: 03 de set. de 2015].

Estudo de potencialidades em ambientes de treinamento em 2D e em Realidade Virtual para o mapeamento de processos cognitivos na aprendizagem de subestações de energia

Aline F. Furtado Silva, Alexandre Cardoso, Elise B. Mendes e Edgard Lamounier

ABSTRACT

This article describes the proposal of a mapping prototype conceptual knowledge and cognitive processes for power substations learning, in order to evaluate the potential of 2D and virtual reality environments (RV) in the training of Energy dealership operators. This is a qualitative approach, which outlines an application example. First, it is a mapping of concepts and understanding of processes necessary for the substation operator function, then draws up a checklist for evaluating the potential of 2D environments and RV. The main objective of this work is to create a knowledge system prototype to identify the contributions of 2D objects and RV for conceptual learning electrical substations, their layouts, and active components by developing checklists and metrics.

RESUMO

Este artigo descreve a proposta de um protótipo de mapeamento de conhecimentos conceituais e processos cognitivos para a aprendizagem de subestações de energia, com o objetivo de avaliar as potencialidades dos ambientes de 2D e realidade virtual (RV) no treinamento de operadores de Subestações de Energia. Trata-se de uma abordagem qualitativa que delineia um plano de aplicação. Primeiramente, é feito um mapeamento de conceitos e processos de entendimento necessários à função de operador de subestações, em seguida é elaborado um *checklist* para a avaliação das potencialidades dos ambientes em 2D e RV. O principal objetivo deste trabalho é criar um protótipo de sistema de conhecimento para identificação das contribuições dos objetos 2D e RV para a aprendizagem conceitual de subestações de energia elétrica, seus layouts, ativos componentes através do desenvolvimento de *checklist* e métricas.

Palavras Chave: potencialidades, 2D, realidade virtual, checklist, métricas.

INTRODUÇÃO

Por meio do uso de ferramentas computacionais, diversas tecnologias podem ser exploradas, tanto para o desenvolvimento de novos processos e produtos, quanto para treinamento e educação, como, por exemplo, as tecnologias aplicadas em ambientes 2D e RV (Realidade Virtual).

A técnica de representação de imagens 2D (bidimensional) conhecida na matemática desde a Grécia antiga através dos números figurais gregos em 2D (ALVES; NETO; MAIA; 2012) é usada para representar objetos e identidades de duas dimensões (largura e comprimento), tanto no papel, na escrita e na arte, como no computador, em jogos e diversas outras aplicações multimídias.

Algumas das características de sistemas de visualização em 2D é que tais sistemas permitem a análise de detalhes, a precisão de medições, a navegação à distância (TORY et al., 2006) sempre privilegiando a visão de todo o sistema. São ainda características de aplicações em 2D (PIMENTEL; DIAS; SANTOS, 2008) para garantia de sua usabilidade: a consistência das aplicações, o requerimento de sequências consistentes de ações em situações similares, o uso de terminologias idênticas nos menus e telas de ajuda, a consistência na utilização das cores, caracteres e a diferenciação entre letras maiúsculas e minúsculas.

De maneira natural, a tecnologia de representação de imagens e objetos evoluiu para o uso de técnicas em 3D (tridimensional), com a necessidade de expressar objetos e identidades de maneira mais realística (MANSSOUR, COHEM, 2006). Contudo, uma vez que os dispositivos de visualização são bidimensionais, a representação gráfica de objetos no espaço tridimensional passa obrigatoriamente por um processo de mapeamento da imagem 3D para 2D, chamado de projeção (MACHADO, 1997).

No entanto, “(...) não demorou muito para que todas essas tecnologias convergissem, rompendo a barreira da tela e do monitor e passassem a gerar ambientes tridimensionais interativos em tempo real, através da realidade virtual (...)” (TORI, KIRNER, 2006).

Nesse mesmo processo de transição se insere este exemplo de aplicação, no qual observamos a mudança do ambiente de treinamento dos operadores dos Centros de Controle e Operação de uma Concessionária de Energia, até então baseado em ambientes 2D, utilizando-se principalmente dos tradicionais diagramas unifilares, para um ambiente predominantemente RV.

Em 2013, o Laboratório de Realidade Virtual da Faculdade de Engenharia Elétrica da Universidade Federal de Uberlândia começou a desenvolver um Sistema de treinamento baseado na tecnologia de Realidade Virtual (RV) com o intuito de apresentar soluções pontuais aos ambientes 2D, empregados para o treinamento de operadores da Concessionária de energia CEMIG. Com este novo sistema, o operador em treinamento tem a oportunidade de navegar, por meio do ambiente de RV, em cada uma das subestações de energia da Concessionária e assim ter acesso a todos os seus

ativos. Porém, a maioria dos operadores tem apresentado forte esquiva em utilizar esta nova tecnologia e apresenta resistência à realização do treinamento em ambientes puramente RV.

Na modalidade tradicional de treinamento desses operadores, é necessária a estimulação de processos cognitivos em ambientes 2D relacionados a conhecimentos conceituais, como **reconhecer** formas e características específicas de equipamentos de medição e de proteção, **recordar** o modo de operação e manobra de equipamentos, **interpretar** a simbologia e **associar (o que é?)** ao equipamento a que se refere. Tais processos precisam também ser estimulados nos novos ambientes de realidade virtual.

Deste contexto, surge a seguinte questão: Como avaliar as potencialidades cognitivas dos ambientes de aprendizagem em 2D e RV para a assimilação conceitual de subestações de energia elétrica, seus layouts, ativos e componentes?

Com o intuito de responder a esta questão, este trabalho propõe uma metodologia de avaliação fundamentada na tecnologia de Realidade Virtual, na taxonomia dos processos cognitivos de aprendizagem conceitual, e nos conceitos de subestações de energia.

Esta pesquisa se justifica pelo fato de desconhecermos publicações que relacionem as potencialidades cognitivas do uso de ambientes 2D e de realidade virtual em processos de treinamento de operadores de subestações de energia, apesar do rigoroso levantamento bibliográfico.

OBJETIVO

Criar um protótipo de sistema de conhecimento para identificação das contribuições dos objetos 2D e RV para a aprendizagem conceitual de subestações de energia elétrica, seus layouts, ativos e componentes através do desenvolvimento de checklist e métricas.

FUNDAMENTAÇÃO TEÓRICA

Estudos considerados relevantes, como Miranda et al (2005) tratam do desenvolvimento do protótipo de um sistema, utilizando como interface um modelo virtual tridimensional que permite centralizar algumas informações, de forma que o usuário possa navegar pela subestação e selecionar os equipamentos para acessar informações armazenadas em um banco de dados, todavia relega as potencialidades e processos cognitivos envolvidos.

Pimentel, Dias e Santos (2008) discutem as características de usabilidade de ambientes 2D sem, contudo, avaliar os processos cognitivos estimulados nesses ambientes.

Com relação à avaliação de sistemas utilizando-se métricas, três trabalhos foram fundamentais: o primeiro de Siemer e Angelides (1998) publicado pela revista Elsevier Science trata da proposta de um método de avaliação tanto para a arquitetura quanto para o aprendizado adquirido por meio de sistemas completamente inteligentes. O principal objetivo desse estudo foi o de propor uma abordagem de

avaliação que sirva para fornecer um feedback detalhado sobre um sistema de tutoria inteligente completo com a finalidade de melhoria do sistema. O método de avaliação proposto neste artigo elimina a fraqueza de métodos adaptados pelo sistema, como os propostos por Nwana (1990), permanecendo independente do sistema, a fim de fornecer uma ferramenta para a avaliação de qualquer sistema de tutoria inteligente.

O segundo trabalho usado como referência foi o de Scholtz, Morse e Steves publicado na revista Elsevier B. V. em 2006 que também trata do desenvolvimento de métricas de avaliação para sistemas, os autores coletam e analisam dados do trabalho de analistas de inteligência em um sistema experimental, por volta de dois anos e desenvolve uma avaliação formativa do software pesquisado.

E, por último, nos apoiamos na pesquisa de Santos (2009) a qual desenvolve uma metodologia de avaliação para ferramentas computacionais para educação, utilizando-se uma estratégia de análise que compreende a organização da estrutura cognitiva do pensamento lógico-matemático, segundo a teoria operatória de Piaget.

Nesta seção, faremos ainda uma breve revisão teórica a respeito de dois assuntos principais, sobre os quais se fundamenta esta pesquisa:

- Taxonomia revisada de Bloom para os processos cognitivos de aprendizagem; e
- Subestações de energia

TAXONOMIA DOS PROCESSOS COGNITIVOS E APRENDIZAGEM CONCEITUAL

Várias pesquisas em Educação e Psicologia Cognitiva buscam compreender e classificar os processos de aprendizagem, investigando a forma como as pessoas aprendem. Em 1956, Benjamin Bloom escreveu a *Taxonomia dos Objetivos Educacionais: Domínio Cognitivo* e, desde então, sua descrição em seis níveis do raciocínio foi amplamente adotada e usada em inúmeros contextos. (BLOON, 1956).

Em 1999, Dr. Lorin Anderson, um antigo aluno de Bloom, e seus colegas publicaram uma versão atualizada da Taxonomia de Bloom que considera uma gama de fatores que afetam o ensino e a aprendizagem. Essa taxonomia revisada tenta corrigir alguns problemas da taxonomia original. Diferentemente da versão de 1956, a taxonomia revisada diferencia “saber o quê” (o conteúdo do raciocínio) de “saber como” (os procedimentos para resolver problemas) (ANDERSON et al, 2001; KRATHWOHL, 2002).

Nesta nova versão, adotada até hoje, a Dimensão do Conhecimento é o “saber o quê”. Ela se divide em quatro categorias: factual, conceitual, procedural e metacognitiva. O conhecimento factual inclui elementos isolados de informação, como definições de vocabulário e conhecimento de detalhes específicos. O conhecimento conceitual consiste em sistemas de informação, como classificações e categorias. O conhecimento procedural (saber como fazer) inclui

algoritmos, heurísticas ou métodos empíricos e técnicas, bem como o conhecimento sobre quando usar esses procedimentos. O conhecimento metacognitivo (refletir sobre o que se sabe) refere-se ao conhecimento dos processos cognitivos e das informações sobre como manipular esses processos de forma eficaz (KRATHWOHL, 2002).

O conhecimento procedural ou processual, também conhecido como aprendizagem procedural tem sido caracterizada como o saber como fazer alguma coisa e envolve seguir uma sequência de certas ações definidas que se destinam a produzir um resultado desejado (Hallett; Nunes; Bryant, 2010).

A aprendizagem procedural é pensada, na maioria das vezes, para apoiar a aquisição e retenção de conhecimento em função da melhoria de desempenho de trabalhadores qualificados e é indexada por tarefas em que a memória é expressa implicitamente por mudanças no desempenho, como resultado de experiência prévia (Willingham, Bullermer, Nissem, 1989).

Mas não há como falar em aprendizagem procedural sem possuir o domínio conceitual. No presente estudo, é de suma importância que o operador possua bem definidos os conceitos de Subestação de energia, Equipamentos de Manobra e de Medição, para então desenvolver as habilidades procedimentais necessárias às tomadas de decisão, mediante situações rotineiras e críticas. A este domínio damos o nome de aprendizagem conceitual (Willingham, Bullermer, Nissem, 1989).

Além disso, a taxonomia revisada é baseada em uma ampla visão de aprendizado que inclui não apenas a aquisição de conhecimento, mas também a habilidade de usar o conhecimento em uma variedade de novas situações (KRATHWOHL, 2002).

Dois dos mais importantes objetivos educacionais são promover retenção e promover transferência de conhecimento. Retenção é a habilidade de lembrar o que foi aprendido durante a instrução depois de algum tempo após tê-la recebido. Transferência é a habilidade de usar o que foi aprendido para resolver problemas, responder novas questões, ou facilitar o aprendizado de novos conceitos (BRANSFORD et al, 2000).

A aprendizagem ocorre quando os aprendizes constroem o conhecimento e os processos cognitivos necessários para a resolução de problemas de forma bem-sucedida. A solução de problemas envolve conceber uma forma de atingir um objetivo que nunca foi alcançado anteriormente; ou seja, descobrir como mudar a situação de seu estado dado em um objetivo (Mayer, 1992 apud Mayer 2002). Para isto, é necessário que conceitos tenham sido assimilados pelo aprendiz.

Quando o objetivo da instrução é promover retenção do material apresentado da mesma forma em que foi ensinado, a relevante categoria de processo cognitivo é “Lembre-se”. Lembrar envolve recuperação de conhecimentos relevantes na

memória de longo prazo. Os dois processos cognitivos associados são **Reconhecer** e **Recordar** (KRATHWOHL, 2002).

No entanto, quando o objetivo da instrução é promover a transferência de conhecimento, seu foco passa a ser as outras cinco categorias de processos cognitivos, **Entender** até **Criar**. Destes, sem dúvida, a maior categoria dos objetivos educacionais baseados em transferência enfatizados nas escolas e faculdades é Entender. Os alunos entendem, quando são capazes de construir significado, a partir de mensagens de instrução, incluindo comunicações orais, escritas e gráficas, e material apresentado durante as aulas, nos livros, ou em monitores de computador (KRATHWOHL, 2002). O processo cognitivo **Aplicar** envolve a utilização de procedimentos para executar, exercer ou resolver problemas e está intimamente ligada ao Conhecimento procedural. A categoria **Aplique** consiste em dois processos cognitivos: **executar**, quando a tarefa é um exercício (isto é, para o aluno familiarizado), e **implementar**, quando a tarefa é um problema (ou seja, desconhecida para o aluno) (KRATHWOHL, 2002).

Em ambientes de treinamento o objetivo da instrução geralmente é o de promover retenção, seja este ambiente 2D (tradicional) ou RV. Neste exemplo de aplicação, o operador em treinamento precisa ser capaz de saber o que é uma subestação de energia e saber como agir diante de possíveis falhas.

SUBESTAÇÕES DE ENERGIA

As subestações são entidades dentro do sistema elétrico, responsáveis pela elevação ou abaixamento de níveis de tensão (PINHEIRO, 1995). Há também subestações que fazem a interligação entre linhas de transmissão sem alterar os níveis de tensão (PINHEIRO, 1995).

As subestações de potência geralmente contêm diversos equipamentos de medição, os quais fornecem informações importantes para a operação do sistema elétrico, como supervisão e controle de fluxo de carga, acionamento da proteção contra faltas e coordenação de manobras no sistema. Os equipamentos de medição mais comuns são os TPs (transformadores de potencial) e os TCs (transformadores de corrente) cujas saídas analógicas são cabeadas para a sala de controle da subestação (BARBOSA et al, 2012). Também podem ser citados os Disjuntores e Religadores, cuja função principal é a proteção contra as correntes de curto-circuito (PINHEIRO, 1995). As subestações também possuem uma quantidade variada de chaves, linhas e barramentos, que podem ser identificados apenas como “ativos”. (PINHEIRO, 1995).

Os ativos em uma subestação são conectados em forma de Arranjos físicos. Para qualquer diagrama de uma linha elétrica existem normalmente vários arranjos possíveis: Barra simples, Barra principal e transferência, Barra dupla disjuntora simples, Barra dupla disjuntor duplo, Barra dupla disjuntora e meio, Barra em anel, entre outros.

Uma única falha em algum componente de uma subestação pode comprometer seriamente o fornecimento normal da energia elétrica, causando graves danos a consumidores e à própria concessionária (PINHEIRO, 1995).

As subestações estão sujeitas a muitos transtornos, como falhas nos mecanismos automáticos, falhas nos sistemas de proteção e nos equipamentos (PINHEIRO, 1995). Tais falhas podem causar interrupção de energia nos alimentadores e linhas de transmissão, fazendo necessárias as realizações de operações para reenergização do sistema. Essas operações também ocorrem quando se necessita realizar a manutenção de algum equipamento, realocação de carga e outros tipos de manobra (PINHEIRO, 1995).

Para realizar tais operações, o usuário do ambiente, seja ele 2D ou RV, precisa passar por um processo de treinamento rigoroso que estimule os processos cognitivos necessários a uma efetiva aprendizagem.

METODOLOGIA

Para o desenvolvimento deste estudo foi escolhida uma abordagem metodológica qualitativa devido ao caráter investigativo avaliativo, pois uma pesquisa qualitativa compreende um conjunto de variadas técnicas interpretativas que visam a descrever e a decodificar os componentes de um sistema complexo de significados. (MAANEM, 1979, p.520, tradução nossa).

Ademais, esta pesquisa trata sobretudo da elaboração de um plano de aplicação, definido por Tonus et al como “um projeto que possui todas as características e potencialidades para se efetivar em um produto tangível e intangível”.

O objeto da pesquisa foi o processo de transição do uso de ambientes 2D para ambientes RV no treinamento dos operadores da Concessionária de Energia CEMIG.

A proposta aqui apresentada é identificar processos cognitivos e as potencialidades dos ambientes 2D e RV para a aprendizagem de subestações de energia elétrica, seus layouts, ativos e componentes.

Para alcançar o objetivo proposto pretende-se desenvolver um protótipo de sistema avaliativo baseado em checklist e métricas que será aplicado aos operadores da Cemig em treinamento.

