



Pedro Alexandre Rodrigues dos Santos

Rastreamento virtual da face: um sistema para espelhar emoções.

Universidade do Minho
Escola de Engenharia





Universidade do Minho
Escola de Engenharia

Pedro Alexandre Rodrigues dos Santos

Rastreamento virtual da face: um sistema
para espelhar emoções.

Dissertação de Mestrado
Mestrado Integrado em Engenharia Eletrónica industrial e
Computadores

Trabalho efetuado sob a orientação do
Professora Doutora Filomena Soares

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.



AGRADECIMENTOS

A realização desta dissertação de mestrado não seria possível sem o apoio e dedicação de algumas pessoas pelas quais estarei sempre grato e aos quais queria aqui deixar as minhas palavras de agradecimento.

- À orientadora desta dissertação, a Professora Doutora Filomena Soares por todo o apoio oferecido, pela disponibilidade que sempre ofereceu e pela valorização do trabalho realizado que sempre defendeu colmatando na publicação de um artigo científico.

- Ao Vinicius Silva, Estudante de Doutoramento na Universidade do Minho, pela partilha de conhecimento e apoio na resolução de algumas dificuldades surgidas durante o desenvolvimento deste trabalho.

- Um agradecimento especial ao Professor Doutor Alberto Simões do IPCA por toda a disponibilidade que sempre demonstrou e pela partilha de conhecimento sobre o motor de jogo Unity 3D. Sem a sua ajuda todo o processo de aprendizagem teria sido mais complicado.

-Ao Professor Doutor João Miguel Sena por todas as sugestões, ideias e feedback oferecidos durante toda a realização do trabalho.

-À escola EB1 de Gualtar, seus professores e alunos pela oportunidade de realização de testes em ambiente escolar e pela simpatia e prestabilidade com que me recebeu.

-À Doutora Ana Paula Pereira pelo seu contributo e partilha de conhecimento sobre crianças com PEA de uma forma sempre muito profissional e sincera.

-Aos meus pais pela oportunidade de ingressar no ensino superior e por todo o apoio que sempre me deram durante o curso e realização deste trabalho.

-À Ana Rita Sá pelo empenho e dedicação na gravação das vozes utilizadas neste trabalho e por todo o apoio e incentivo que sempre me ofereceu.

DECLARAÇÃO DE INTEGRIDADE

Declaro ter atuado com integridade na elaboração do presente trabalho académico e confirmo que não recorri à prática de plágio nem a qualquer forma de utilização indevida ou falsificação de informações ou resultados em nenhuma das etapas conducente à sua elaboração.

Mais declaro que conheço e que respeitei o Código de Conduta Ética da Universidade do Minho.

RESUMO

A capacidade de interação social, comunicação e reconhecimento de emoções é uma característica que permite ao ser humano a integração no meio social contribuindo para a sua autoestima e o seu bem-estar. No entanto, existem pessoas com dificuldades em exprimir e interpretar estados emocionais, o que por vezes pode contribuir para uma exclusão na sociedade. Um caso particular são as crianças com Perturbações do Espectro do Autismo (PEA). Estas crianças possuem dificuldades na interação social e caracterizam-se por manifestarem padrões repetitivos e gostos muito restritos.

Na literatura encontram-se vários estudos que utilizam jogos sérios e outras ferramentas para promover a interação com crianças com PEA. A utilização de robôs para este fim é também muito comum. Estas ferramentas são capazes de reconhecer emoções e promover desenvolvimentos das capacidades cognitivas destas crianças.

Nesta dissertação desenvolveu-se um jogo sério em que um avatar (ZECA Avatar) é capaz de interagir com a criança, lançando-lhe desafios que poderão facilitar a imitação e o reconhecimento de emoções. Este jogo cria uma alternativa à utilização do robô na interação com crianças podendo estas ter acesso a ele em qualquer lugar, sem depender do robô.

O jogo está dividido em três níveis. No primeiro nível a criança tem de executar alguns movimentos faciais que são replicados pelo avatar e validados por software. No segundo tem-se uma imitação diferida onde o avatar ensina a criança a expressar emoções através de expressões faciais. O terceiro é o modo história. Neste nível são contadas histórias e a criança tem de ser capaz de detetar qual a emoção presente em cada uma.

Para o desenvolvimento gráfico do jogo foi usado o *software* Unity 3D e o OpenFace para a monitorização da face da criança.

Por último, foram feitos testes laboratoriais e em ambiente escolar para validar os níveis criados bem como a robustez do sistema. Assim, com este jogo espera-se que as crianças com PEA desenvolvam as suas capacidades cognitivas e melhorem a sua capacidade de comunicação e interação com os pares.