Para isso, primeiramente foi feito um levantamento bibliográfico a despeito de ambientes de treinamento e aprendizagem baseados em técnicas de visualização de imagens 2D e 3D e em tecnologias de realidade virtual. Seguiu-se à pesquisa bibliográfica de metodologias e protótipos de avaliação de sistemas ditos educacionais.

Em seguida foram realizados estudos comparativos de técnicas de representação de imagens em ambientes 2D e RV em relação aos processos cognitivos de aprendizagem para o entendimento de subestações de energia. A partir destes estudos, foi construído um quadro cujo exemplo segue abaixo (quadro 1):

Quadro 1 - Recorte Taxonomia dos Processos Cognitivos aplicados à Subestação de Energia

		Técnicas de representação de objetos		
		2D	3D	RV
Lembrar	Reconhecer	Equipamentos específicos através de sua simbologia gráfica	Equipamentos específicos através de suas formas.	Equipamentos específicos através de suas formas.
	Recordar	Função de cada equipamento através do diagrama unifilar.	Função de cada equipamento através da sua visualização e identificação.	Função de cada equipamento através da sua visualização, identificação e localização na subestação.

Em seguida, foram criados Mapas Conceituais (NOVAK, CAÑAS, 2010) de subestações de energia elétrica, seus layouts, ativos e componentes, com o intuito de alinhar conhecimentos, processos cognitivos e objetos de aprendizagem em 2D e RV.

Para a criação deste mapeamento foi utilizada a ferramenta Cmap Tools (CAÑAS *et al.*, 2004, <http://cmap.ihmc.us>). Fez-se um recorte dos principais conceitos de subestações de energia para exemplificar os conteúdos que o operador em treinamento precisa reter independentemente do tipo de ambiente utilizado, a partir dos quais obtivemos vários mapas conceituais. Na figura 1 é observado um dos mapas construídos, o mapa conceitual dos arranjos físicos de uma subestação de energia.

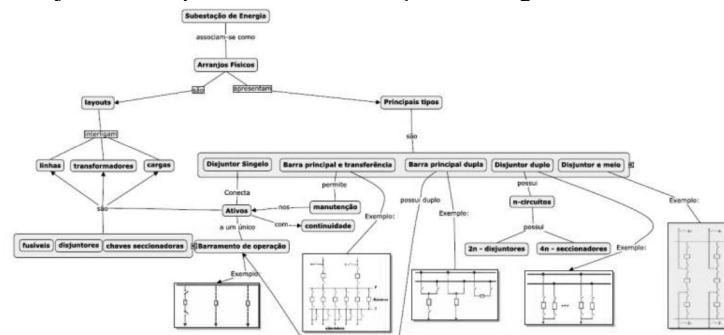


Figura 1 - Mapa conceitual dos layouts de subestações de energia.

Com base no mapeamento realizado, foi possível construir um quadro de entendimentos requeridos para cada um dos processos cognitivos de interesse em ambientes de treinamento de subestações de energia (quadro 2).

Quadro 2 - Processos cognitivos de entendimento em ambientes de treinamento de subestações de energia.

Processo cognitivo	Entendimento requerido
Interpretar	O que é uma subestação de energia;
Exemplificar	Tipos de equipamentos constituintes da subestação;
Classificar	Tipos de subestação de energia;
Deduzir	O que são falhas, o que as provocam, tipos de falhas, suas consequências.
Comparar	Tamanho e forma dos equipamentos de manobra, de transformação, de medição, de proteção.
Explicar	Quais são e como funcionam os equipamentos de manobra, de transformação, de medição, de proteção, etc.

A partir dos mapas e quadros criados, partiu-se para a elaboração de uma tabela para a compreensão sobre o

desenvolvimento da aprendizagem conceitual, segundo as características de modelagem do sistema e os processos cognitivos dos conhecimentos conceituais.

Cada um desses fatores foi considerado como um bloco ou conjunto de interrogativas capazes de elencar parâmetros desejáveis que visam à estimulação de processos cognitivos do simples ao complexo, formando um *checklist* (tabela 1) para o qual foram atribuídos pesos, de forma que a soma dos pesos para cada questão em uma mesma categoria seja igual a 1.

Tabela 1 - Checklist

Categoria A: Questões relativas às características de modelagem do sistema	
<i>Questões: Ao utilizar o Sistema RV Cemig</i>	<i>Pesos</i>
Você encontrou alguma dificuldade para diferenciar os ativos do ambiente?	0,2
Você encontrou dificuldades para distinguir os detalhes dos ativos?	0,2
Você encontrou alguma dificuldade ao se movimentar no ambiente usando o avatar? (3ª pessoa)?	0,2
Você encontrou alguma dificuldade ao se movimentar no ambiente usando a visão de Deus (1ª pessoa)?	0,2
Você sente algum desconforto visual durante o tempo em que fica exposto ao sistema?	0,2
Categoria B: Questões relativas a conhecimentos conceituais (Lembrar, reconhecer, recordar)	
<i>Questões: Ao utilizar o Sistema RV Cemig</i>	<i>Pesos</i>
Você apresentou alguma dificuldade para reconhecer ou identificar algum dos ativos da subestação?	x
Você apresentou alguma dificuldade para recordar a função de algum dos ativos da subestação?	x
Você apresentou alguma dificuldade para recordar como se opera algum dos ativos da subestação?	x
Você encontrou dificuldade para reconhecer os arranjos da subestação?	x
<i>Obs.: Fazer upload dos ativos e ou arranjos do banco de dados, repetir a pergunta para cada ativo ou arranjo do banco de dados, onde x = 0,25 / n° de ativos carregados</i>	
Categoria C – Questões relativas a conhecimentos conceituais (Entender, interpretar, exemplificar, classificar)	
<i>Questões: Ao navegar pelo Sistema RV Cemig</i>	<i>Pesos</i>
É possível compreender a lógica de funcionamento da subestação?	0,167
É possível compreender a relação entre os equipamentos?	0,167
Você consegue identificar quais equipamentos estão no estado ligado e no estado desligado?	0,167
Você consegue identificar os valores medidos nos equipamentos de medição?	0,167
Você consegue ler e compreender os valores dos transformadores de tensão?	0,167
Você consegue classificar os estados de operação em Normal, Emergência ou Restaurativo?	0,167
Categoria D – Questões relativas a conhecimentos conceituais (Resumir, deduzir, comparar e explicar):	
<i>Questões: Com base na observação do Sistema RV Cemig</i>	<i>Pesos</i>
Você encontrou alguma dificuldade para deduzir a ocorrência de falhas?	0,2
Você encontrou alguma dificuldade para explicar a ocorrência de falhas?	0,2
Você encontrou alguma dificuldade para fazer uma síntese do funcionamento da subestação visualizada?	0,2
Você encontrou alguma dificuldade para distinguir equipamentos com aparências semelhantes, mas com funcionalidades diferentes?	0,2
Você encontrou alguma dificuldade para identificar o estado atual de operação do sistema?	0,2
Categoria E – Questões relativas a conhecimentos conceituais no Sistema 2D (reconhecer, entender, interpretar, deduzir)	
<i>Questões: Ao usar o tradicional Sistema em 2D (diagramas unifilares)</i>	<i>Pesos</i>
Você encontra alguma dificuldade para reconhecer e identificar todos os elementos da subestação?	0,2
Você encontra alguma dificuldade para compreender a lógica de funcionamento da subestação?	0,2
Você encontra alguma dificuldade para identificar o funcionamento de cada um dos equipamentos?	0,2
Você encontra alguma dificuldade para interpretar os Estados de Operação dos equipamentos?	0,2

Você encontra alguma dificuldade para deduzir a ocorrência de falhas?	0,2
---	-----

A partir dessas interrogativas foram deduzidas métricas que ajudarão a caracterizar os processos cognitivos. Estas métricas foram organizadas em grades de análise no formato de questionário, visando à diminuição da subjetividade do processo. Para a quantificação do resultado final, foram elaboradas medidas e expressões para as métricas, inspiradas no modelo de Qualidade de software de McCall et al. (1977) apud Pressman (2002).

As métricas derivadas das questões são dadas assinalando, as opções de 1 a 5 que aparecerão para o usuário como uma régua ou termômetro, e ele através de um cursor móvel irá selecionar a opção desejada, que terá um peso de 0 a 4, de forma que:

Tabela 2 - Pesos atribuídos às respostas dos usuários

Opção selecionada	Peso
1. Nenhum / Nunca	0
2. A minoria / Na minoria das vezes	1
3. Alguns / Algumas vezes	2
4. A maioria / Na maioria das vezes	3
5. Todos / Sempre	4

A métrica utilizada será: $F_q = c_1m_1 + c_2m_2 + \dots + c_nm_n$

Onde:

F_q	Habilidades cognitivas desenvolvidas com o uso da ferramenta SRV Cemig
m	As notas atribuídas às métricas, definidas de 1 a 5.
c	Coeficientes que serão interpretados como pesos para as habilidades e de forma que $\Sigma = 1$.

De forma que:

<i>Se F(c) = 0</i>	O sistema estimula os processos de conhecimentos conceituais esperados.
<i>Se F(c) = 4</i>	O sistema não estimula os processos de conhecimentos conceituais esperados.
<i>Se 0 > F(c) < 4</i>	Alguns processos são estimulados pelo sistema, mas não todos os esperados.

Como as notas atribuídas pelo usuário variam de 0 a 4, e F(c) é função do somatório dos produtos das notas atribuídas pelos coeficientes para cada habilidade, o máximo valor possível encontrado, utilizando-se a métrica proposta, será igual a 4, valor atribuído conforme o máximo peso total possível, e o mínimo valor possível será 0.

RESULTADOS ESPERADOS

Esta métrica será aplicada para cada uma das categorias, proporcionando um feedback a respeito das potencialidades para cada processo cognitivo observado, fornecendo na forma de um gráfico de barras o grau de potencialidade propiciado por cada categoria, esteja ela relacionada ao ambiente tradicional em 2D ou ao ambiente em realidade virtual.

Pretende-se, ao final desta pesquisa, desenvolver um protótipo em HTML, no qual será possível fazer o upload de um banco de dados, contendo os ativos de cada subestação que aparecem no Sistema RV e então o próprio Sistema de avaliação irá generalizar as questões do Checklist, referentes à cada Categoria, replicando-as para cada um dos ativos presentes no banco de dados.

Este protótipo será validado mediante sua instalação e aplicação nos computadores dos operadores em treinamento e pretende proporcionar ao Grupo de Pesquisas do Laboratório de Realidade Virtual da Faculdade de Engenharia Elétrica da UFU um *feedback* do sistema de treinamento desenvolvido, possibilitando intervenções que busquem aprimorar a qualidade das pesquisas por eles produzidas.

CONSIDERAÇÕES FINAIS

O desenvolvimento da metodologia proposta pretende contribuir com a comunidade acadêmica, sugerindo um sistema de conhecimento para aferir medições à potencialidade de aquisição da aprendizagem conceitual em ambientes de aprendizagem em 2D e em RV.

AGRADECIMENTOS

À equipe do Projeto GT 411 pela disponibilização de acesso às informações contidas nesta pesquisa.

REFERENCIAS

- Alves, F. R. V., Neto, H. B., Maia, J. A. D. "História da matemática: Os números figurais em 2D e 3D". Conex. Ci. e Tecnol. Fortaleza/CE, v.6, n.2, p. 40-56, jul. 2012.
- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., Wittrock, M. C. (2001) "A Taxonomy for Learning, Teaching, and Assessing" — A Revision of Bloom's Taxonomy of Educational Objectives. Addison Wesley Longman, Inc.
- Barbosa, C. F. et al. "Avaliação do Uso de Tecnologia ZigBee no Ambiente de uma Subestação de Potência". Simpósio Brasileiro de Sistemas Elétricos (SBSE), v. 4, 2012. Disponível em: www.swge.inf.br/anais/sbse2012/PDFS/ARTIGOS/96668.PDF. Acesso em 04 nov. 2014
- Bloom, Benjamin S. "Taxonomy of Educational Objectives: The Classification of Education Goals. Cognitive Domain". Handbook 1. Longman, 1956.
- Bransford, J. D., Brown, A. L., Cocking, R. R., editors. "How people learn: brain, mind, experience, and school". Committee on Developments in the Science of Learning and Committee on Learning Research and Educational Practice, Commission on Behavioral and Social Sciences and Education, National Research Council. — Expanded ed. 2000.
- Cañas, A. J.; Hill, G.; Carff, R.; Suri, N.; Lott, J.; Eskridge, T.; et al. "CmapTools: A knowledge modeling and sharing environment". In: CAÑAS, A. J.; NOVAK, J. D.; GONZÁLEZ, F. M. (Eds.). Concept maps: theory, methodology, technology. Proceedings of the first international conference on concept mapping. Pamplona, Spain: Universidad Pública de Navarra, 2004b. v. 1, p.125-133.
- Hallett, D., Nunes, T., Bryant, P. Individual Differences in Conceptual and Procedural Knowledge When Learning Fractions. Journal of Educational Psychology, 2010 American Psychological Association, Vol. 102, No. 2, 395–406. 2010.
- Krathwohl, D. R., (2002) "A Revision of Bloom's Taxonomy: An Overview. THEORY INTO PRACTICE", Volume 41, Number 4, Autumn Copyright. College of Education, The Ohio State University.
- Maanem, J. V, Reclaiming qualitative methods for organizational research: a preface, In Administrative Science Quarterly, vol. 24, nº 4, December, 1979, pp 520 – 532.
- Machado, L. S. "A realidade virtual em aplicações científicas". Dissertação de Mestrado em Computação aplicada, INPE, São José dos Campos, Abril. 1997.
- Manssour, I. H., Cohen, M. "Introdução à computação gráfica". RITA, volume XXIII, número 2, 2006.
- Miranda, M. F. et al. Interface 3D aplicada à visualização de informações de uma subestação de energia armazenadas em banco de dados. In: Workshop De Aplicações De Realidade Virtual. 2005.
- Novak, J. D., Cañas, A. J. "A teoria subjacente aos mapas conceituais e Como elaborá-los e usá-los".
- Nwana, H.S., "The evaluation of an intelligent tutoring system, Intell. Tutoring Media", 1990.
- Pimentel, A., Dias, P., Santos, B. S. "Avaliação de usabilidade em sistemas de realidade virtual e aumentada: principais métodos". Revista do Detua, v. 4, nº 9, junho 2008.
- Pinheiro, H.B. (1995) "Sistema especialista para operação de subestações de energia elétrica". Disponível em: <http://www.bibliotecadigital.unicamp.br/document/?code=vtls000093990> Acesso em 03 nov. 2014.
- Pressman, R. S. "Engenharia de software". Trad. Mônica Maria G. Travieso. 5. ed. Rio de Janeiro: McGraw-Hill, 2002.
- Scholtz, J., Morse, E., Steves, M.P. Evaluation metrics and methodologies for user-centered evaluation of intelligent systems. Elsevier 2006.
- Siemer J., Angelides, M. C. A comprehensive method for the evaluation of complete intelligent tutoring systems. Elsevier Science, 1998.
- Tori, R., Kirner, C. "Fundamentos de realidade virtual", capítulo 1 in Fundamentos e Tecnologia de Realidade Virtual e Aumentada. Livro do Pré Simpósio VIII Symposium on Virtual Reality. Belém – PA, 02 de maio de 2006.
- Tory, M., Kirkpatrick, A. E., Atkins, M. S., Möller, T. "Visualization tasks performance with 2D, 3D, and combination displays". IEEE Transactions on visualization and computer graphics, 2-13. 2006.
- Willingham, D. B., Bullemer, P., Nissen, M. J. On the Development of Procedural Knowledge Journal of Experimental Psychology: Learning, Memory, and Cognition. Vol. 15, No. 6, 1047-1060. 1989.

Uma Plataforma para o Desenvolvimento de Múltiplos Atlas 3D Colaborativos para Educação da Anatomia Humana.

A Platform for Development of Multiple 3D Collaborative Atlas for Education on Human Anatomy

Paulo Vinícius de Farias Paiva¹

¹paulo.fariaspaiva@gmail.com

Resumo- Os Ambientes Virtuais (AVs) baseados em Realidade Virtual (RV) tem trazido benefícios para diferentes áreas do conhecimento. Especialmente na educação em saúde, estes sistemas possibilitam a diminuição dos custos envolvidos no processo educacional, possibilitam maior interatividade e auxiliam na avaliação dos conhecimentos adquiridos por parte dos estudantes. Neste contexto, surgiram diferentes AVs que constituem atlas virtuais 3D do corpo humano, em auxílio do ensino de anatomia. Neste trabalho será apresentada uma arquitetura para o desenvolvimento de atlas virtuais 3D, bem como os resultados preliminares da implementação de um protótipo. Também são apresentados os resultados de uma análise comparativa entre diferentes atlas virtuais pesquisados na literatura, os quais fundamentaram o processo de análise de requisitos da aplicação gerada.