PALAVRAS-CHAVE

Jogo sério, Emoções, Perturbação do Espectro do Autismo (PEA), ZECA (Avatar)

ABSTRACT

The social capacity to interact, communicate, understand and express emotions is a human characteristic that allow us to live in an integrated social environment. Though, there are people with difficulties in expressing and distinguishing emotional stages which contribute for exclusion and discrimination. A particular case are children with Autism Spectrum Disorders (ASD). These children have some difficulties in social interaction. Usually they are characterized by manifesting repetitive patterns of behavior and very singular preferences. There are several studies in the literature that use games and other tools to promote the interaction with children with ASD. In our days, the use of robots for this purpose is also very common. Some of these tools can recognize emotions and improve the cognitive abilities of these children.

The goal of this dissertation is to develop a digital game in which an avatar (ZECA Avatar) is able to interact with the child, throwing him/her challenges that may facilitate the imitation and recognition of emotions. This game will create a virtual alternative to the physical robot, to interact with children, being able of be accessed anywhere.

The game is divided into three levels. At the first level the child must perform certain facial movements that are replicated by the avatar and validated by the software. In the second, the software has a deferred imitation where the avatar teaches the child to express emotions through facial expressions. The last level is the story mode where stories are told and the child should be able to understand which emotion is presented. Unity 3D software engine was chosen to develop the game and the OpenFace software for the monitoring and the face-tracking of the children.

Finally, some tests were performed in the lab and with children in the school environment to validate all game challenges and measure the robustness of the system. Thus, with this game it is expected to help children with ASD to develop their cognitive skills and improve their ability to communicate and interact with peers/world.

KEYWORDS

Serious Game, Emotions, Autism Spectrum Disorders (ASD), ZECA (Avatar)

ÍNDICE

Agradecimentos.....	iii
Resumo.....	v
Abstract.....	vi
Índice de Figuras.....	ix
Índice de Tabelas.....	xi
Lista de Abreviaturas.....	xii
1. Introdução.....	13
1.1 Motivação.....	14
1.2 Enquadramento do tema.....	14
1.3 Objetivos.....	14
1.4 Resultados científicos do desenvolvimento do trabalho.....	15
1.5 Organização do documento.....	15
2. Estado de arte.....	17
2.1 Avatar.....	17
2.2 Robôs Interativos.....	17
2.3 Espelhamento de movimentos.....	18
2.4 Jogos criados para crianças com PEA.....	19
3. Métodos.....	22
3.1 Sistema de codificação da ação facial (<i>FACS</i>).....	22
3.2 Unidades de Ação (AU).....	22
3.3 Motor de jogo - Unity 3D.....	25
3.4 Adobe Fuse – Criação de avatares humanoides.....	26
3.5 Mixamo – Criação de animações.....	27
3.6 TCP-IP – protocolo de comunicação.....	28
4. Desenvolvimento do jogo – HiZeca.....	29
4.1 Público-alvo.....	29
4.2 Projeto ZECA Robô.....	29
4.3 Objetivo do jogo.....	30
4.4 ZECA Avatar – Desenvolvimento do objeto 3D da personagem principal do jogo.....	31
4.5 Cenários de jogo.....	32

Cenário 1 – Treino de movimentos.....	33
Cenário 2 – Treino de Expressões faciais	37
Cenário 3 – Modo História	40
5. Arquitetura e Design do HiZeca	44
5.1 Menus iniciais e menus comuns.....	44
5.2 Funções comuns a todos os níveis.....	47
5.3 Cenário 1	49
Animação do Avatar	49
Arquitetura e Design	51
5.4 Cenário 2.....	56
Animações.....	56
Arquitetura e Design	58
Seleção de emoções	60
5.5 Cenário 3.....	61
5.6 HiZeca – Junção de módulos	62
6. Resultados.....	64
6.1 Testes Laboratoriais	64
6.2 Testes em Ambiente Escolar.....	69
Crianças tipicamente desenvolvidas	69
Crianças com PEA	73
7. Conclusões e Trabalhos Futuros.....	79
Referências Bibliográficas	83

ÍNDICE DE FIGURAS

Figura 1 – AU utilizadas para a deteção das emoções [30].....	23
Figura 2 - Interface gráfica para utilização do OpenFace	24
Figura 3 - Interface gráfica Unity 3D	25
Figura 4 - Adobe Fuse - Criação de Avatares.....	26
Figura 5 - Mixamo.....	27
Figura 6 - Comunicação TCP-IP.....	28
Figura 7 - ZECA robô	29
Figura 8 – ZECA avatar desenvolvido no Photoshop	31
Figura 9 - ZECA avatar versão final.....	32
Figura 10 - Linha lógica do HiZeca	33
Figura 11 - Imitação por Modelo	34
Figura 12 - Aparência gráfica de alguns momentos de jogo do cenário 1	34
Figura 13 - Dados gravados cenário 1	37
Figura 14 - Imitação diferida	37
Figura 15 - Ciclo de jogo Cenário 2	38
Figura 16 - Barra de controlo de resposta	39
Figura 17 - Dados gravados cenário 2	40
Figura 18 - Modo história	41
Figura 19 - Imagens exemplo do modo história	41
Figura 20 - Menu identificação da emoção	42
Figura 21 - Sequência de respostas erradas.....	42
Figura 22 - Dados gravados cenário 3.....	43
Figura 23 - Menu identificador do utilizador.....	44
Figura 24 - Menu inicial	45
Figura 25 - Fluxograma menu principal	45
Figura 26 - Menu início de cada nível	46
Figura 27 - Menu pausa.....	46
Figura 28 - Menu fim de nível	47
Figura 29 - Fluxograma função do início do nível	47
Figura 30 - Fluxograma função continuar jogo	48
Figura 31 - Fluxograma função jogar outra vez	48

Figura 32 - Fluxograma função sair do jogo	49
Figura 33 - Filtro de Smoothing [38]	50
Figura 34 - Arvore de objetos do avatar	51
Figura 35 - Estrutura de objetos do Mixamorig:Hips.....	51
Figura 36 - BlendShapes Unity 3D	52
Figura 37 - String de dados enviados	52
Figura 38 - Fluxograma geral Cenário 1	53
Figura 39 - Sub-rotina para verificação dos movimentos	54
Figura 40 - Lista de falas cenário 1	55
Figura 41 - Arvore de objetos cenário 1	56
Figura 42 - Animação da expressão facial de medo	56
Figura 43 - Animação criada pelo Mixamo para dança do avatar.....	57
Figura 44 - Bloco de controlo das animações criadas	58
Figura 45 - Arvore de objetos cenário 2	58
Figura 46 - Fluxograma geral Cenário 2	59
Figura 47 - Dados recebidos no Unity 3D	60
Figura 48 - Fluxograma Sub-Rotina verificação de expressões.....	61
Figura 49 - Fluxograma geral cenário 3	62
Figura 50 - Assets do HiZeca	62
Figura 51 - Cenas disponíveis para o jogo	63
Figura 52 - Excerto de código responsável por abrir o nível 1.....	63
Figura 53 - Posicionamento para realização de testes laboratoriais	64
Figura 54 - Resultados nível 1 criança D	73
Figura 55 - Resultados nível 2 criança D	74
Figura 56 - Resultados nível 3 criança D	75
Figura 57 - Resultados nível 1 criança E.....	76
Figura 58 - Resultados nível 2 criança E.....	77
Figura 59 - Resultados nível 3 criança E.....	78

ÍNDICE DE TABELAS

Tabela 1 – Resumo comparativo dos jogos para crianças com PEA.....	21
Tabela 2 - Unidades de ação utilizadas	23
Tabela 3 - Expressões faciais e respetivas UA.....	24
Tabela 4 - Lista de movimentos cenário 1	35
Tabela 5 - Numeração das emoções enviadas para o Unity 3D.....	60
Tabela 6 - Lista de adultos utilizados nos testes laboratoriais.....	65
<i>Tabela 7 - Tabela resumo com resultados dos testes laboratoriais nível 1.....</i>	<i>66</i>
<i>Tabela 8 - Tabela resumo com resultados dos testes laboratoriais nível 2.....</i>	<i>67</i>
Tabela 9 - Lista de crianças utilizadas nos testes em ambiente escolar.....	69
Tabela 10 - Tabela resumo com valores médios dos resultados dos testes em ambiente escolar nível 1 com crianças tipicamente desenvolvidas.....	70
Tabela 11 - Tabela resumo com valores médios dos resultados dos testes ambiente escolar nível 2 com crianças tipicamente desenvolvidas.....	71
Tabela 12 - Tabela resumo com resultados dos testes ambiente escolar nível 3 com crianças tipicamente desenvolvidas	72

LISTA DE ABREVIATURAS

CERT	C omputer E xpression R ecognition T oolbox
DAE-Collada	COLLA borative D esign A ctivity
FBX	F ilm BoX
IGU	Interface G ráfica do U tilizador
IHC	Interação H umano C omputador
IP	Internet P rotocol
PEA	P erturbação do E spetro do A utismo
TCP	T ransmission C ontrol P rotocol
SCAF	S istema de C odificação de A ção F acial (FACS terminologia inglesa)
UA	U nidade A ção (AU terminologia inglesa)
ZECA	Z eca E ngaging C hildren with A utism

1. INTRODUÇÃO

O ser humano tem a capacidade de espelhar no rosto diversas emoções que caracterizam o seu estado emocional. Este espalhamento é feito de uma forma consciente de acordo com o que está a sentir num determinado momento, conseguindo fazer uma leitura dessas emoções através de expressões faciais feitas por outros seres humanos. Isto contribui para que a interação e compreensão dos pares seja facilitada. No entanto, a leitura dessas expressões pode ser uma tarefa complicada para crianças com Perturbação do Espectro de Autismo (PEA) [1]. Estas crianças apresentam dificuldades ao nível da interação social, do uso convencional da comunicação e da linguagem e pela restrita variedade de interesses [2].