Keywords-component; *Atlas Virtual 3D, Anatomia, Realidade Virtual, Colaboração, Interação Natural*

I. INTRODUÇÃO

A Realidade Virtual (RV) oferece aos seus usuários a possibilidade de interagirem em ambientes virtuais 3D, e estimular os canais sensoriais humanos, possibilitando a simulação de experiências do mundo real. A medicina tem se apresentado certamente como uma das áreas mais beneficiadas com o advento da RV [2]. Atualmente no processo de formação médica das universidades, existem algumas limitações particularmente no estudo da fisiologia humana. No que se refere às atividades práticas, existem as dificuldades relacionadas ao uso de cadáveres, tais como: problemas éticos envolvidos, alto custo de aquisição e de manutenção, tempo de uso limitado devido à degradação das peças e a impossibilidade de uso fora das universidades. Deste modo, como alternativas didáticas, são utilizados diagramas, manequins artificiais e os atlas de papel. Os atlas são livros, que servem como verdadeiras encyclopédias das estruturas internas e externas anatômicas humanas (sistemas e órgãos), contendo imagens e informações gerais ou específicas [2].

No entanto, é notória a existência de dificuldades encontradas pelos aprendizes quando realizam o estudo de estruturas anatômicas impressas em papel, pela natural dificuldade de representação fidedigna destes modelos 3D. Soluções baseadas em RV, têm sido propostas por diferentes pesquisadores nas últimas décadas, que exploram

diferentes recursos e capacidades destas aplicações, tais como a possibilidade de geração de modelos volumétricos, aplicação de transparência e outros efeitos de visualização (ex: estereoscópica) e diferentes meios de interação.

Visando a obtenção o conhecimento dos principais recursos utilizados nestes ambientes, bem como auxiliar na definição dos requisitos da arquitetura proposta neste trabalho, foi realizado um levantamento de aplicações de RV desta natureza. Este trabalho apresenta uma plataforma web colaborativa, voltada para o desenvolvimento de múltiplos Atlas 3D. Dentre as principais vantagens oferecidas, destacam-se a sua escalabilidade, ou seja, capacidade de suporte a múltiplos usuários, bem como a sua persistência. A plataforma se utiliza do conceito de conectividade em nuvem, contando com uma estrutura de servidores dedicados. Deste modo, as formas de consultas se expande para o contexto de colaborativo.

II. ESTADO DA ARTE

Neste tópico, serão apresentadas algumas aplicações encontradas de atlas virtuais tridimensionais, sendo estas desenvolvidas com diferentes tecnologias, objetivos e funcionalidades. O principal objetivo deste levantamento, foi realizar uma análise dos principais recursos computacionais disponibilizados e que vêm sendo explorados por estes tipos de aplicações tais como: tipos de visualização de modelos abordados, tipos de operações permitidas, dispositivos de interação utilizados, dentre outros. Deste modo, uma série de requisitos foram definidos e serviram de base para a construção do modelo de arquitetura proposto por este trabalho. As aplicações analisadas foram as seguintes: *AICoH* [1], *Anatomi3D* [2], *AVIDHa* [3], *Zigotte Body*[4], *Visible Body* [5] e *Anatronica* [6]. Um quadro comparativo entre as ferramentas pode ser visualizado na Tabela 1, onde as células que contenham o símbolo '---' representam a não especificação do atributo pesquisado nas referências.

No que se refere as *tecnologias de desenvolvimento* utilizadas, destacamos que algumas das aplicações podem ser executadas via *web*, graças aos padrões específicos (*webGL*, *HTML 5.0*, *jQuery* e *JSON*) tais como a *AICoH* (*Atlas Interativo do Corpo Humano*), o *Zigotte Body* e a *Anatronica*. Em relação à interação com os modelos dos

objetos virtuais são possíveis basicamente todas as transformações usuais (translado, rotação e escala), bem como outras formas de interação mais complexas, tal como a seleção de entidades fisiológicas específicas (veias, ossos, órgão vital, etc).

Tabela 1. Estudo comparativo entre aplicações de Atlas Virtuais.

Aplicação	Tecnologias	Interações com os Modelos	Interação / Visualização	Tipo de Colaboração
AICoH [1]	1- WebGL 2- HTML 5.0 3- Jquery-Ui 4- XTK 5- JSON	1- Operações básicas de transformação. 2- Transparência. 3- Zoom.	Mouse	---
Anatomi 3D [2]	1- VRML 2- C/C++ 3- OpenGL	1- Operações básicas de transformação. 2- Transparência. 3- Zoom.	Plataforma VirtWall	Presencial via projeção
AVIDHa [3]	1- Placa de aceleração física PhysX *Linguagem não especificada	1- Operações básicas de transformação. 2- Transparência. 3- Zoom. 4- Modelos volumétricos 5- Planos de corte 6- Colisão/deformação	Mouse 3D, Joystick e PHANTOM Omni	Presencial via projeção
Zigotte Body[4]	1- WebGL 2- Javascript 3- JSON	1- Operações básicas de transformação. 2- Transparência	Mouse	---
Visible Body [5]	---	1- Operações básicas de transformação. 2- Seleção de entidades fisiológicas 3- 2.500 estruturas disponíveis	Mouse	---
Anatronica [6]	1- Unity 3D	1- Operações básicas de transformação. 2- Seleção de entidades fisiológicas 3- Transparência	Mouse	---

Em geral, a aplicação de efeitos especiais como a transparência e geração de imagens estereoscópicas sobre os modelos também são utilizados. A aplicação *AVIDHa* (Atlas Virtual Interativo Distribuído Háplico) mostrou-se ser a que disponibiliza maior número de recursos, especialmente a possibilidade de aplicação de diferentes planos de corte nos modelos, uso de modelos volumétricos (múltiplas camadas), colisão e deformação dos modelos. Além disto, é a única que possibilita interação haptica via uso do dispositivo *PHANTOM Omni®* e visualização em alta definição. A *Visible Body* é a que disponibiliza maior número de modelos tridimensionais realistas (2.500 estruturas) projetadas por uma equipe multidisciplinar de especialistas [5].

Como observado, o recurso colaborativo é pouco utilizado dentre as aplicações analisadas, sendo esta possibilidade apenas explorada no *Anatomi 3D* e *AVIDHa*, onde é oferecida a possibilidade de um grupo de usuários interagirem colaborativamente por meio de uma plataforma de projeção. Como observado nos sistemas pesquisados, a colaboração torna-se limitada por ter como requisito a presença física dos participantes que interagem via telas de projeção, ou mesmo via sistemas *web* não-escaláveis. Nestes casos, além do espaço físico (ou virtual) ser limitado, existem restrições no que se refere ao compartilhamento de conhecimentos especialistas para diferentes níveis de aprofundamento. Por exemplo, requisitos como a qualidade

gráfica dos sistemas fisiológicos virtuais se distinguem entre estudantes de ensino médio e estudantes de pós-graduação.



Figura 1. Atlas com recurso colaborativo, ou seja, estudo em grupo via tela de projeção: *Anatomi3D* e o *AVIDHa*.

II. ARQUITETURA PROPOSTA

Segundo orientações encontradas na literatura [XX] de boas práticas de desenvolvimento destas aplicações, alguns requisitos devem ser previamente estabelecidos na fase de desenvolvimento, tais como: (a) Autenticação de usuários; (b) Armazenamento da base de modelos gráficos; (c) Protocolos de compartilhamento de dados; (d) Plataforma de execução (*web*, *standalone*, etc); (e) Modos de colaboração; (f) Técnicas interativas, dentre outras.

Neste sentido, a arquitetura de uma plataforma de múltiplas aplicações colaborativas e imersivas é proposto, de modo que a execução e manutenção destas aplicações seja escalável. A plataforma *web* denominada *UniColAtlas-3D (Unity Collaborative Atlas)* é um sistema colaborativo de larga escala que permite a criação e manutenção de diversos Atlas 3D Colaborativos simultaneamente. A unidade básica da plataforma é intitulada de *uniAtlas* e refere-se a cada AVC gerado. Cada *uniAtlas* possui diferentes salas, que são criadas e gerenciadas por um único usuário (*roomOwner*), designado por padrão pelo administrador do *uniAtlas* em questão (*atlasOwner*). O administrador possui controle absoluto de uma sala, incluindo o controle de todos modelos 3D visualizados, controle sobre os usuários conectados, dentre outras funções, podendo passar este controle para outros usuários. Neste sentido, os usuários de cada *uniAtlas* podem se conectar e transitarem entre diferentes salas relacionadas a diferentes contextos, como por exemplo, diferentes tipos de sistemas fisiológicos do corpo humano.

Na criação dos Atlas, são passados todos os parâmetros e recursos desejáveis (número de usuários e de salas suportados, recursos de RV, etc) para a plataforma, que serão utilizados *a priori* por todas as suas salas, que são criadas *a posteriori* pelos seus participantes. O conjunto de *k-uniAtlas* gerados referem-se a *k* diferentes níveis de aprofundamento do conhecimento (ex: ensino fundamental, médio e graduação). Desta forma, o uso de diferentes técnicas interativas podem ser variado facilmente, de acordo com o perfil do grupo de estudantes de determinado *uniAtlas*. Assim, o uso de técnicas de interação/imersão mais simples (ex: estereoscopia, mouse e teclado) até as

mais robustas (ex: dispositivos hápticos, *CAVEs*, luvas de dados) podem ser variadas.

Outro aspecto importante a ser destacado, é que diferentes abordagens pedagógicas podem ser exploradas pelos administradores dos AVCs, que no papel de tutores, podem se utilizar de quizzes, *serious games* embutidos, dentre outros recursos. Segundo Machado e Moraes [X] não haveria sentido o uso de AVs para educação/treinamento sem que houvesse a possibilidade da provisão de avaliação dos seus usuários. Assim, a classificação do conhecimento de cada estudante em particular, bem como de um grupo torna-se essencial. Inicialmente, um usuário deve realizar um cadastro e se conectar ao servidor *web* onde está armazenada a plataforma *uniColAtlas*. Logo após sua autenticação, o usuário é conectado a um servidor em nuvem, sendo gerado um *uniAtlas* de acordo com parâmetros requisitados pela plataforma ao usuário. Após devidamente processado, é retornada a visualização do ambiente, onde o usuário poderá interagir e colaborar com antigos usuários (previamente conectados) (Figura 2).

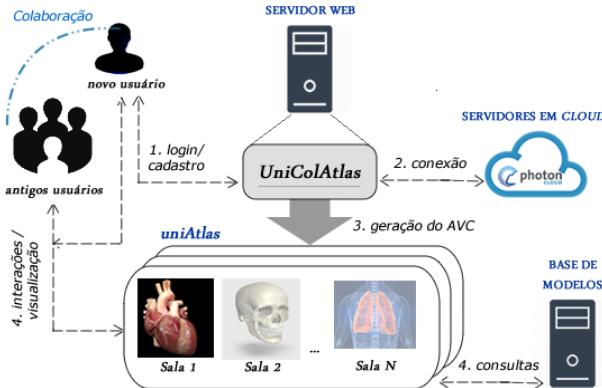


Figura 2. Arquitetura da Plataforma *uniColAtlas*.

Este é um diferencial da plataforma, pelo fator de persistência do AVC e sua escalabilidade, visto que a estrutura de servidores em nuvem providenciam maior garantia de conectividade e suporte a múltiplos usuários. A arquitetura de *software* da plataforma encontra-se dividida da seguinte forma:

- **Módulo de Avaliação:** Diferentes técnicas de I.A. (modelos probabilísticos, lógica fuzzy, árvores de decisão, etc.) podem ser utilizados para o reconhecimento dos padrões e classificação do conhecimento de cada estudante. Logo após o término das interações, são exibidos logs ou relatórios de desempenho.
- **Base de dados Multimídia:** Responsável pelo armazenamento dos modelos 3D, arquivos contendo as definições e conceitos pedagógicos usados.
- **Módulo de Gerenciamento de AVCs:** Gerencia a criação e atualização dos *Atlas* gerados, sendo cada aplicação

responsável pela lógica de suas salas (com o auxílio do módulo de redes).

- **Módulo de Rede:** autenticação dos usuários, inicialização e manutenção das conexões, inserção de usuários em diferentes salas e *Atlas 3D*.

III. RESULTADOS PARCIAIS: IMPLEMENTAÇÃO DE UM PROTÓTIPO DE UNITLAS

Inicialmente, foi implementada uma aplicação do tipo *uniAtlas* a fim de serem avaliadas as funcionalidades e tecnologias que serão utilizadas e disponíveis automaticamente e posteriormente pela plataforma *uniColAtlas*. Para o desenvolvimento da aplicação foram utilizadas as seguintes tecnologias:

- **Uniy 3D** - Motor de desenvolvimento de jogos e simuladores gráficos 3D, contendo uma série de recursos importantes para este contexto previamente implementados, tais como: tratamento de colisões, física, renderização gráfica, comunicação em rede, etc.
- **PUN (Photon Unity Network)** - Solução para criação e gerenciamento de ambientes virtuais escaláveis em rede com suporte a múltiplos usuários, particularmente aplicações de jogos 3D que pode ser integrado ao *Unity 3D*.
- **Linguagens de Programação/Marciação** - C#, PHP, XML,
- **Leap Motion** - Dispositivo periférico para rastreamento manual. Trata-se de um controlador USB, projetado para ser colocado em um *desktop* físico, voltada para cima (Figura X). O dispositivo utiliza duas câmeras e três LEDs infravermelhos monocromáticos, para observar uma área hemisférica, a uma distância de aproximadamente um metro. Os LEDs geram um padrão 3D de pontos de luz infra-vermelho [21] e as câmeras geram ~300 fps de dados refletidos, construindo o mapa ou região de rastreamento (Figura 3).

A aplicação de teste é constituída por uma lista de objetos relacionados ao sistema ósseo do corpo humano, sendo estes carregados remotamente, dentro de diferentes salas. Inicialmente foram geradas duas salas, contendo dois modelos (crânio e coluna vertebral) que podem ser visualizados em estereoscopia passiva por anaglifo (Figura 3) e facilmente manipulados via teclado e mouse. Também é possibilitado ao usuário a interação via gestos de Interação Natural (IN), ou seja, de modo intuitivo, para realização de transformações como translação, rotação e escala (Figura 4).

A aplicação foi desenvolvida com uma arquitetura de rede cliente-servidor. Cada modelo é armazenado em um servidor *web* que recebe as requisições dos clientes sobre atualizações dos modelos, para que seja efetuada a renderização gráfica nos clientes locais.

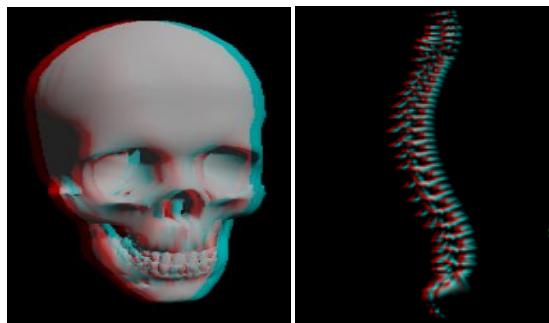


Figura 3. Visualização estereoscópica por anaglifo dos modelos utilizados (crânio e coluna vertebral).

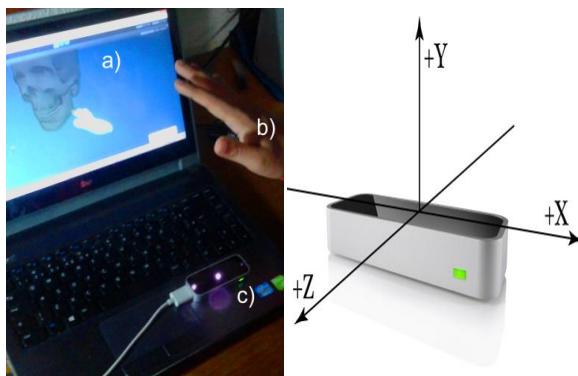


Figura 4. Imagem da aplicação à esquerda: a) modelos virtuais de um crânio e da mão; b) mão do usuário; c) Dispositivo de IN *Leap Motion*. Sistema de coordenadas do dispositivo.

Desta forma, é garantido o sincronismo das diferentes instâncias do *uniAtlas*. Ao conectar-se é dada a possibilidade de o usuário visualizar textos informativos das peças anatômicas e realização de questionários acerca do modelo estudado. Estas informações encontram-se armazenadas em estruturas de arquivos do tipo XML referentes a cada modelo e que são carregados dinamicamente. Como forma de avaliação inicial dos questionários, foi implementado um algoritmo que computa a estatística Alpha (α) de Cronbach [7]. Esta estatística é frequentemente utilizada para se mensurar estimativas da confiabilidade de um dado questionário, pelo grau de correlação entre seus itens da seguinte forma:

$$\alpha = \frac{k}{k - 1} \cdot \left[1 - \frac{\sum_{i=1}^k S_i^2}{S_t^2} \right]$$

Onde k = nº de itens (perguntas) do questionário; S_i^2 = variância do i -ésimo item ($i = 1, \dots, k$); S_t^2 = variância do total T de cada estudante j nos k itens.

Após computação da estatística α , o algoritmo de avaliação realiza a classificação do valor obtido de acordo com as classes propostas por Freitas e Rodrigues em [7]. Logo após a interação dos usuários com o *uniAtlas* é apresentado um relatório de avaliação por indivíduo e do

grupo de estudantes. Os dados de interação, refletindo o acesso a determinadas estruturas da peça anatômica, screenshots gerados na interação, os textos informativos consultados e as questões respondidas são registrados para futuras análises.

IV. CONSIDERAÇÕES FINAIS

Como observado, os atlas virtuais baseados em RV apresentam diferentes recursos, tais como a possibilidade de interação com os modelos anatômicos 3D volumétricos, com maior realismo gráfico, uso de interfaces intuitivas, e dispositivos de interação não convencionais de RV. O enfoque deste trabalho é particularmente sobre a capacidade de colaboração entre diferentes estudantes, numa perspectiva de promoção de educação/avaliação continuada. Desta forma, é apresentada a plataforma *uniColAtlas* capaz de gerar, armazenar e gerenciar Atlas 3D com múltiplos usuários em nuvem. Finalmente, foi apresentada uma aplicação de protótipo, para validação de alguns conceitos da arquitetura que serão implementados futuramente.

V. REFERÊNCIAS

- [1] Cardia, L.C.; Thomaz, V.A.; Oliveira, J.C. (2014). "AICoH 3D: Atlas Interativo do Corpo Humano AICoH 3D: Interactive Atlas of Human Body." In Proceedings of Symposium on Virtual Reality (SVR'2014).
- [2] Monteiro, S. B., Valdek, M. C. O., Cunha, I. L. L., Moraes, M. R., Machado, L. S., AnatomI 3D: Um Atlas Digital Baseado em Realidade Virtual para Ensino da Medicina. In Proceedings of Symposium on Virtual Reality (SVR'2006), Belém. pp. 3-14.
- [3] S.M. Malfatti, L.M. Fraga, J.C. de Oliveira, S.R. dos Santos, P.F.F. Rosa, "Um Atlas 3D Háplico para o Estudo de Anatomia" Workshop de Informática Médica, Belém, PA, Brazil, 2008.
- [4] Visible Body Atlas. Disponível em: <<http://www.visiblebody.com/index.html>>. Acessado em: 8 de jan. de 2015.
- [5] Zygote Body Atlas. Disponível em: <<http://www.zygotebody.com/index.html>>. Acessado em: 8 de jan. de 2015.
- [6] Anatronica. Disponível em: <<http://anatronica.com/>>. Acessado em: 8 de jan. de 2015.
- [7] A. L. P. Freitas; S. G. Rodrigues. A avaliação da confiabilidade de questionários: uma análise utilizando o coeficiente Alfa de Cronbach. XII SIMPED, Bauru/SP, 2005.
- [8] Burdea, G. C., Coiffet, P. Virtual Reality Technology. Wiley-Interscience, second edition, 2003.
- [9] Paiva, P.V.F.; Machado, L.S.; Oliveira, J.C. A peer-to-peer Multicast Architecture for Supporting Collaborative Virtual Environments (CVEs) in Medicine In: Symposium on Virtual and Augmented Reality 2012. Niterói/RJ - Brazil. pp. 165-173, 2012.

Tangible Touch Screen Mobile Device for Visually Impaired People: A Braille Text Entering Approach

Fabrizzio Alphonsus A. M. N. Soares, Joyce Siqueira, Mateus M. Luna*, Luciana de Oliveira Berretta, Cristiane B. R. Ferreira, Ronaldo M. da Costa, Leandro L. G. de Oliveira, Rogério Salvini

Instituto de Informática
Universidade Federal de Goiás - UFG
Goiânia, Brasil

{fabrizzio, joycesi, luciana, cristiane, ronaldocosta, leandroluis, rogeriosalvini}@inf.ufg.br, *mateus.m.luna@gmail.com

Abstract—Mobile devices with touch screen have brought greater task dynamicity and Interaction. However, the lack of a physical keyboard makes mobile devices hard to be used by visually impaired people. The aim of this work is to develop a method and a interface to support the needs of this audience in text entering. Thereby, the Braille method is used and a Android app is being developed for simulating a virtual keyboard. In order to allow user screen handling and interaction, a plastic film with embossed dot marks is being built. This plastic film has the 6 Braille dots and some others for control and interaction. The interface validation is being performed with support of a visually impaired collaborator, with congenital blindness and Braille literate. In order to improve the interface and the text entering software the volunteer performs a lot of experiments and evaluations. The specialist evaluation shows the proposed method has a great potential. Once this stage is completed usability tests with greater number of volunteers will be performed.

Keywords—Braille, Interface, Text Entering, Touch Screen, Visually Impaired People.

Resumo—Dispositivos móveis com telas sensíveis ao toque trouxeram maior dinamicidade e interatividade às tarefas. Todavia, a ausência do teclado físico torna seu uso por deficientes visuais uma atividade muito complexa. Por isso, o objetivo desta pesquisa é desenvolver um método e uma interface que atendam às necessidades deste público, no que tange a entrada de texto. Para isto foi adotado o método braile, implementado em um aplicativo para Android, que simula um teclado virtual. Para a manipulação e interação da tela, uma película, com pontos em alto relevo, está sendo desenvolvida. Esta película contém, além dos 6 pontos do braile, pontos para controle e interação. A validação da interface está sendo realizada por uma colaboradora, deficiente visual congênita e alfabetizada em braile. A cada teste, tanto a interface quanto o aplicativo são aprimorados. Os resultados parciais demonstram que o método proposto possui potencial, desta forma, finalizada a avaliação preditiva, pretende-se realizar testes de usabilidade com um número maior de voluntários.

Palavras Chaves—Braile, Interface, Entrada de Texto, Telas Sensíveis ao Toque, Deficientes Visuais.

I. INTRODUÇÃO

No Brasil há cerca de 580 mil pessoas cegas [4]. No mundo, este número chega a 39 milhões [8]. Para estas pessoas, investimentos em tecnologias assistivas oportunizam uma maior integração com a sociedade, atualmente envolta em ferramentas tecnológicas que progridem rapidamente e influenciam na comunicação, no trabalho e no lazer, transformando o cotidiano.

Existem diferentes iniciativas voltadas à acessibilidade para computadores, principalmente, softwares leitores de tela. Contudo, os computadores vêm perdendo espaço para maciça utilização de smartphones, que mesmo munidos com recursos de

acessibilidade, ainda não atendem muitas das necessidades básicas para seu uso por deficientes visuais.

Neste contexto, o escopo desta pesquisa se restringe ao desenvolvimento de um método e de uma interface que viabilizem a entrada de texto, utilizando o método braile, em aplicativos como, por exemplo, de envio de SMS, em smartphones com telas sensíveis ao toque sem o teclado físico acoplado.

Desta forma, apresenta-se neste artigo os resultados parciais obtidos durante o processo de desenvolvimento, que por considerar a especificidade do público alvo culminou em dois diferentes protótipos, o aplicativo para Android e a película tátil, que convergem em uma única interface.

A primeira etapa, a avaliação preditiva, está sendo realizada com uma colaboradora com cegueira congênita, ou seja, adquirida antes dos 5 anos de idade e alfabetizada em braile, a qual será identificada como C1. A segunda etapa, avaliação experimental, será realizada por um grupo de colaboradores (C2) com o mesmo perfil, detalhado no item Materiais e Métodos, que utilizará e posteriormente avaliará os protótipos por meio de questionários de usabilidade.

As próximas seções apresentam conteúdo relevante para entendimento acerca dos resultados parciais obtidos. Na seção II, uma introdução ao alfabeto braile. Na III, Referencial Teórico, os trabalhos relacionados. Na seção IV, Materiais e Métodos, detalhes sobre a realização do experimento. Na V, descreve-se os Resultados Parciais e por último, seção VI, as Considerações Finais.

II. O ALFABETO BRAILE

O principal incentivo para criação do método braile foi a tentativa, no século XVIII, de oferecer educação aos cegos. Neste período, Valentin Haüy partiu do pressuposto que o principal problema consistia “em fazer que o visível se tornasse tangível” [1].

Em 1819, o capitão de artilharia Carlos Barbier de la Serre realizou novos estudos sobre a escrita dos cegos, com o intuito de manter o sigilo das mensagens militares e diplomáticas. Suas pesquisas subsidiaram Louis Braille, que aos 5 anos perdeu totalmente a visão, a desenvolver o Sistema Braile que após várias adaptações, constitui-se atualmente por 63 sinais, gerados pela combinação de 6 pontos agrupados em duas fileiras, que não excedem o campo tátil e podem ser facilmente identificados [1].

A Fig. 1 apresenta os caracteres braile de “a” a “z”, em minúsculo, sem números, acentuação ou caracteres especiais.

Um texto em braile é apresentado com pontos em alto relevo, grafados no papel, onde o deficiente visual realiza a leitura tateando o documento, a fim de identificar os símbolos inseridos, conforme Fig. 2.

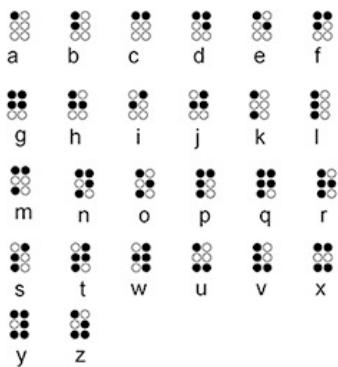


Fig. 1. Alfabeto braile básico



Fig. 2. Leitura em braile.

Para a construção deste documento, os sinais podem ser inseridos com o auxílio de um reglete de mesa, que é uma prancheta com uma régua de orientação para gravar ponto a ponto utilizando a punção, uma espécie de furador (Fig. 3).



Fig. 3. Reglete e punção

Mais recentemente foram desenvolvidas alternativas para auxiliar a escrita, como máquinas de escrever (Fig. 4) e com o advento do computador, as Impressoras de Impacto para Braile (Fig. 5).

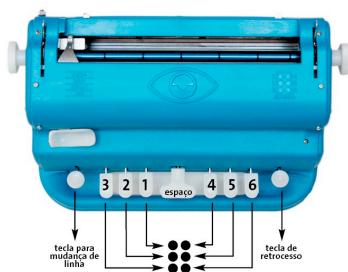


Fig. 4. Máquina de escrever

Os aparelhos diferem na facilidade e modernidade, mas produzem o mesmo resultado: um papel com pontos em alto relevo, semelhante à Fig. 2.

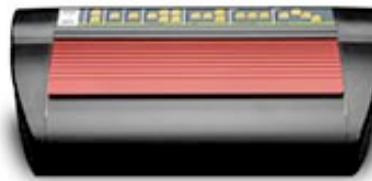


Fig. 5. Impressora de impacto para braile

A relevância do sistema braile e sua aceitação em diversos países resultou em pesquisas sobre a entrada de texto em *smartphones* com telas sensíveis ao toque, utilizando o braile. Pois assim, pretende-se tornar possível a utilização produtiva desta ferramenta, cuja evolução acarretou em problemas de usabilidade, por ter eliminado o teclado físico, essencial para este público.

III. REFERENCIAL TEÓRICO

Romero et al (2011) desenvolveu o projeto *BrailleTouch*, com a finalidade de possibilitar uma alternativa a entrada de texto em *smartphones* com telas sensíveis ao toque, para deficientes visuais. O *BrailleTouch* não requer peça adicional de hardware, contribuindo para mobilidade. Tem como aspecto chave na interface ter menos botões do que dedos, pois desta forma, o usuário não precisaria mover na tela para encontrar as sequências e combinações corretas. Os botões correspondem espacialmente ao carácter braile e fornece feedback sonoro a medida que os símbolos são inseridos. Também foram implementados um botão central, que representa a barra de espaço e o retrocesso, representado pela combinação de tecla 6 e a barra de espaço [7].

Frey, Southern e Romero (2011) testou o *BrailleTouch* com o propósito de verificar se o mesmo poderia ser um produto acessível, do ponto de vista financeiro, em relação as demais opções do mercado. Como resultado, concluiu-se que em comparação a soluções existentes o *BrailleTouch* tem o potencial para ter um custo reduzido [3].

Oliveira et al (2011) propôs o *BrailleType*, um método de entrada de texto baseado no alfabeto braile. Destaca que evita gestos *multi-touch* a fim de criar uma interação mais simples, com o uso de um único dedo. Neste artigo também realizou-se a comparação do desempenho do *BrailleType* em relação ao *VoiceOver*, da *Apple*. Como resultado, o *BrailleType* apresentou-se um pouco mais lento, porém, significativamente mais fácil e com menos erros [5].

Oliveira et al (2011) realizou um estudo comparativo entre quatro diferentes métodos para entrada de texto: *QWERTY*, *MultiTap*, *NavTouch* e *BrailleType*, com o objetivo de identificar as vantagens e desvantagens destas abordagens e para reconhecer se, de fato, os deficientes visuais poderiam se beneficiar de um método sobre os outros. Analisou-se desempenho, preferências e principalmente as características individuais dos usuários, dando destaque a relevância de não se negligenciar as diferenças entre os usuários, para desta forma, sejam capazes de fornecer variações mais inclusivas, maximizando cada performance [6].

El-Glaly et al (2013) apresenta a ideia de projetar uma sobreposição para proporcionar um feedback tangível para a leitura, por deficientes visuais, em dispositivos de toque. A sobreposição foi testada em um estudo de usabilidade e os resultados mostraram o papel da tangibilidade em alavancar a acessibilidade dos dispositivos de toque [2].

IV. MATERIAIS E MÉTODOS

Esta pesquisa é desenvolvida em seu processo metodológico por dois caminhos. No primeiro, realiza-se a revisão da literatura recente com o intuito de identificar os principais estudos sobre interfaces para telas sensíveis ao toque, voltadas aos deficientes visuais. No segundo, a criação de protótipos, que desdobra-se em duas novas fases: a avaliação preditiva e a avaliação experimental.

Os protótipos do aplicativo estão sendo desenvolvidos com o software *Android Studio 1.3.2* e da película, construídos com papel contact transparente e lantejoulas, que simulam a sensação tática de botões e cliques. Os testes são realizados em um *smartphone* Motorola, Moto G, 2^a geração, com tela de 5" e dimensões 70,7mm x. 141,5mm, com sistema operacional *Android™ 5.0 Lollipop*.

A avaliação preditiva está sendo realizada por C1, 36 anos, com cegueira congênita, alfabetizado em braile, com formação superior e experiente no uso de computadores e celulares com softwares leitores de tela.

Durante esta fase, C1 realiza testes semanais verificando, dentre outras coisas, a melhor posição do aparelho em suas mãos, a localização dos botões, a sensação tática da película e o funcionamento do aplicativo como um todo, para analisar se está sendo proporcionada a “visão” da interface.

Estes testes têm como finalidade principal alcançar agilidade e eficiência na escrita de textos, para posteriormente, na avaliação experimental, serem testados por C2, que avaliarão a usabilidade dos protótipos.

V. RESULTADOS PARCIAS

Para início do desenvolvimento dos protótipos foi preciso compreender o método braile, para então, com as contribuições de C1 prototipar os 6 pontos necessários à escrita.

Antes do desenvolvimento do aplicativo, foi construída a versão inicial da película tática e a partir dela, os primeiros testes com C1 foram realizados. Conforme descrito no item Materiais e Métodos, utilizou-se lantejoulas e papel contact.

Esta versão considerou as experiências relatadas nos trabalhos relacionados, conforme Fig. 6a, onde os botões braile foram inseridos nas laterais e os botões de confirmação no centro, com as funções: espaço, confirmar letra, ouvir letra e ouvir o texto escrito. Na Figura 6b, a versão inicial do aplicativo.

Os testes foram realizados semanalmente, e a partir deles muitas adaptações, listadas a seguir, foram propostas, aperfeiçoando o modelo, principalmente no que tange a agilidade na escrita.

A Fig. 7 apresenta a versão atual do protótipo da interface. Esta versão considerou as modificações propostas por C1 e as considerações discutidas durante as reuniões no grupo de pesquisa.

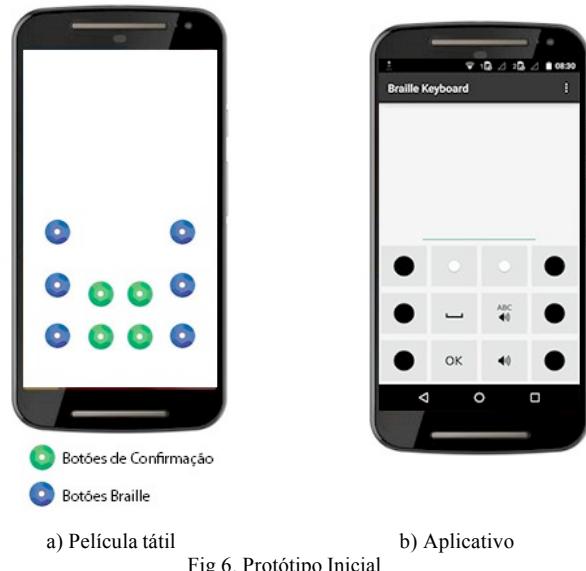
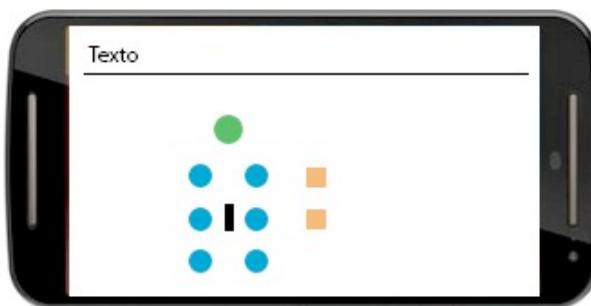


Fig 6. Protótipo Inicial

1. o botão de OK confirma a letra digitada e na sequência, emite feedback sonoro;
2. os botões do braille foram colocados na mesma posição do reglete e os botões de confirmação redistribuídos na tela;
3. na película tática, o botão OK é maior que os botões braille e os botões de confirmação são quadrados;
4. observada a maneira como a colaboradora segura o aparelho, a localização dos botões foi alterada para vertical;
5. para seleção dos botões, inicialmente era necessário uma pressão longa, o que foi alterado para toque;
6. os botões foram colocados mais próximos, para facilitar a locomoção na tela;
7. o mesmo botão realiza mais de uma tarefa, para assim deixar a interface mais leve. Exemplos: o botão de enviar mensagem emite feedback sonoro com a leitura de todo o texto e o botão para apagar, se pressionado 1 vez, exclui o caractere anterior, 2 vezes, a palavra, 3 vezes, o texto;
8. para inserção do espaço, basta clicar no botão OK sem pressionar nenhum botão braille;
9. para transitar entre as letras, gesto na tela;
10. foi inserido um relevo adicional para colaborar com a localização do usuário.