A utilização de robôs para a interação com as crianças com PEA tem sido muito testada [3], mas atualmente a criação de aplicações para o apoio a estas crianças tem vindo a aumentar visto ser uma alternativa à utilização dos robôs. Estas aplicações são desenvolvidas para promover o desenvolvimento de certas capacidades cognitivas de uma forma didática.

Para que todas estas aplicações sejam bem aceites, é necessário que a Interface Gráfica do Utilizador (GUI em terminologia inglesa) seja cativante e intuitiva para que as crianças não tenham dificuldades em usá-las. Como tal, a utilização de jogos para o apoio no desenvolvimento de capacidades de comunicação e interação em crianças com PEA é uma das potenciais soluções a utilizar. Foram já realizados testes com algumas crianças [4] onde se comprovou avanços significativos no seu comportamento no que diz respeito ao relacionamento com colegas e professores, por exemplo.

No entanto, a Interação Humano Computador (HCI em terminologia inglesa) necessita de alguns progressos no que diz respeito à compreensão de reações do utilizador e adaptação do comportamento por parte da máquina. É importante que esta interação seja fluida e dinâmica evitando que o humano em algum momento sinta que do outro lado existe apenas uma máquina programada para determinadas tarefas predefinidas e estáticas. Através do rosto, o ser humano pode transmitir diversas expressões [5] e é com essa capacidade que se consegue relacionar com os outros permitindo assim que seja aceite e introduzido na sociedade. Para isso é importante que crianças com PEA percebam cada emoção e a consigam relacionar com expressões faciais características.

1.1 Motivação

A oportunidade de criação de uma aplicação para ajudar as crianças com PEA foi desde logo um incentivo para a realização deste trabalho pois este apoio é adequado para a sua evolução, desenvolvimento e crescimento humano. Neste projeto está identificada uma dificuldade do público-alvo e como tal a criação de um jogo cooperativo pode vir a ser determinante na formação destes seres humanos.

A integração num projeto inovador no qual este tema se insere e a oportunidade de utilizar o conceito de HCI também são motivos que contribuíram para a vontade de desenvolvimento do jogo já referido. A oportunidade de utilização de novas ferramentas de trabalho e a aprendizagem de novas linguagens de programação foram também um incentivo.

1.2 Enquadramento do tema

Para o ser humano conseguir interagir é essencial que consiga entender o que o par está a sentir e compreender as suas reações. Na face espelha-se aquilo que sentimos e a partir disso qualquer um pode interpretar uma mensagem, verbal ou não-verbal. Essa mensagem pode ser passada na forma de gestos e expressões. Como as crianças com PEA possuem alguns problemas cognitivos e sociais é importante criar um suporte no sentido de melhorar as suas capacidades de forma a tornar o seu relacionamento com os outros mais fácil e fluido [6]. Para isso a criação de aplicações ou jogos capazes de auxiliar na aprendizagem de emoções e compreensão de estados de emocionais torna-se uma solução relevante. Essas aplicações passam a ser uma base de apoio importante por conseguirem promover o desenvolvimento das capacidades destas crianças de uma forma didática através de entretenimento construtivo. O jogo desenvolvido enquadra-se num projeto desenvolvido na Universidade do Minho relacionado com o autismo (Projeto Robótica Autismo, <http://robotica-autismo.dei.uminho.pt/>), onde foi utilizado um robô humanoide para interagir com crianças portadoras de PEA.

1.3 Objetivos

Nesta dissertação pretende-se desenvolver um jogo sério e didático para que o utilizador (em especial crianças com PEA) o encare como um divertimento a fim de adquirir a capacidade de associar expressões faciais a emoções. Os desafios que serão lançados à criança irão permitir que esta compreenda o tipo de emoção e consiga interpretar estados emocionais. Para isso será feita a monitorização da face do utilizador com intuito de verificar a execução dos seus movimentos faciais e apoiar no aperfeiçoamento dos mesmos. Será pedido ao utilizador que replique movimentos feitos pelo avatar e que esse mesmo avatar seja capaz de interagir e incentivar na correção de possíveis movimentos incorretos. Pretende-se

também que o jogo seja de fácil utilização e apelativo criando interfaces simples devido à facilidade de distração por parte do público-alvo. Por fim, pretende-se criar uma alternativa à utilização de robôs interativos dispensando o transporte dos mesmos e facilidade de acesso ao jogo independentemente do local.

1.4 Resultados científicos do desenvolvimento do trabalho

Durante o desenvolvimento deste trabalho foi submetido e aceite um artigo científico com apresentação oral numa conferência:

- Pedro Santos, Vinicius Silva, Filomena Soares, Alberto Simões, “Facial Virtual Tracking: A System to Mirror Emotions”, 19th EPIA Conference on Artificial Intelligence, 3-6 Setembro, Vila Real, 2019.