- Botões do braille
- Botão para confirmação das letras e espaço
- Relevo para localização
- Botões para apagar e enviar mensagem

Fig 7. Protótipo atual

VI. CONSIDERAÇÕES FINAIS

Apresentamos neste artigo os primeiros protótipos, aplicativo e película tática, para entrada de textos em smartphones com tela sensível ao toque, utilizando o Sistema Braille. Nas etapas iniciais contamos com a colaboração de uma deficiente visual congênita, que mune o projeto com informações relevantes sobre a vivência do público alvo.

Os estudos preliminares mostraram a importância da película tática na percepção do usuário, que o deixou mais seguro em relação às tarefas solicitadas.

Durante os testes foram observados: sensação tática de diferentes tipos e tamanhos de botões, distância entre eles, principais tarefas a serem implementadas, os momentos mais importantes para emissão do feedback sonoro, a pressão na qual os botões devem ser pressionados, dentre outros, sempre com a intenção de transformar a experiência do usuário agradável, ágil e eficiente.

Prosseguiremos a pesquisa com novos testes nos protótipos, ainda na avaliação preditiva, para posteriormente os avaliar em cenários reais, com um grupo de voluntários cegos, para que, no futuro, os protótipos possam ser implementados e beneficiem os deficientes visuais.

AGRADECIMENTOS

Agradecemos a colaboradora Lílian Cristina dos Santos, que participa do desenvolvimento dos protótipos, e ao grupo de pesquisa da Universidade Federal de Goiás. Este trabalho está sendo realizado com apoio financeiro dos editais 5, 6 e 8/2012 da Fundação de Apoio à Pesquisa do Estado de Goiás (FAPEG).

REFERENCIAS

- [1] BAPTISTA, J. A. L. S.. A Invenção do Braille e a sua Importância na Vida dos Cegos. Lisboa: Comissão Braille. 2000. Disponível em: <<http://www.gesta.org/braille/braille01.htm>>. Acesso em 02 out 2015
- [2] EL-GLALY, Yasmine N., QUEK, Francis, SMITH-JACKSON, Tonya e DHILLON, Gurjot. “Touch-screens are not tangible: fusing tangible interaction with touch glass in readers for the blind,” in Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction, 2013, pp. 245–253.
- [3] FREY, Brian, SOUTHERN, Caleb e ROMERO, Mario. BrailleTouch: Mobile Texting for the Visually Impaired . 2012. Georgia Institute of Technology School of Interactive Computing. Atlanta. GA.
- [4] IBGE. Instituto Brasileiro de Geografia e Estatística. Censo Demográfico 2010: Características gerais da população, religião e pessoas com deficiência. 2012. Disponível em: <http://www.ibge.gov.br/home/estatistica/populacao/censo2010/caracteristicas_religiao_deficiencia/caracteristicas_religiao_deficiencia_tab.xls>. Acesso em 30 set 2015.
- [5] OLIVEIRA, João, GUERREIRO, Tiago, NICOLAU, Hugo, JORGE, Joaquim e GONÇALVES, Daniel. BrailleType: Unleashing Braille over Touch Screen Mobile Phones. Technical University of Lisbon. Portugal. 2011.
- [6] OLIVEIRA, João, GUERREIRO, Tiago, NICOLAU, Hugo, JORGE, Joaquim e GONÇALVES, Daniel. Blind People and Mobile Touch-based Text-Entry: Acknowledging the Need for Different Flavors. Technical University of Lisbon. Portugal. 2011.
- [7] ROMERO, Mário, FREY, Brian, SOUTHERN, Caleb e ABOWD, Gregory D. “BrailleTouch: designing a mobile eyes-free soft keyboard,” in Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services, 2011, pp. 707–709.
- [8] WHO. World Health Organization. Visual Impairment and Blindness. 2014. Disponível em: <<http://www.who.int/mediacentre/factsheets/fs282/en/>>. Acesso em 30 set 2015.

Aspectos de Implementação de um Serviço de Voz em Aplicações de Realidade Aumentada Móvel

Implementation Aspects of a Voice Service in Mobile Augmented Reality Applications

Tiago Araújo*, Brunelli Miranda†, Carlos Santos*, Nikolas Carneiro*, Marissa Carvalho*, Gleidson Sousa†, Bianchi Serique*†, Nelson Neto*

*Programa de Pós-Graduação em Ciência da Computação - PPGCC

†Faculdade de Computação – FACOMP

Universidade Federal do Pará, Belém, Pará, 66075-110

Resumo – O uso de aplicações de realidade aumentada móvel tem aumentado nos últimos anos, permitindo a integração de vários tipos de interações que antes não eram tão difundidas. Existe atualmente uma grande quantidade de tecnologias, entretanto não há um padrão de desenvolvimento para aplicações que aceitam diversas interações. Este artigo propõe alguns aspectos de implementação com boas práticas de programação para aplicações de realidade aumentada móvel utilizando interação por voz e sistema Android. Para se chegar nesses aspectos, foi feito uma revisão técnica nas guias de desenvolvimento e API's das bibliotecas utilizadas na ferramenta. Foi condensado e apresentado boas práticas de implementação utilizando interação por voz em aplicações de realidade aumentada móvel.

Abstract – The use of mobile augmented reality applications has increased in recent years, enabling the integration of several types of interactions that were not as widespread. There is currently a lot of technology, but there isn't a pattern of development for applications that accept several interactions. This paper proposes some implementation aspects with good programming practices for mobile augmented reality applications using voice interaction and Android system. To reach these aspects, a technical review was made focusing in development guides and API's used in the tool library. It was condensed and presented good practical implementation using voice interaction of mobile augmented reality applications.

Keywords—realidade aumentada, voz, implementação, aplicação móvel, interação, Android, reconhecimento de voz

I. INTRODUÇÃO

A Realidade Aumentada Móvel (RAM) ganhou um importante destaque em múltiplas áreas de aplicação, dentre elas: engenharias, saúde, educação turismo, automação e outras. Este destaque está ligado a vários fatores, como: o desenvolvimento tecnológico dos dispositivos móveis, aumento da popularidade e comercialização de *smartphones* e *tablets*, e a inovação e facilidades da Realidade Aumentada (RA).

Apesar do desenvolvimento tecnológico e da popularidade crescente da RAM serem motivações para sua ampla utilização, são os estudos na usabilidade das ferramentas RAM que podem manter uma boa experiência de uso, consolidando assim a vida útil da utilização dessas ferramentas nas mais diversas áreas.

Martínez *et al.* [1] apresenta alguns desafios para o desenvolvimento de aplicações de RAM. Dentre os desafios apresentados, pode-se destacar a falta de padrões de desenvolvimento e o pouco espaço para a apresentação de informações.

A falta de padrões de desenvolvimento pode ser considerada tanto escassez de padrões para codificação dessas aplicações de RAM, assim como para a organização e *design* das interfaces gráficas. E o pouco espaço para apresentação das informações se dá pelo fato de que a RA precisa mostrar ao usuário o ambiente real acrescido de informações virtuais em sincronia [2], e isto, em uma tela de dispositivos móveis como um *smartphone*, que pode variar de 3 a 8 polegadas.

Inevitavelmente, o projetista terá que optar por deixar algumas informações escondidas (geralmente em menus que podem apresentar várias opções) infringindo alguns princípios de usabilidade, como a primeira heurística de Nielsen [3], em troca de uma melhor visibilidade do conteúdo principal e deixando a tela do usuário menos carregada. Porém, esconder funcionalidades em menus pode acarretar em um maior esforço cognitivo para encontrar e utilizar tais funcionalidades.

Um importante recurso que pode contribuir com a organização da interface gráfica no pouco espaço de tela é a utilização de *feedbacks* e comandos voz na qual o usuário pode interagir e receber informações utilizando a comunicação oral. Um modelo de desenvolvimento pronto para desenvolvedores pode ajudar a integração desse serviço em aplicações que antes só tinham interação por toques na tela.

Este trabalho apresenta um padrão de desenvolvimento de um serviço de voz voltado a aplicações RAM. Ele cobre a arquitetura geral desse tipo de serviço, vai refinando em um

modelo de implementação para sistemas Android e encerra com os aspectos que devem ser levados em consideração no momento de implementar esse serviço.

II. INTERAÇÃO POR VOZ

A voz é um dos meios de comunicação mais utilizados pelos seres humanos, estando entre as três formas de comunicação (voz, gestos e expressões faciais) mais utilizadas no dia-a-dia de uma pessoa [6]. A voz, quando utilizada em uma aplicação, está dentro do conceito de linguagem natural.

A interação por linguagem natural pode ser definida como a comunicação entre humano e máquina utilizando uma linguagem familiar ao humano [7]. A área de Interação Humano-Computador (IHC) atribui papel importante para esse tipo de interação, devido aos benefícios oferecidos ao usuário, entre eles, o fato das ações serem mais intuitivas, minimizando assim o esforço cognitivo e permitindo que o usuário se concentre na tarefa e não na interação em si [8]. A interação por voz também pode facilitar a busca do usuário por uma funcionalidade, ou até mesmo servir como atalho para a execução da funcionalidade em questão [9].

Apesar da utilização de comandos por voz poder variar de acordo com o tipo de aplicação, de maneira geral, o reconhecimento de voz nas aplicações segue um modelo padrão, que pode ser observado na figura 1.

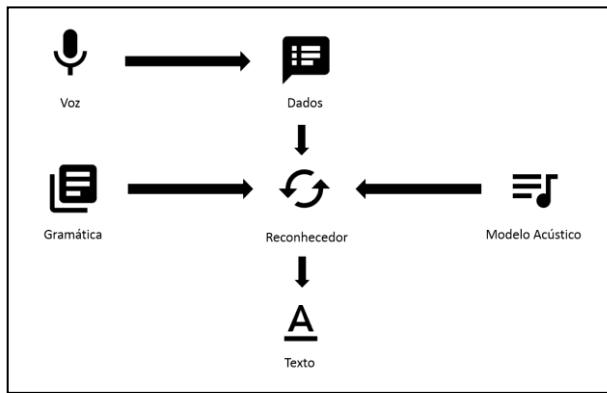


Fig 1. Funcionamento padrão de um modelo de reconhecimento de voz

No modelo apresentado, o usuário entra com a informação em forma de comando de voz, que é repassada para o reconhecedor, onde ela é analisada juntamente com a gramática e modelo acústico empregados na aplicação. O resultado desse processo é então retornado ao usuário em forma textual. Esse modelo geral exemplifica aplicações de interação por voz encontrados em vários sistemas de reconhecimento de voz [10].

III. ARQUITETURA

Um serviço de voz precisa de alguns elementos chaves para poder realizar a síntese e reconhecimento de voz. O processo visto na figura 1 mostrou o modelo geral independente da aplicação utilizada. Para aplicações de realidade aumentada móvel esse modelo se encaixa com

algumas mudanças, relacionadas aos componentes de realidade aumentada.

Uma arquitetura para aplicações de RAM que utilizam serviços de voz pode ser utilizada para definir ações de implementação para os principais componentes desse tipo de aplicação. Os componentes utilizados são baseados em componentes de um dispositivo Android, que já contém alguns elementos prontos para reconhecimento e síntese de voz.

Estes componentes foram divididos em dois nós principais, relacionados aos dois principais módulos para esse modelo, serviço de voz e a Aplicação de RAM. Os componentes do serviço de voz são um reconhecedor, um sintetizador e as gramáticas. Os componentes de uma aplicação RA foram divididos em quatro, os trabalhos de [11], [12], [13], [14] e [15] apresentam componentes semelhantes quando se foca em desenvolvimento. Foi definido um padrão dentro dessas aplicações, definidos pela sua função dentro do seu escopo:

- Dados: Dados que a aplicação utiliza para gerar conteúdo para o usuário.
- Renderizador: Motor gráfico para gerar modelos de realidade aumentada na tela da aplicação.
- Reconhecimento: Conjunto de algoritmos ou ferramentas utilizados pela aplicação para gerar interação com usuário, como reconhecimento de imagens, posição georeferenciada, utilização de sensores.
- Coordenador: O componente que liga o Renderizador e o Reconhecimento para gerar o ambiente de realidade aumentada para o usuário.

A figura 2 mostra um diagrama desses componentes para a obtenção deste serviço de voz para aplicações de RAM.

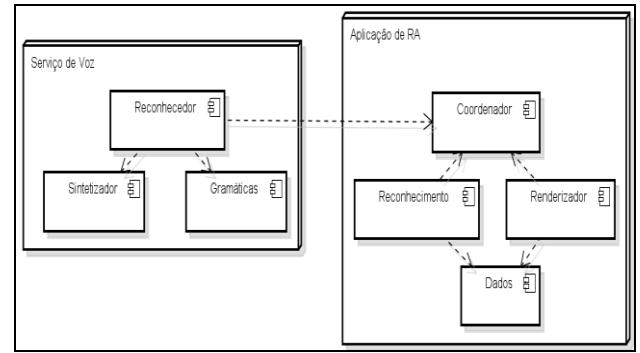


Fig 2. Diagrama de componentes de um serviço de voz em aplicações RAM

A ligação entre o reconhecedor do serviço de voz e do coordenador da aplicação de RAM é importante para manter a sincronização das interações. Um comando que tem sua ação muito atrasada ou uma ação da aplicação que seja inapropriada para um contexto podem dificultar a aceitação do serviço pelo usuário[5].

O diagrama de componentes da figura 3 mostra o conceito geral da aplicação, detalhes mais específicos podem ser vistos

na figura 2, que mostra um diagrama de classes para o serviço de voz em uma aplicação de RAM.

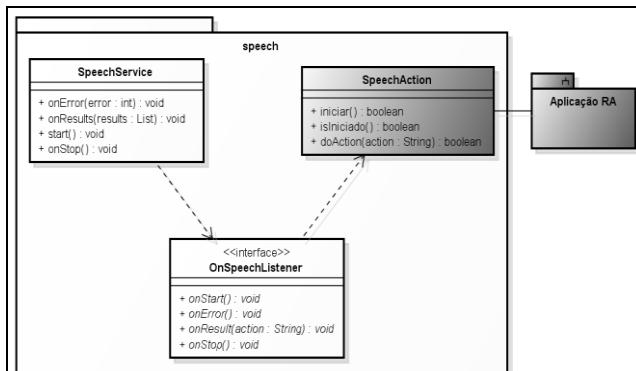


Fig 3. Diagrama de classes de um serviço de voz em aplicações RAM

A especificação de cada classe do diagrama e suas relações pode ser vista abaixo:

1) **SpeechService**: Essa classe é responsável pela comunicação realizada com o sintetizador online. Deve cercar os eventos para enviar e receber os resultados do sintetizador. Os métodos dessa classe são explicados abaixo:

- **start**: Esse método inicia o reconhecimento do serviço de voz ligando o microfone. É encerrado assim que percebe silêncio por parte da entrada do usuário.
- **onStop**: Método que é chamado quando é percebido silêncio, para que a entrada de áudio puro seja encerrada, permitindo manipulação por parte da classe. Esse é o momento de enviar os resultados para o sintetizador.
- **onResults(results : List)**: Nesse método deve-se encaminhar o conjunto de palavras para o reconhecedor poder realizar as ações necessárias fazendo comparação com a gramática.
- **onError(error : int)**: Esse método é chamado em caso de qualquer erro no processo de envio e recebimento do processo de síntese de voz. Erros que podem acontecer estão relacionados a erros de rede, áudio não ser compreendido como voz, não entendimento do que foi dito e falta de resposta do sintetizador.

2) **OnSpeechListener**: É uma interface para ligar os eventos dependentes do sintetizador que ocorrem no serviço de voz. Os métodos são chamados em `SpeechService` e suas chamadas manipuladas em `SpeechAction`.

3) **SpeechAction**: É a classe que implementa os serviços do reconhecedor. Ela recebe por meio da interface que implementa os resultados da síntese de voz e aplica as ações necessárias relacionadas a gramática. O método `doAction` recebe uma palavra da gramática utilizada pela aplicação e realiza uma ação ou cancela a mesma mediante definição já feita.

Essa arquitetura mostra um conjunto básico de classes que devem ser implementados e como devem ser relacionados

com componentes de uma aplicação de RAM. Serviços já implementados no sistema Android podem ser utilizados, como o acesso ao microfone e acesso ao que foi gravado.

IV. SERVIÇO DE VOZ

A ligação entre os componentes de uma aplicação de RAM com reconhecimento de voz pode ser feita através de serviços já presentes no sistema Android, o que garante o desempenho otimizado da aplicação, já que usa chamadas nativas do próprio sistema operacional.

A figura 4, de autoria própria, mostra o fluxo do serviço de voz implementado utilizando o Google Now [16], o serviço de síntese e reconhecimento de voz padrão do Android.

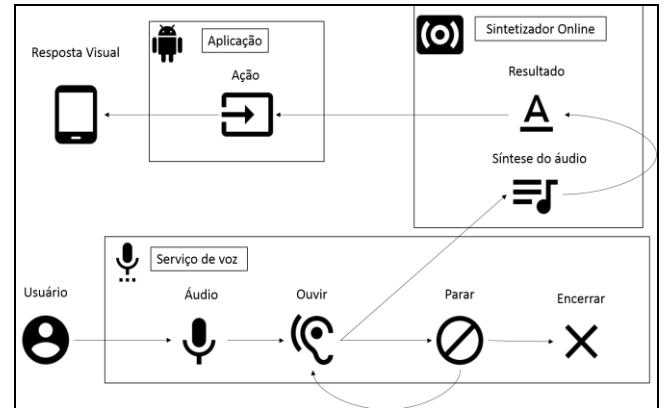


Fig 4. Fluxo de troca de mensagens na implementação do serviço de voz

O serviço de voz implementado possui alguns estados para garantir a eficiência, acurácia e corretude do reconhecedor. Os estados e funções do reconhecedor são listados a seguir:

1) **Ouvir**: Esse estado liga o microfone do dispositivo e espera a entrada de áudio do usuário. Nesse estado o áudio puro é convertido em um formato de dados compatível com o o reconhecido pelo sintetizador. Assim que ele detecta silêncio, o reconhecedor muda de estado para o estado de Parar.

2) **Parar**: Nesse estado é necessário parar o reconhecimento da entrada do usuário, e enviar os dados para o sintetizador. Abrir uma conexão e esperar o sinal positivo de entrega são as etapas para esse envio. De acordo com a ação realizada, pode-se passar pro estado de encerrar, ou voltar para o estado de Ouvir.

3) **Encerrar**: Esse estado finaliza o serviço de reconhecimento. Faz as verificações necessárias para liberar o microfone para o sistema, e encerra qualquer conexão ainda aberta com a internet usada em função do reconhecedor.

V. ASPECTOS DE IMPLEMENTAÇÃO

Os aspectos podem ser divididos em duas partes principais: relacionadas a aplicação e relacionadas a tecnologia. Aspectos de aplicação são as medidas que devem ser tomadas para que o serviço de voz seja integrado a aplicação de RA. A figura 5 mostra o processo de aplicação

das diretrizes na aplicação de RA. As diretrizes que devem ser seguidas para implementação do serviço estão listadas abaixo:

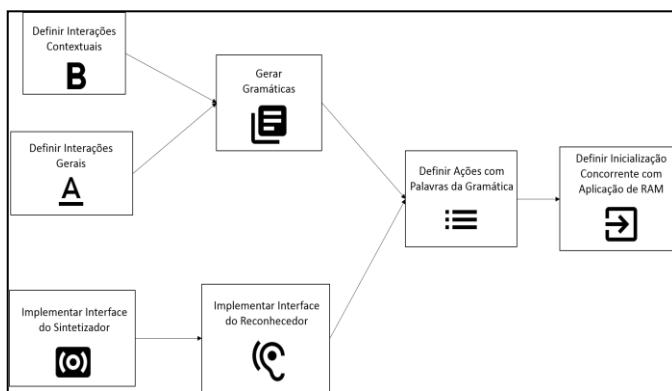


Fig 5. Aspectos gerais de implementação

1) Definir Interações Contextuais: Definir comandos contextuais é necessário para que os comandos que são dependentes de certo contexto sejam executados somente nesse contexto. Exemplos de contexto podem ser a visualização de certo dado, a navegação em um mapa, a ação diante de um objeto virtual. As interações em que se deseja ter interação devem ser mapeadas em um conjunto de palavras e deve ser associada a uma condição de contexto presente na aplicação.

2) Definir Interações Gerais: Comandos gerais são necessários para ações que são gerais em mais de um contexto da aplicação. São comandos gerais ações de menu que podem ser acessadas em qualquer momento, comandos padrões da própria tecnologia utilizada, como voltar, página principal (home) e configurações. Essas interações devem ser mapeadas para que o contexto não seja aplicado a elas, o conjunto de palavras não pode ser concorrente com nenhuma palavra de nenhum contexto.

3) Gerar Gramáticas: Definidas as interações que serão possíveis em contextos e de forma geral, o conjunto dessas palavras deve ser usado para geração de uma gramática que será usada para as ações da aplicação. Essa gramática deve ser organizada de modo que o reconhecedor possa utilizar essas palavras para implementar essas ações, que depende do formato que o sintetizador e o reconhecedor podem processar.

As diretrizes que devem ser seguidas para tecnologia estão listadas abaixo:

1) Implementar Interface do Sintetizador: O sintetizador pode estar integrado ou não ao telefone, mas é necessário ter uma interface de acesso a ele, ao microfone, ou ambos. Essa interface deve ser criada para processamento ou direcionamento da voz do usuário. Ela deve ter métodos para retorno do resultado do que foi captado para a interface do reconhecedor.

2) Implementar Interface do Reconhecedor: Essa interface deve ser implementada para direcionar o resultado do sintetizador em alguma ação da aplicação. Deve ter métodos para esperar os resultado do sintetizador e métodos para cercar

os estados possíveis do reconhecimento, como ouvindo, começando a ouvir, parando de ouvir e iniciando síntese.

3) Definir Ações com Palavras da Gramática: As ações devem ser associadas no reconhecedor, com as palavras que foram definidas na gramática. Essas ações devem podem ser as gerais ou de contexto, desde que sejam as mesmas acessíveis através da interface gráfica. Mais de uma palavra pode acionar a mesma ação, desde que esteja mapeada na gramática.

4) Definir Inicialização Concorrente com Aplicação de RAM: O passo final para implementação desse serviço é definir como o serviço vai ser iniciado. Um botão que aciona o serviço é usado em algumas aplicações de interação por voz [17], mas pode ser definido do modo que for necessário para a aplicação. Essa implementação é altamente dependente da tecnologia, já que envolve paralelismo de execução e interação, e sistemas operacionais móveis implementam de forma diferenciada acesso a ferramentas como microfone e execução do modo de ouvir dos mesmo.

Esses aspectos são relevantes no desenvolvimento de qualquer serviço de voz em uma aplicação RAM. O exemplo de reconhecedor de voz utilizado é padrão de sistemas Android, sendo assim esse modelo pode ser replicado com fidelidade nesses sistemas.

CONSIDERAÇÕES FINAIS

Esse trabalho propõe uma série de aspectos de implementação que devem ser levados em consideração para que serviços de voz possam ser integrados com sucesso em aplicações de RAM.

Esses aspectos foram usados como base para o desenvolvimento do serviço de voz de uma aplicação de RAM, que pode ser vista na figura 6. Um breve conjunto de palavras usadas para as gramáticas está presente também, divididas em seu determinado contexto.



Fig 6. Exemplo de aplicação e palavras definidas nas gramáticas

Foi definido com uma base em sistema Android, então certas decisões feitas em relação a tecnologia foram feitas escolhendo o que era mais eficaz dentro desse contexto. Futuros trabalhos devem cobrir exceções de outros sistemas.

Os aspectos apresentados procuram facilitar o desenvolvimento de aplicações RAM que tem como objetivo integrar um serviço de voz. Uma arquitetura e diagramas foram apresentados para suportar esse modelo.

A aderência a esses aspectos deve facilitar que aplicações de RAM possam melhorar suas interações e que os usuários se sintam mais à vontade para escolher variadas interações em seus ambientes de RAM.

REFERÊNCIAS

- [1] H. Martínez, D. Skournetou, J. Hyppola, S. Laukkonen, and A. Heikkilä, "Drivers and bottlenecks in the adoption of augmented reality applications," *Journal of Multimedia Theory and Application*, 2014.
- [2] Pulli, P.; Pyysalo, T.; Metsavainio, J.-P.; Komulainen, O., "CyPhone-experimenting mobile real-time telepresence," in *Real-Time Systems*, 1998. Proceedings. 10th Euromicro Workshop on , vol., no., pp.10-17, 17-19 Junho 1998
- [3] NIELSEN, J. "Heuristic Evaluation". Em: NIELSEN, J.; MACK, R. L. Usability Inspection Methods. New York, NY: Katherine Schowalter, 1994. Cap. 2
- [4] Cawood, S., Fiala, M., "About Augmented Reality", *Augmented Reality: A Practical Guide*. Pragmatic Bookshelf, USA 2008, cap 1.
- [5] T. Olsson and M. Salo. "Online User Survey on Current Mobile Augmented Reality Applications". *Proceedings of the 2011 10th IEEE International Symposium on Mixed and Augmented Reality*. Washington: IEE Computer Society. 2011. pp 75-84.
- [6] A. Teixeira et al. "Speech-Centric Multimodal Interaction for Easy-To-Access Online Services – A Personal Life Assistant for the Elderly". *5th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion – DSAI*, 2013.
- [7] B. Shneiderman. "Designing the User Interface: Strategies for Effective Human-Computer Interaction", 3rd Edition, pp 293-295.1998.
- [8] N. Vidakis, M. Syntychakis, G. Triantafyllidis and D. Akoumianakis. "Multimodal Natural User Interaction for Multiple Applications: The Gesture – Voice Example". *International Conference on Telecommunications and Multimedia – TEMU*, 2012.
- [9] L. Xia, K. Kai, W. Xiaochun and W. Dan. Research and Design of the "Voice-Touch-Vision" Multimodal Integrated Voice Interaction in the Mobile Phone, 2010.
- [10] Dolezal, J.; Kencl, L., "A unifying architecture for easy development, deployment and management of voice-driven mobile applications," in *Network and Service Management (CNSM)*, 2011 7th International Conference on , vol., no., pp.1-5, 24-28 Oct. 2011
- [11] Tang, L.Z.W.; Kian Sin Ang; Amirul, M.; Bin Mohamed Yusoff, M.; Chee Keong Tng; Bin Mohamed Alyas, M.D.; Joo Ghee Lim; Kyaw, P.K.; Folianto, F., "Augmented reality control home (ARCH) for disabled and elderly". Em *Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, 2015 IEEE Tenth International Conference on , vol., no., pp.1-2, 7-9 April 2015
- [12] Vullamparthi, A.J.; Nelaturu, S.C.B.; Mallaya, D.D.; Chandrasekhar, S., "Assistive Learning for Children with Autism Using Augmented Reality". Em *Technology for Education (T4E)*, 2013 IEEE Fifth International Conference on , vol., no., pp.43-46, 18-20 Dec. 2013
- [13] Buti Al Delail, Luis Weruaga, and M. Jamal Zemerly. 2012. "CAViAR: Context Aware Visual Indoor Augmented Reality for a University Campus". Em *Proceedings of the The 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology - Volume 03 (WI-IAT '12)*, Vol. 3. IEEE Computer Society, Washington, DC, USA, 286-290.
- [14] Loris D'Antoni, Alan Dunn, Suman Jana, Tadayoshi Kohno, Benjamin Livshits, David Molnar, Alexander Moshchuk, Eyal Ofek, Franziska Roesner, Scott Saponas, Margus Veane, and Helen J. Wang. 2013. Operating system support for augmented reality applications. In *Proceedings of the 14th USENIX conference on Hot Topics in Operating Systems (HotOS'13)*.
- [15] C. Santos, N. Carneiro, B. Miranda, B. Serique. "Uma Aplicação de Realidade Aumentada Móvel para Ambientes Indoor e Outdoor". XV Workshop de Realidade Virtual e Aumentada, Marília – SP, pp 120-126, November 2014.
- [16] Google Inc., "Google Now", <https://www.google.com/landing/now/>. Dezembro 2014.
- [17] Buti Al Delail, Luis Weruaga, and M. Jamal Zemerly. 2012. "CAViAR: Context Aware Visual Indoor Augmented Reality for a University Campus". Em *Proceedings of the The 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology - Volume 03 (WI-IAT '12)*, Vol. 3. IEEE Computer Society, Washington, DC, USA, 286-290.

Aplicando *Visual Domain Specific Languages* para Definição de Sistemas Multi-Agentes em Aplicações de Realidade Virtual

Appling Visual Domain Specific Languages for Definition of Multi-Agent Systems in Virtual Reality Applications

Daniel Abella C. M. de Souza; Fernando da Fonseca de Souza

Centro de Informática (Cin)
Universidade Federal de Pernambuco (UFPE)
Recife, Brasil
{dacms,fdfd}@cin.ufpe.br

Raphael Mendonça da Nóbrega; Fabrício da Costa Dias; Pacelly J. C. A. Xavier

Departamento de Graduação
Centro de Ensino Superior e Desenvolvimento (CESED)
Campina Grande, Brasil
{rmendonca,fdias,pxavier}@facisa.edu.br

Resumo — Com intuito de amparar o desenvolvimento de ambientes virtuais de aprendizagem, este artigo apresenta *Agent Modeling Tool (AMT)*, uma ferramenta que possibilita a modelagem visual de agentes inteligentes em ambientes de Realidade Virtual, abstraindo a complexidade no desenvolvimento de sistemas multi-agentes aplicados a estes ambientes. Por meio desta ferramenta, tem-se a possibilidade de modelar agentes, seus respectivos comportamentos, fluxo de mensagens, entre outras características definidas pelo padrão *Foundation for Intelligent Physical Agents (FIPA)*.

Palavras-Chave —realidade virtual; agentes inteligentes; linguagens de domínio específico

Abstract—In order to support the development of virtual learning environments, this paper presents Agent Modeling Tool (AMT), a tool that provides a visual modeling of intelligent agents in Virtual Reality context, reducing the complexity of developing multi-agent systems that can be applied to these learning environments. Using this tool is possible to model agents, their behavior, messages flow and other features standardized by Foundation for Intelligent Physical Agents (FIPA).

Keywords— virtual reality; intelligent agent; domain specific languages

I. INTRODUÇÃO

Agentes Inteligentes (AI) estão sendo integrados a ambientes de Educação à Distância [1][2], especialmente em aplicações de *e-Learning*, visando tornar o processo de aprendizado mais adaptável. Neste sentido, os agentes podem também ser aplicados para realizar uma avaliação contínua do usuário.

Segundo Oliveira [3], por meio da utilização de agentes inteligentes em aplicações de *e-Learning* e Realidade Virtual (RV), torna-se possível a adaptação do conteúdo apresentado

ao usuário, baseado no seu conhecimento e em suas respectivas habilidades. Além disto, o mesmo autor indica que isto atuará como um fator revolucionário no mundo educacional, tornando o processo de aprendizado mais dinâmico.

De maneira a abstrair a complexidade no desenvolvimento de Software, tem-se a aplicação da técnica de reuso *Domain-Specific Visual Language (DSVL)*, que é uma variante da Domain Specific Language (DSL), que permite o rápido desenvolvimento do comportamento e/ou estrutura das aplicações em domínios bem definidos através de notações visuais [4].

No contexto de desenvolvimento de sistemas multi-agentes, a aplicação de DSVL possibilitará a definição dos agentes e suas interações de maneira visual.

Atuando neste sentido, apresentamos neste artigo uma ferramenta para modelagem visual dos Agentes Inteligentes, de maneira a compor um Sistema Multi-Agentes (SMA), que são utilizados em aplicações que usam RV.

O artigo está organizado da seguinte maneira. Na seção II temos o referencial teórico. Na sequência, as seções III e IV, apresentam respectivamente a ferramenta e sua aplicação no desenvolvimento da ferramenta de RV Honey. Por fim, na seção V, temos a conclusão do trabalho.

II. REFERENCIAL TEÓRICO

Este trabalho envolve pesquisa nas áreas de RV, Inteligência Artificial (IA) e Reuso de Software (RS) que são descritos nesta seção.

A. Realidade Virtual

A área de RV é uma das subáreas da Computação Gráfica que permite a criação de ambientes gerados por computador que lidam com características como imersão onde a realidade é representada [5] [6].

Segundo Braga [7], o termo RV surgiu nos anos 80 por Jaron Lamier, e se refere a uma técnica avançada de interface, onde é prevista a interação do usuário através de objetos tridimensionais. Complementarmente, o mesmo autor explicita que a RV é composta de três idéias básicas: Interação, Envolvimento e Imersão.

Em Ambientes Virtuais (AV), devido à inserção de ambientes tridimensionais que representam à realidade, assim como a associação de atributos como Interação, Envolvimento e Imersão, estes ambientes são principalmente aplicados em sistemas de entretenimento e educação. Entretanto, a sua aplicabilidade não se restringe aos citados anteriormente, podendo ser aplicadas em aplicações de simulação, por exemplo.

A necessidade de novos métodos de aprendizagem é crescente, uma vez que os métodos tradicionais de ensino carecem de meios que possibilitem a diferenciação da capacidade de aprendizado individual, evitando o tratamento das suas necessidades de maneira genérica.

Segundo Braga [7], a RV não deve mais ser tratada mais como uma alternativa para ampliar a aprendizagem, mas como uma maneira de suprir um novo instrumento de ensino e aprendizado.