1.5 Organização do documento

Esta dissertação está dividida em 7 capítulos e está organizada da seguinte forma:

- O capítulo 2 contém uma breve explicação sobre o conceito de avatar, muito usado ultimamente, e a apresentação de alguns projetos sobre robôs interativos com crianças portadoras de PEA. Por último é feita uma apresentação de alguns projetos focados no espelhamento de movimentos e um conjunto de jogos criados para crianças com PEA.

- No capítulo 3 apresentam-se os métodos utilizados para desenvolver este jogo. Como tal é feita uma explicação de alguns conceitos teóricos sobre expressões faciais e movimentos faciais. São também apresentadas todas as ferramentas utilizadas para o desenvolvimento do HiZeca tais como Openface, Unity 3D e Mixamo. É ainda apresentado um breve resumo sobre o protocolo TCP-IP utilizado para a troca de informação entre as ferramentas.

- No capítulo 4 é apresentado os desenvolvimentos feitos neste jogo. Como tal, é feita uma pequena introdução sobre o público alvo e o projeto ZECA robô no qual este trabalho se insere. Em seguida apresentam-se os objetivos do jogo e todos os níveis criados neste jogo.

- No capítulo 5 apresentam-se as técnicas de implementação utilizadas. São apresentados alguns fluxogramas relativos a implementação de menus, funções e níveis criados. Apresentam-se também todas as práticas usadas para o desenvolvimento de cada nível e a forma estes foram criados. Por último é explicado como foi feita a junção de todos os módulos criados gerando assim o HiZeca.

- No capítulo 6 estão expostos todos os resultados obtidos durante os testes realizados. São explicados todos os procedimentos utilizados para a realização dos mesmos em ambiente laboratorial e em ambiente escolar. Apresentam-se algumas tabelas com os tempos de resposta, taxas de validação de movimentos e emoções e respostas certas e erradas no modo história. Por fim é feita uma breve conclusão sobre cada um dos testes.

- O capítulo 7 contém todas as conclusões obtidas ao longo da realização do trabalho e os possíveis trabalhos futuros ou melhorias a fazer. É explicada também a importância de desenvolvimento deste tipo de aplicações e a criação de apoios ao desenvolvimento de crianças com PEA.

2. ESTADO DE ARTE

Neste capítulo apresenta-se uma visão geral sobre o conceito de avatar e a utilização de robôs para a interação com crianças com PEA. Como tal são apresentados alguns projetos envolvendo diferentes robôs humanoides. É também apresentado um conjunto de projetos/jogos que envolvem espelhamento do corpo humano recorrendo a câmaras, como por exemplo a Kinect. Por fim mostram-se alguns jogos criados para crianças com perturbações do espectro do autismo sendo esclarecidos os diferentes objetivos de cada um.

2.1 Avatar

Com o desenvolvimento da tecnologia de produção de jogos nos últimos anos, o mundo dos videojogos torna-se cada vez mais presente na vida das pessoas. Existem alguns termos característicos desta área como é o caso do termo avatar. Um avatar é um elemento totalmente digital que representa uma personagem num jogo. É uma representação 2D ou 3D de um humano ou animal, por exemplo, que pode reproduzir sons ou até mesmo executar movimentos, todos eles pré-programados. No mundo da informática este termo é usado regularmente quando nos referimos à representação virtual das personagens.

2.2 Robôs Interativos

A utilização de robôs para interagir com crianças portadoras de PEA é muito comum hoje em dia. O robô *ZECA (Zeno Engaging Children with Autism)* é um robô humanoide Zeno R-50 da *Hanson Robotics* utilizado num projeto da Universidade do Minho, que atua como promotor na interação com crianças, ensinando estados emocionais e competências académicas num contexto de cenários de jogo educativo, promovendo a interação com os pares [3] [7] [8]. Um outro exemplo é o robô *QRobot* (lançado em 2018) um robô humanoide criado pela *LuxAI S.A.* Este robô apresenta expressões faciais e usa linguagem corporal durante a interação com a criança, praticando habilidades sociais e de comunicação [9]. O *Milo* é também um robô humanoide projetado para ser interessante e acessível para as crianças com PEA. Ele pode andar, falar e até mesmo modelar expressões faciais humanas [10]. Por último, existe também o robô *Face* criado para trabalhar as expressões faciais associadas a emoções com crianças com PEA[11].

2.3 Espelhamento de movimentos

A utilização da linguagem corporal para a manipulação de avatares é muito frequente. Para isso é necessário um controlo sobre todos os movimentos feitos pelo ser humano quando este utiliza uma tecnologia de espelhamento de movimentos. Esse controlo pode ser feito através de comandos ou outros objetos equivalentes que são suportados pelo utilizador. A consola de jogos *Wii*, lançada pela *Nintendo*, faz uso de um comando (*Wii remote*) equipado com um acelerómetro capaz de fazer a deteção dos movimentos exercidos pelo jogador [12]. Com esses movimentos monitorizados, o jogador consegue mover objetos como por exemplo uma raquete, dando-lhe uma experiência de jogo muito real dando a sensação que está de facto no mundo real a jogar. Esta consola é também utilizada na área da saúde nomeadamente na área da fisioterapia. O uso desta consola por pessoas com a doença de Parkinson por exemplo, ajuda na melhoria das suas debilidades funcionais com o equilíbrio da mobilidade e desempenho motor, por exemplo [13]. A *Playstation Move* é também uma consola que utiliza comandos para a leitura de movimentos corporais[14].