B. Agentes Inteligentes e Sistemas Multi-Agentes (SMA)

Nos últimos anos, AI se tornaram um instrumento eficaz em sistemas de busca, automação de sistemas complexos e jogos. Na literatura, segundo Wooldridge e Jenninds [8], não há um conceito universalmente aceito para a definição de agentes em Computação.

No entanto, na área acadêmica, uma definição aceita para agentes classifica-os como tudo o que pode ser considerado capaz de perceber seu ambiente por meio de sensores e agir por intermédio de atuadores [9].

Em casos de alta complexidade, apenas um agente pode não ser o suficiente. Ao partilhar a responsabilidade entre agentes, de maneira que estes atuem de maneira colaborativa, temos um SMA.

Nos SMA's os agentes são o foco principal, tendo uma responsabilidade específica no contexto da aplicação, onde o resultado da atividade é fruto da cooperação e troca de informações entre os agentes que estão envolvidos para a resolução de dada tarefa.

C. Domain Specific Languages (DSL)

Entre as técnicas de programação, o RS é uma das menos evidenciadas, visto a sua complexidade e ao baixo conhecimento do corpo técnico a respeito do assunto. Com a utilização desta técnica, para um código desenvolvido, testado e disponibilizado em ambiente de produção, este poderia ser reutilizado em outros cenários similares.

No ano de 1992, Charles Krueger relatou um cenário que as empresas buscavam melhorias em seu processo de criação de software, principalmente no que diz respeito a qualidade e produtividade. Neste contexto, Krueger [10][11] define RS

como o processo de criar novos aplicativos a partir de já existentes ao invés de construí-los desde o início.

Nesta área, existem diversas abordagens, entre elas o DBC [12] e a DSL [13].

De maneira a abstrair a complexidade no desenvolvimento, tem-se a aplicação da técnica de reuso *Domain-Specific Visual Language*, que é uma variante da DSL descrita neste trabalho que permite o rápido desenvolvimento do comportamento e/ou estrutura das aplicações em domínios bem definidos, através de notação visual [4].

Na próxima seção apresentaremos a ferramenta que utiliza os conceitos supracitados.

III. APRESENTAÇÃO DA FERRAMENTA

A ferramenta *Agent Modeling Tool (AMT)* possibilita a modelagem visual de agentes inteligentes, dentro de um contexto de ambientes de RV.

Apresentamos a arquitetura da ferramenta AMT na Figura 1, enquanto na Figura 2, apresentamos a interface do sistema, onde temos a modelagem de dois agentes e o código gerado pela ferramenta dos dois agentes.

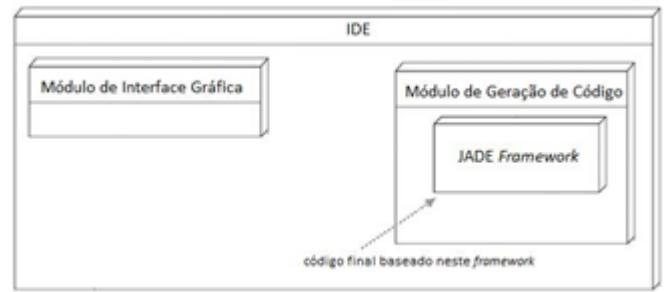


Figura 1: Arquitetura do AMT

A arquitetura é composta inicialmente pelo módulo de interface gráfica, na qual possibilita ao usuário modelar agentes inteligentes de maneira similar a diagramas de sequência da linguagem *Unified Model Language (UML)*.

Após a modelagem, o módulo de geração de código é chamado, onde temos a geração do código final executável dos agentes, escrito na linguagem Java, utilizando-se do framework *Java Agent Development Framework (JADE)* [14].

O framework JADE [14] adere as especificações da *Foundation for Intelligent Physical Agents (FIPA)* [15], que é uma organização para desenvolvimento e estabelecimento de padrões de software para agentes.

A integração de ferramentas de RV com o código gerado pelo AMT se dá através da implementação de uma interface definida pelo AMT, que permite assim a interação com os ambientes de RV.

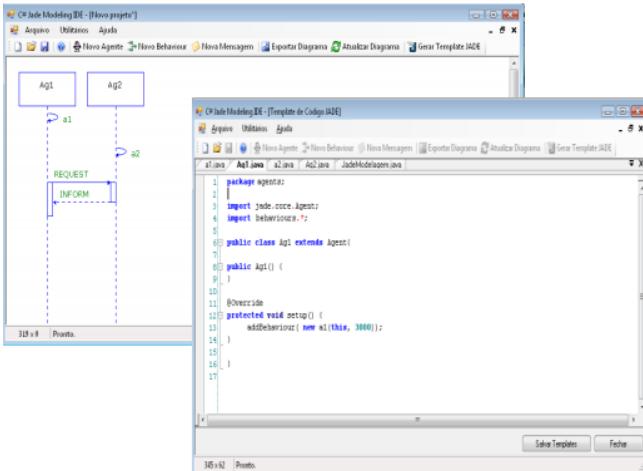


Figura 2: Interface do AMT

Na próxima seção apresentaremos brevemente ferramentas que atuam na mesma área.

A. Trabalhos Relacionados

Para a modelagem visual de aplicações baseadas em Agentes, temos uma série de fatores que ampliam a sua complexidade como, por exemplo, o fluxo de controle de mensagens entre os agentes, assim como a migração em um sistema multi-agentes.

A ferramenta ITM [16], apresentada na Figura 3, permite aos leigos em programação, a definição de agentes móveis utilizando-se da notação de diagrama de sequência. Em resumo, a aplicação é estritamente restrita ao ensino, uma vez que não é possível gerar um modelo executável dos agentes definidos visualmente.

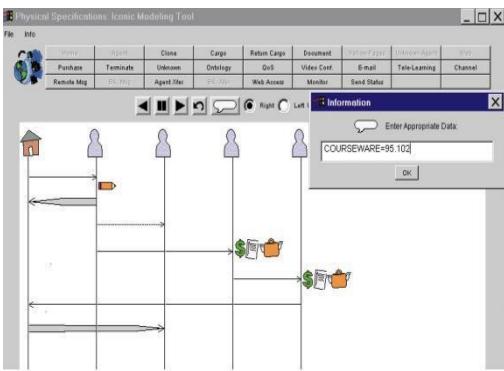


Figura 3: Ferramenta ITM

Outras iniciativas na mesma linha de pesquisa propõem a utilização da linguagem UML para a modelagem visual de

agentes, porém possuem a mesma necessidade da apontada na ferramenta ITM.

Neste sentido, a ferramenta proposta neste artigo permite além da modelagem, a geração do código fonte e a execução do SMA em um contexto de aplicações de RV.

IV. APlicando o AMT PARA O DESENVOLVIMENTO DA FERRAMENTA HONEY

Honey é um ambiente virtual de aprendizagem aplicado ao ensino de Engenharia de *Software*, conforme apresentado na Figura 4, na qual apresentamos a técnica *Stand up Meeting*, da metodologia *Extreme Programming (XP)* [17].

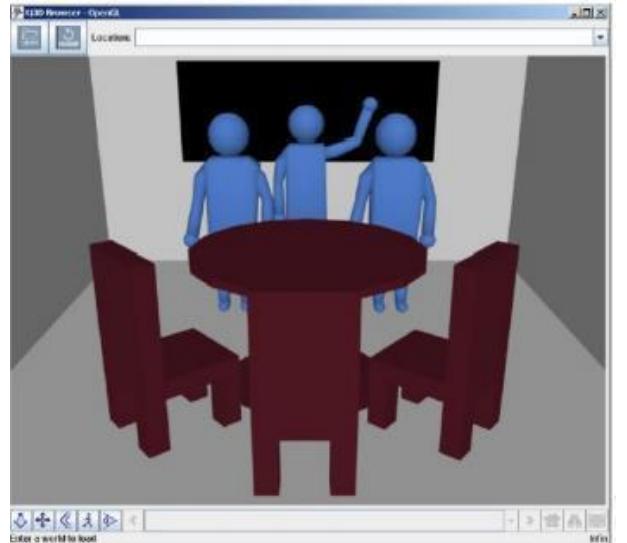


Figura 4: Demonstração da prática *Stand up meeting*.

O *browser* deste ambiente foi construído utilizando-se o framework Xj3d [18], que permite a visualização de ambientes tridimensionais descritos na linguagem X3D [19].

Para o gerenciamento deste ambiente foram definidos dois agentes inteligentes, o Agente Ambiente (AA) e o Agente Gerente (AG), que foram modelados na ferramenta AMT e estes agentes interagem diretamente com o browser.

V. CONCLUSÃO

Com a realização deste trabalho, pode-se concluir que por meio da aplicação das tecnologias em RV, IA e RS, podemos produzir um modelo para modelagem visual de agentes inteligentes, facilitando desta maneira a implementação de Ambientes de RV.

Para trabalhos futuros, sugere-se o desenvolvimento de um repositório público para sistemas multi-agentes desenvolvidos previamente, permitindo a análise da utilização deste modelo.

REFERÊNCIAS

- [1] Soula, G.; Pagesy, R.; Giorgi, R.; Fieschi, D.; Gouvernet, J.; Daniel, L.; Fieschi, M. An adaptive medical e-learning environment: the MEDIDACTE project. *Studies in health technology and informatics*, (2), 1076-1080, 2001.
- [2] Aquino, M. S.; Souza, F. F.; Frery, A. C.; Souza, D. A. C. M.; Fujioka, R. C. (2007), Supporting Adaptive Virtual Environments with Intelligent Agents. 7th International Conference on Intelligent Systems Design and Applications 2007, October.
- [3] Oliveira, Carlos DC, Marcos E. Cintra, and Francisco Milton Mendes Neto. "Jogo sério para o ensino da Gestão de Riscos em Projetos de Softwares usando Inteligência Artificial." *RENOTE* 11.1 (2013).
- [4] Sprinkle, J.; Karsai, G. A domain-specific visual language for domain model evolution. *Journal of Visual Languages & Computing*, pp.291-307, 2004.
- [5] Machado, L. S. A Realidade Virtual no Modelamento e Simulação de Procedimentos Invasivos em Oncologia Pediátrica: Um Estudo de Caso no Transplante de Medula Óssea. Tese (Tese de Doutorado) São Carlos: Pós-Graduação em Ciência da Computação da Escola Politécnica da Universidade de São Paulo – USP, 2003.
- [6] Kirner C., Siscoutto R., Fundamentos de Realidade Virtual e Aumentada . Em: Realidade Virtual e Aumentada: Conceitos, Projeto e Aplicações. Eds. Cláudio Kirner e Robson Siscoutto. Livro do Pré-Simpósio, IX Symposium on Virtual and Augmented Reality , Petrópolis – RJ, 2007.
- [7] Braga M. Realidade Virtual e Educação. *Revista de Biologia e Ciências da Terra*, nº 1, 2001.
- [8] Jennings, Nicholas R., Wooldridge, Michael. Software Agent. IEE Review. USA, v. 42, n. 1, p. 17-21, Jan. 1997.
- [9] Russel,S.J e Novig, P. (1995), Artificial Intelligence – A Modern Approach. Prentice-Hall
- [10] Maras, Josip, Maja Štula, and Ivica Crnković. "Towards specifying pragmatic software reuse." *Proceedings of the 2015 European Conference on Software Architecture Workshops*. ACM, 2015.
- [11] Krueger, Charles W. "Software reuse." *ACM Computing Surveys (CSUR)* 24.2 (1992): 131-183.
- [12] Sametinger, J. *Software Engineering with Reusable Components*. Springer-Verlag, 1997.
- [13] Fowler, M. *Domain-specific languages*. Pearson Education, 2010
- [14] SILVEIRA, Ricardo Azambuja. Modelagem orientada a agentes aplicada a ambientes inteligentes distribuídos de ensino: JADE: Java Agent framework for Distance learning Environments. 2001.
- [15] FIPA Specifications. <http://fipa.org/specifications/index.html>. Último acesso em janeiro de 2008.
- [16] B. Falchuk, A. Karmouch, “Visual Modeling for Agent-Based Applications”, 2013.
- [17] Beck, K. *Extreme Programming Explained: Embracing Change*, Boston, Addison-Wesley. 2000.
- [18] Xj3D - Java based X3D Toolkit and X3D Browser. <http://www.web3d.org/x3d/xj3d>. Último acesso em abril de 2015.
- [19] Extensible 3D (X3D). <http://www.web3d.org>. Ultimo acesso em outubro de 2015

Persistência e Distribuição de Dados em Ambientes Virtuais Distribuídos

Persistence and Data Distribution in Distributed Virtual Environments

Ronaldo Celso Messias Correia, Pedro Henrique Reis, Rogério Eduardo Garcia, Celso Olivete Júnior

Departamento de Matemática e Computação – Faculdade de Ciências e Tecnologia

UNESP - Universidade Estadual Paulista “Júlio de Mesquita Filho”

Presidente Prudente, SP - Brasil

ronaldo@fct.unesp.br, pedro.hreis22@gmail.com, {rogerio, olivete}@fct.unesp.br

Resumo—A persistência e distribuição de dados são fatores a serem considerados no projeto e implementação de Ambientes Virtuais Distribuídos (AVDs). A incompatibilidade entre sistemas de banco de dados relacional e objetos de linguagens orientadas a objetos cria a necessidade de uma forma de mapear (seja para persistir ou recuperar) os atributos do objeto criado pela aplicação, para os atributos da base de dados. Em uma aplicação distribuída é desejável que seja possível a um objeto invocar métodos de um objeto remoto que ofereça um serviço desejado, e cabe ao ambiente virtual viabilizar isto. Neste trabalho apresentamos um framework para a persistência e distribuição de dados em AVDs. Para avaliá-lo, foram propostos estudos de caso por meio de aplicações tridimensionais, executadas através de uma arquitetura distribuída. Com a utilização do framework, diminuiu-se o overhead sobre as bases de dados e obteve-se boa eficiência no mapeamento dos atributos do objeto para a base de dados. Em relação a distribuição de objetos, tem-se uma alternativa eficiente, já que serviços podem ser executados remotamente apenas com a implementação de uma interface nativa.

palavras-chave: Ambientes virtuais distribuídos, persistência de dados, distribuição de dados

Abstract—The persistence and data distribution are factors to be considered in the design and implementation of Distributed Virtual Environments (DVEs). The incompatibility between relational database systems and objects of object-oriented languages creates a need for a way to map (whether to persist or retrieve) the attributes of the object created by the application, to the attributes of database. In a distributed application it is desirable to be able to an object invoke methods of a remote object that offers a desired service, and it is up to the virtual environment enable this. In this work we present a framework for persistence and distribution of data in VDEs. To evaluate it, case studies were proposed by two three-dimensional applications, implemented through a distributed architecture. Using the framework, decreases the overhead on the database and obtained a good efficiency in mapping the attributes of the object for database. Regarding the distribution of objects, it has been an efficient alternative, since services can be remotely executed only by implementing a native interface.

keywords: Distributed virtual environments, data persistence, data distribution

I. INTRODUÇÃO

Ambientes Virtuais Distribuídos (AVDs) permitem a múltiplos usuários (pessoas ou programas) interajam entre si, compartilhando informações e objetos em tempo real. Cabe ao ambiente prover realismo grau de imersão, incorporando gráficos 3D sons de alta qualidade. Cada usuário tem sua própria visão deste mundo e suas alterações de estado são compartilhadas dentre os demais usuários [1] [2].

Os participantes de um AVD têm a sensação de estarem no mesmo ambiente físico, mesmo quando distantes geograficamente. O espaço compartilhado deve apresentar as mesmas características a todos os seus participantes, independente da heterogeneidade das estações conectadas. A heterogeneidade das estações deve ser tratada pelo ambiente virtual de modo transparente aos usuários.

Ambientes Virtuais Distribuídos são particularmente difíceis e complexos de serem projetados e implementados de forma eficiente. Tal complexidade se deve, principalmente, ao fato de que um AVD resulta da junção de três áreas de aplicação da computação: aplicação distribuída, computação gráfica e interação humano-máquina [1].

De modo geral, é desejável que as características do ambiente virtual e também dos objetos manipulados por ele persistam. Ainda que nenhum participante esteja conectado ao ambiente, as informações e objetos do ambiente devem ser armazenados, através de alguma técnica, de modo que seja possível recuperá-los para a próxima execução. Está dentre as responsabilidades do AVDs armazenar e distribuir os dados e objetos do ambiente virtual [2].

Por se tratar de uma aplicação distribuída, ainda é desejável que um dado objeto tenha acesso a serviços de um objeto remoto. Logo, o projeto e implementação de um AVD também deve-se atentar-se a técnicas de distribuição de objetos, possibilitando a invocação de métodos de um objeto remoto.

Neste trabalho, é apresentado um *framework*, que de forma eficiente, gerencia a persistência de dados de AVDs para uma base de dados relacional e a distribuição de objetos sobre o ambiente virtual. Como estudo de caso foram utilizadas

aplicações tridimensionais executadas através da AVDNet. A AVDNet é uma arquitetura para criação e gerenciamento de aplicações em realidade virtual distribuídas não específicas, capaz de suportar uma grande quantidade de participantes e com suporte à gráficos 3D [3] [4].