No entanto, a utilização de comandos pode ter algumas limitações/requisitos, o que pode pôr em causa a correta leitura dos movimentos ou até mesmo limitá-los. Com o uso de comandos torna-se impossível ter acesso a todos os movimentos do corpo humano pois esses movimentos são lidos pela deslocação do comando, que se encontra sempre na mão do utilizador. Nesse sentido a utilização de câmaras para a monitorização de movimentos pode ser uma alternativa adequada. Com a utilização de imagens consegue-se detetar um movimento não dependendo da presença de nenhum objeto no utilizador. A *Kinect*, sensor desenvolvido pela *Xbox 360* e *Xbox One*, é um exemplo de um sensor para deteção de movimentos. Com este sensor é possível detetar um conjunto de movimentos corporais. Atualmente as potencialidades da *Kinect* são usadas para vários fins não só no mundo de jogos como também em aplicações, onde se torna oportuno a deteção de movimentos corporais. Um exemplo é a aplicação onde o utilizador pode experimentar roupas virtuais que são visualizadas num espelho virtual. Os movimentos da mão servem para a alteração da peça de roupar a experimentar podendo assim o utilizador perceber qual a peça que mais gosta sem ter de vestir nenhuma [15]. É ainda possível detetar qual o tamanho da peça de roupa a utilizar evitando que o utilizador gaste tempo em filas de espera para provedores nas lojas de roupas [16]. O *FakeMi* é um sistema que também utiliza a *Kinect* para rastreamento do corpo humano incluindo o rosto. Com a utilização de um espelho falso (espelho virtual), este sistema faz o espelhamento do corpo e da face do utilizador em tempo real para incorporação de avatares que pode ser usado para terapias [17].

No entanto, o espelhamento de movimentos pode localizar-se apenas no rosto do ser humano. No mundo cinematográfico este espelhamento já se torna comum. Os atores utilizam sensores e câmaras para dar movimento aos avatares previamente criados tal como aconteceu na produção do famoso filme *Avatar* lançado em 2009 dirigido pelo James Cameron. Este espelhamento pode ser feito a partir da extração das *Action Units (AU)* (explicado no capítulo 3) e replicação das mesmas num avatar [18].










2.4 Jogos criados para crianças com PEA

Atualmente a criação de aplicações para o apoio a crianças com PEA é uma área que tem vindo a aumentar visto ser uma alternativa à utilização dos robôs. Estas aplicações são desenvolvidas para promover o desenvolvimento de certas capacidades de forma didática. O *Livox* é uma aplicação [19] que contém vários modos de jogos onde a criança recebe estímulos para desenvolver a capacidade de comunicação. Utilizando a sua base com mais de 20.000 imagens, recursos de áudio e vídeo o utilizador pode escolher um item e o *Livox* produz o áudio ou vídeo que está associado a esse mesmo item. Outro caso é a *Sono Flex*[20], uma aplicação de associação de palavras a imagens que é um bom treino para estas crianças, que tem dificuldades em distinguir por exemplo as emoções. O *TippyTalk* [21] permite que uma pessoa com deficiência verbal comunique traduzindo as imagens em mensagens de texto, que são enviadas para um membro da família ou para o responsável. Isso permite que a criança comunique e expresse um desejo, necessidade ou sentimento.

Dentro do conjunto de jogos para crianças com PEA foi feita uma pesquisa onde foi encontrado um elevado número de aplicações. O *Autism Games* é uma colaboração entre o Programa de Design Multimídia da Universidade de *Swinburne*, a Escola Especialista em Colunas de *Bulleen*, a Iniciativa de Bio-Pesquisa de Autismo de *Swinburne (SABRI)* e o Centro Nacional de e-Terapia (*NeTC*) [22]. Esta ferramenta *online* contém dois *websites*, um para os pais ou professores onde estes podem recolher informação sobre todos os jogos, e o outro para aceder aos jogos. Estes jogos relacionam-se com expressões faciais associados a emoções, alterações de rotinas do dia-a-dia, treino de gestos não-verbais entre outros. Em 2017 foi lançado também o projeto *Emotiplay*, um programa interativo de treino emocional para crianças com autismo. Os diferentes jogos foram projetados para ajudar as crianças a reconhecer, compreender e expressar emoções através de expressões faciais, voz e gestos corporais. Dentro destes jogos tem-se diferentes abordagens à compreensão e reconhecimento de emoções. Isso é feito através de jogos de memória, associação de expressões faciais de diferentes pessoas que estão a sentir ou expressar a mesma emoção, e ainda o reconhecimento de emoções através da audição de frases curtas. Segundo os estudos realizados, este programa melhorou significativamente as habilidades