Para apresentar este trabalho, este artigo está organizado em 5 Seções, incluindo esta introdução. A Seção 2 aborda conceitos de distribuição de dados em AVDs, sobretudo utilizando a tecnologia Java RMI; A Seção 3 apresenta o *framework* para a persistência e distribuição de dados em AVDs; A Seção 4 tem por objetivo apresentar dois estudos de casos utilizados para validar o *framework*; por fim, a Seção 5 apresenta as conclusões deste trabalho.

II. DISTRIBUIÇÃO DE DADOS COM JAVA RMI EM AVDs

Um sistema de objetos distribuídos é aquele que permite a execução de operações remotas. Deve ser possível, a um cliente do ambiente virtual, obter referências de um objeto remoto, que ofereça um serviço desejado, e utilizar-se remotamente de métodos deste objeto [3].

Java oferece nativamente dois mecanismos para o desenvolvimento de aplicações sob o conceito de objetos distribuídos: Java RMI (*Remote Method Invocation*) e Java IDL (*Interface Definition Language*).

O Java RMI oferece suporte ao desenvolvimento de aplicações distribuídas que operam exclusivamente com objetos Java. Seu propósito é fazer com que objetos separados em máquinas virtuais diferentes do Java assemelhem-se e comportem-se como objetos locais [5] [6].

Por sua vez, a arquitetura Java IDL utiliza o padrão CORBA, permitindo a integração de aplicações Java com aplicações desenvolvidas em outras linguagens [7].

Para a utilização do mecanismo Java RMI em aplicações cliente-servidor, é preciso definir uma interface com os serviços (métodos) que são oferecidos pelo servidor. A especificação de uma interface remota em Java se dá da mesma maneira que qualquer outra interface em Java, apenas com a necessidade de herdar-se a interface *Remote*, além disso todos os métodos devem possuir a exceção *RemoteException*.

Para a invocação remota de métodos pelo cliente utilizando a tecnologia Java RMI, os argumentos são empacotados e enviados ao servidor, que os desempacota e realiza o processamento do serviço solicitado. Após a realização do processamento, o objeto responsável pelo serviço, empacota o resultado e devolve ao objeto solicitante.

III. O framework DE PERSISTÊNCIA E DISTRIBUIÇÃO DE DADOS

Um dos problemas encontrados na persistência de dados ao utilizar-se uma linguagem orientada a objetos se dá pela incompatibilidade entre sistemas de bancos de dados relacionais e os os objetos da linguagem. Somado a isso, ainda existe a necessidade de sistemas distribuídos oferecerem a possibilidade de invocação de métodos de objetos remotos.

Tendo em vista o problema de inadequação entre linguagens de programação orientadas a objeto e banco de dados

relacionais, e oferecer um alternativa de distribuição de objetos em AVDs, é apresentado neste artigo um *framework*.

Com uma solução baseado sobre o conceito de *framework* – um conjunto de classes cooperantes que constroem um projeto reutilizável para uma específica classe de *software* – tem-se como benefício modularidade, extensibilidade e inversão de controle, além alto índice de reusabilidade em aplicações do domínio específico do *framework* [8] [9].

Um Ambiente Virtual Distribuído deve persistir dados referentes a criação, apresentação e gerenciamento de dados. Estes dados são constituídos não somente de informações referentes à descrição geométrica e aparência de objetos, bem como as normas que regem e o que nela está incluído [3] [10].

De modo geral, os objetos de um ambiente virtual possuem características em dualidade: objetos de estado dinâmico, tipicamente alocados em memória volátil em tempo de execução; e objetos de estado persistente, que devem ser mantidos em um sistema não-volátil, de modo que seja passíveis de recuperação na próxima execução, mesmo após o encerramento do programa que os criou.

O *framework* apresentado responsabiliza-se pela persistência dos objetos de comportamento persistente. Os dados destes objetos são armazenados em uma base de dados em períodos pré-determinados de tempo e para que o *overhead* sobre o servidor do ambiente virtual seja reduzido, apenas objetos que sofreram alterações são salvos.

A Figura 1 ilustra o diagrama de classes do framework para persistência de objetos de um AVD. A classe *AVDConectaBanco* realiza a conexão entre o Banco de Dados relacional utilizado pelo ambiente virtual e o *framework*. Por meio desta classe é possível definir os parâmetros da conexão com o banco, como endereço, usuário e senha.

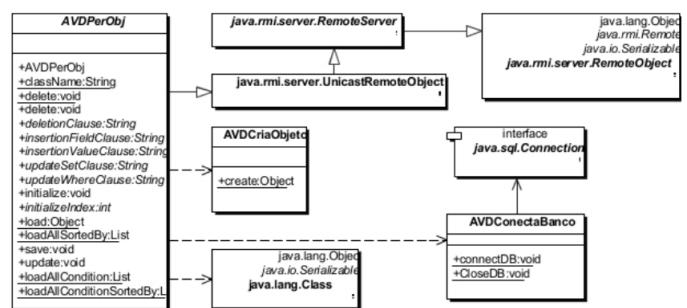


Figura 1. Diagrama de Classes do *framework* – Fonte: [3][p. 76]

A classe abstrata *AVDPerObj* contém os métodos necessários para a criação dos objetos do ambiente virtual através dos dados persistidos no banco de dados, além de métodos para o mapeamento dos atributos dos objetos do ambiente virtual para o banco de dados relacional.

Cada uma das classes persistentes devem herdar a classe abstrata *AVDPerObj*. Por meio de classes abstratas garante-se que qualquer objeto possa ser persistido através do *framework*.

A Figura 2 ilustra um exemplo de mapeamento de um objeto tridimensional hipotético de um ambiente virtual da

para o modelo entidade-relacionamento do banco de dados de modelo relacional.

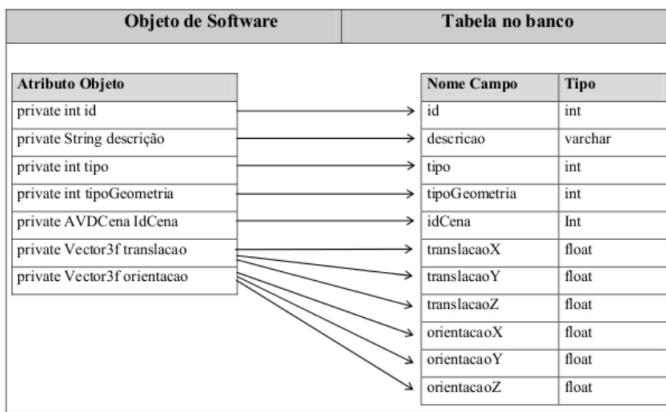


Figura 2. Exemplo de Mapeamento de um objeto tridimensional para um banco de dados relacional através do *framework* – Fonte:[3][p. 76]

Como pode-se notar, o número de campos necessários para armazenar um objeto no banco de dados é proporcional ao número de atributos do objeto persistente. Neste exemplo apresentado, o atributo *translacao* de tipo *Vector3f* é armazenado em três campos (*translacaoX*, *translacaoY* e *translacaoZ*) da tabela do banco de dados. De forma semelhante, o mapeamento do atributo *orientacao* também acarreta em três atributos da tabela do banco de dados.

Com a abordagem adotada pelo *framework*, não é necessário armazenar todo o objeto (apenas seus atributos). Assim, informações desnecessárias do objeto, que podem ser obtidas em sua reconstituição, não são armazenadas, evitando a sobrecarga sobre o servidor de banco de dados.

Para a distribuição dos dados pelo *framework*, foi adotada a tecnologia Java RMI. Com a adoção do Java RMI reduz-se a necessidade de centralização de processamento e armazenamento, garantindo maior escalabilidade ao sistema e independência para o servidor [11].

O processo interno de desempacotamento utilizado pelo mecanismo Java RMI é feito por meio de serialização nativo em Java através da interface *Serializable*. Embora a interface *Serializable* seja nativa em Java e adotada pelo mecanismo de distribuição de objetos RMI, as classes da API Java 3D não podem ser serializadas. Para contornar esta incompatibilidade, para todas classes remotas respectivas classes cliente são criadas, com os atributos do Java 3D e um atributo do tipo de dados da classe remota. Deste modo, são mantidas as funcionalidades de transformar objetos de software em objetos tridimensionais e inseri-los na cena provida pelo Java 3D de forma transparente, rápida e fácil utilizando o *framework*.

IV. ESTUDO DE CASO SOBRE A AVDNET

Para a realização de testes sobre o *framework*, em cenários próximos aos reais, foram elaborados dois estudos de caso. Para tanto foram desenvolvidas as aplicações “Sala de Atendimento Virtual do Professor” e “Armazenamento Virtual 3D”, ambas executadas sobre a arquitetura AVDNet.

A AVDNet é uma arquitetura de software flexível para criação e gerenciamento de aplicações de realidade virtual distribuída em larga escala de propósito geral que suporta a modelagem e tratamento do comportamento de objetos tridimensionais. Adota bases de dados MySQL, distribuídas em uma estrutura híbrida, onde as informações, como objetos estáticos (modelos geométricos, texturas, sons, etc.) e dinâmicos (posição, velocidade, aceleração, etc), são armazenados e gerenciados por uma estação servidora [3] [4].

A estação servidora realiza a replicação dos dados ao clientes sob demanda, onde a estação cliente, através de uma base de dados local, pode armazenar e gerenciar as informações e estados obtidas da estação servidora. Com esta abordagem, a carga de processamento sobre a estação servidora é distribuída e garante a consistência entre as demais bases de dados das estações clientes.

Como a AVDNet é uma arquitetura para ambientes virtuais não específicos, as simulações foram realizadas através de estudos de caso. Para tanto foram utilizadas as aplicações “Sala de Atendimento Virtual do Professor” e “Armazenamento Virtual 3D”. As aplicações foram desenvolvidas utilizando a API Java 3D e utilizam a arquitetura AVDNet e seus módulos, com a adição de novas classes e métodos específicos para a aplicação.

A Sala de Atendimento Virtual do Professor possibilita ao professor e alunos interagirem através de uma sala remota virtual, na qual, através de mensagens de texto podem ser apresentadas e esclarecidas dúvidas. O Aluno pode, ainda, interagir com o professor e com os objetos do ambiente. A Figura 3 ilustra a aplicação Sala de Atendimento Virtual do Professor em execução.



Figura 3. Visão da aplicação Sala de Atendimento Virtual do Professor – Fonte:[3][p. 107]

Na aplicação Armazenamento Virtual 3D, os usuários do ambiente virtual interagem com os objetos, representado por um arquivo com gavetas e pastas, além de acessarem informações armazenadas no banco de dados, de acordo com o objeto selecionado. As informações acessadas pelos usuários podem ser modificadas em tempo real, desde que respeite as restrições de acesso ao banco de dados e ambiente virtual, mantendo-os em estado consistente.

O framework foi introduzido à AVDNet em seu Módulo de Modelagem. Como o Java RMI faz uso da interface *Serializable*, nativa do Java, e os objetos da API Java 3D não pode ser serializados é necessário a criação de classes auxiliares que contenham os atributos do objeto tridimensional. Por meio da Figura 4 é possível observar que para cada objeto tridimensional da API Java 3D existe outra classe que implementa a interface *Remote*.

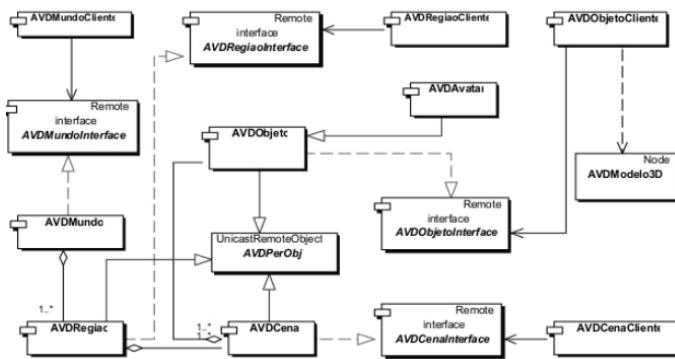


Figura 4. Visão da aplicação Sala de Atendimento Virtual do Professor – Fonte:[3][p. 107]

As classes *AVDMundo*, *AVDRegiao*, *AVDCena* e *AVDOBJeto* contém os métodos responsáveis pela recuperação e atualização dos dados do ambiente virtual na base de dados. As classes que seus objetos devem ser persistidos herdam a classe *AVDPerObj* do framework. Por sua vez, as interfaces *AVDMundoInterface*, *AVDRegiaoInterface*, *AVDCenaInterface* e *AVDOBJetoInterface* definem os serviços oferecidos remotamente e para tanto implementam a interface *Remote*.

A AVDNet foi implementadas em Java com base de dados MySQL. As aplicações de caso de uso também foram implementadas com Java, utilizando recursos da API Java 3D. Deste modo, tem-se um cenário de boa aderência para a realização dos testes, visto que trata-se de aplicações desenvolvidas sobre o paradigma de orientação a objetos persistindo seus objetos em uma base de dados relacional.

Os estudos de casos foram executados com múltiplos usuários conectados ao ambiente virtual simultaneamente. O framework se comportou da maneira esperada. Os objetos, sobretudo os tridimensionais, foram persistidos em bases de dados relacional através do mapeamento dos atributos realizado pelo framework de maneira eficiente. Além disso, os clientes, por meio framework, conseguiram acessar serviços oferecidos por objetos remotos.

V. CONCLUSÕES

A incompatibilidade entre aplicações desenvolvidas sob o paradigma de orientação a objetos e base de dados relacionais deve ser levada em consideração no projeto e desenvolvimento de um AVD. O framework proposto consegue descrever a relação entre as propriedades do objeto e os campos de uma ou mais tabelas no banco, transformando objetos em registros e vice-versa de maneira eficiente.

O overhead sobre o servidor é minimizado pois apenas objetos modificados são persistidos, além disso, apenas infor-

mações necessárias dos objetos são armazenadas descartando aquelas que podem ser obtidas na reconstrução do objeto.

O framework possui alto grau de reusabilidade, já que adotando classes abstratas, qualquer objeto, seja ele tridimensional ou não, pode herdá-la e implementar seus métodos. Desta maneira, qualquer objeto pode ser persistido sem a necessidade de alteração de classes do framework.

Em relação a distribuição de dados, também tem-se uma proposta eficiente, já que a tecnologia Java RMI oferece a possibilidade de distribuição de dados de forma nativa para aplicações desenvolvidas utilizando a linguagem Java, apenas implementando a interface *Remote*.

Além disso, ao utilizar-se Java RMI, as classes são serializadas através da interface *Serializable*, que também é nativa em Java, de modo que não é necessário uma nova implementação. Já para as classes que fazem uso de Java 3D e que não podem estender a interface *Serializable*, o framework cria novas classes que implementa com os atributos do Java 3D, garantindo a compatibilidade para este conjunto de classes.

O framework apresentado, seja para a persistência ou distribuição de objetos, é flexível e comportou-se bem nos estudos de caso realizados sobre a arquitetura AVDNet. No entanto, o framework não limita-se à AVDNet, e por seu alto grau de reusabilidade, pode ser facilmente adotado em qualquer arquitetura de gerenciamento de aplicações em realidade virtual distribuída.

REFERÊNCIAS

- [1] S. Singhal and M. Zyda, *Networked Virtual Environments: Design and Implementation*. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 1999.
- [2] A. S. Tanenbaum and M. V. Steen, *Sistemas Distribuídos: Princípios e Paradigmas*, 3rd ed. Pearson Prentice Hall, 2007.
- [3] R. C. M. Correia, “Avdnet-arquitetura para ambientes virtuais distribuídos escaláveis baseada na infra-estrutura atual da internet,” Ph.D. dissertation, Instituto Tecnológico de Aeronáutica, São José dos Campos, SP, 2005.
- [4] R. C. M. Correia and S. M. Pellegrino, “Scalable distributed virtual environments using multicast application level,” *Proceedings of the 9th International Conference on Computer Supported Cooperative Work in Design*, vol. 1, pp. 140–145, 2005.
- [5] N. Li, J. Mitchell, and D. Tong, “Securing java rmi-based distributed applications,” in *Computer Security Applications Conference, 2004. 20th Annual*, Dez 2004, pp. 262–271.
- [6] M. Philippsen and M. Zenger, “Javaparty – transparent remote objects in java,” *Concurrency: Practice and Experience*, vol. 9, no. 11, pp. 1225–1242, 1997.
- [7] A. P. G. Serra, “O modelo de arquitetura corba e suas aplicações,” *Integração*, 2004.
- [8] M. E. Fayad, D. C. Schmidt, and R. E. Johnson, *Building Application Frameworks: Object-oriented Foundations of Framework Design*. New York, NY, USA: John Wiley & Sons, Inc., 1999.
- [9] S. M. Lewandowski, “Frameworks for component-based client/server computing,” *ACM Comput. Surv.*, vol. 30, no. 1, pp. 3–27, Mar. 1998. [Online]. Available: <http://doi.acm.org/10.1145/274440.274441>
- [10] G. Popescu and C. Codella, “An architecture for qos data replication in network virtual environments,” in *Virtual Reality, 2002. Proceedings. IEEE*, 2002, pp. 41–48.
- [11] J. Linebarger, C. D. Janneck, and G. Kessler, “Shared simple virtual environment: an object-oriented framework for highly interactive group collaboration,” in *Distributed Simulation and Real-Time Applications, 2003. Proceedings. Seventh IEEE International Symposium on*, Oct 2003, pp. 170–180.