de reconhecimentos de emoções por parte das crianças com PEA [23]. Para a plataforma *iOS* tem-se também a aplicação *Autism Emotion*. Focada apenas em quatro emoções (felicidade, tristeza, calma, orgulho) esta aplicação inclui a narração de textos descritivos sobre uma imagem e músicas da musicoterapeuta *Rachel Rambach*. Para além de crianças com PEA esta aplicação também se destina a crianças com síndrome de *Asperger* [24]. O *Emotion Mirror* foi outro jogo desenvolvido baseando-se na imitação do utilizador. Neste projeto tem-se duas abordagens. Na primeira o utilizador é levado a expressar uma de seis emoções básicas. No caso de acertar, o avatar executa também essa expressão, num ambiente de imitação do participante. Na segunda abordagem, o avatar expressa uma emoção, mantendo-a, para que o participante a possa expressar também copiando/imitando o avatar. Quando essa expressão é executada corretamente o participante é recompensado ganhando um prémio [25]. Para a deteção e monitorização da face do utilizador foi usado o sistema *CERT* [26] que é uma ferramenta de reconhecimento de expressões faciais totalmente automático, em tempo real, e acesso gratuito. O *CERT* pode codificar automaticamente a intensidade de 19 *AU* diferentes do *Facial Action Coding System (FACS)* e as seis expressões faciais básicas. Um outro jogo é o *SmileMaze*. Neste jogo tem-se uma abordagem diferente do anterior. O objetivo do jogo é percorrer um labirinto com uma peça utilizando o teclado, e modo análogo ao jogo *pacman*. No entanto, ao longo do percurso existem obstáculos que são desbloqueados através da expressão de emoções. Se o utilizador expressar corretamente a emoção o obstáculo é retirado e pode continuar o labirinto até terminar [27]. Neste jogo também foi utilizado o sistema *CERT* para a aquisição de dados sobre a face do utilizador. *LIFEisGAME* é também um jogo, criado na Universidade do Porto, que visa ajudar crianças com Perturbações do Espectro do Autismo, sem induzir *stress*, a reconhecer as emoções faciais de um modo interativo. Neste jogo é possível manipular a face de um avatar com o intuito de identificar as diferentes emoções. Este jogo inclui cinco modos de jogo que variam desde jogos de memória, de reconhecimento e de construção de expressões faciais com a face do avatar [28]. Todos estes jogos atrás referidos foram desenvolvidos para o auxílio a crianças com PEA, cada um com diferentes abordagens e objetivos. No entanto, todas visam o melhoramento do relacionamento destas crianças com os outros facilitando assim a sua integração na sociedade. A Tabela 1 apresenta um resumo comparativo de algumas funcionalidades dos jogos apresentados. O *HiZeca*, desenvolvido nesta dissertação, destaca-se no conjunto de jogos apresentados na Tabela 1 por englobar três estilos ou abordagens de jogo que apesar de já existirem não estão agregadas num único jogo. Nesse sentido o *HiZeca* torna-se um jogo que disponibiliza três modos de jogo com uma sequência previamente definida pensada no objetivo geral do jogo, a distinção e reconhecimento de emoções e suas expressões.

Tabela 1 – Resumo comparativo dos jogos para crianças com PEA.

Jogo	Objetivo	Espelhamento Face	Deteção Expressões	Reconhecimento Emoções	Modo História
 <p>Livox</p>	Auxílio na comunicação	Não	Não	Não	Não
 <p>Sono Flex</p>	Auxílio na comunicação	Não	Não	Não	Não
 <p>TippyTalk</p>	Auxílio na comunicação	Não	Não	Não	Não
 <p>Autism Games</p>	Emoções e raciocínio	Não	Não	Sim	Sim
 <p>Autism Emotion</p>	4 Emoções	Não	Não	Sim	Não
 <p>Emotion Mirror</p>	Treino de Expressões	Sim	Sim	Sim	Não
 <p>SmileMaze</p>	Treino de Expressões	Não	Sim	Sim	Não
 <p>LifeisGame</p>	Emoções	Não	Sim	Sim	Não
 <p>EmotiPlay</p>	Emoções	Não	Não	Sim	Sim

3. MÉTODOS

Neste capítulo apresentam-se os métodos utilizados para o desenvolvimento do jogo. Para além dos métodos são apresentadas também as ferramentas escolhidas e algumas características de cada uma. Inicialmente é apresentado o sistema (*FACS*) e o conceito de *AU*. Em seguida é apresentada o motor de jogo utilizado, o Unity 3D e os softwares usados para a criação do avatar e sua animação. Por fim é apresentado o protocolo de comunicação responsável pela troca de informação entre o software de aquisição de dados da face do jogador e o Unity 3D. Para essa comunicação foi escolhido o protocolo TCP-IP.

3.1 Sistema de codificação da ação facial (*FACS*)

As emoções podem ser caracterizadas como negativas (tristeza, raiva, repulsa ou medo), positivas (felicidade ou surpresa) ou neutra. As expressões faciais estão presentes na comunicação e interação entre seres humanos. Estas podem transmitir emoções e opiniões. O *FACS* (terminologia inglesa) é um método utilizado para caracterizar a expressão física das emoções [29]. Este sistema, desenvolvido por *Ekman e Friesen* [5] determina várias posições e movimentos da cabeça, dos olhos e da boca bem como toda a atividade facial visualmente distinguível. Com este sistema o ser humano consegue padronizar expressões faciais através das *AU* que permitem a identificação automática de emoções, sendo apenas necessário reconhecer quais as *AU* presentes num determinado momento, na face de um ser humano.

3.2 Unidades de Ação (*AU*)

O movimento dos músculos da face do ser humano é visível e perceptível a olho nu. Dependendo do nosso estado emocional existem músculos que são movimentados ou contraídos e outros não. Para definir cada músculo ou movimento do músculo da face do ser humano foram criadas as *AU*. Uma *AU* representa ações realizadas por um músculo ou um grupo de músculos e serve de entrada para o sistema *FACS* atrás referido. Na face humana estão presentes 30 *AU* diferentes. Para além dessas existem outras que estão relacionadas com movimentos da cabeça tais como ângulos de inclinação e rotação. No entanto no presente trabalho foram utilizadas apenas 18 dessas *AU*. Na Figura 1 pode ver-se a localização das diferentes *AU* utilizadas.

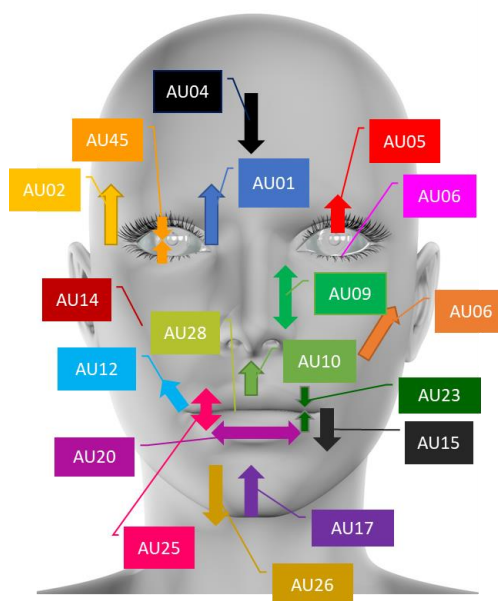


Figura 1 – AU utilizadas para a detecção das emoções [30]

Segundo *Ekman*, diferentes emoções contêm associadas diferentes AU [29] podendo estas estar presentes em mais que uma emoção. Com a detecção das AU pode determinar-se a emoção presente. Na Tabela 2 encontram-se as AU utilizadas e a sua descrição e na Tabela 3 quais as AU presentes em cada emoção.

Tabela 2 - Unidades de ação utilizadas

CÓDIGO AU	DESCRIÇÃO AU
AU01	Levantador da sobrancelha interior
AU02	Levantador da sobrancelha externa
AU04	Abaixador da sobrancelha
AU05	Levantador da pálpebra superior
AU06	Levantador das bochechas
AU07	Apertador da pálpebra
AU09	Enrugar o nariz
AU10	Levantar lábio superior
AU12	Puxador canto do lábio
AU14	<i>Dimpler</i>
AU15	Depressor do canto do lábio
AU17	Levantador do queixo
AU20	Esticador do lábio
AU23	Endurecer o lábio
AU25	Separador de lábios
AU26	Queda da mandíbula
AU28	Sucção dos lábios
AU45	Piscar olhos

Tabela 3 - Expressões faciais e respetivas UA

Expressão Facial	AU
Felicidade	6 + 12
Tristeza	1 + 4 + 15
Surpresa	1 + 2 + 5 + 26
Medo	1 + 2 + 4 + 5 + 7 + 20 + 26
Raiva	4 + 5 + 7 + 23
Desgosto	9 + 15 + 16
Neutro	0

Para a deteção dessas *AU* e monitorização da face do utilizador foi escolhida a biblioteca *OpenFace* [31]. Esta biblioteca com o código aberto, deteta as diferentes *AU* através de redes neuronais artificiais retornando os pontos faciais através do sistema *FACS* com o objetivo de ser feita a verificação necessária em relação a expressões executadas pelo utilizador. Na Figura 2 visualiza-se a interface gráfica desenvolvida para visualização de toda a informação recolhida pela câmara e processada com auxílio do *OpenFace*. Como já foi referido o *OpenFace* permite também a extração dos ângulos de posição e orientação da cabeça da pessoa.

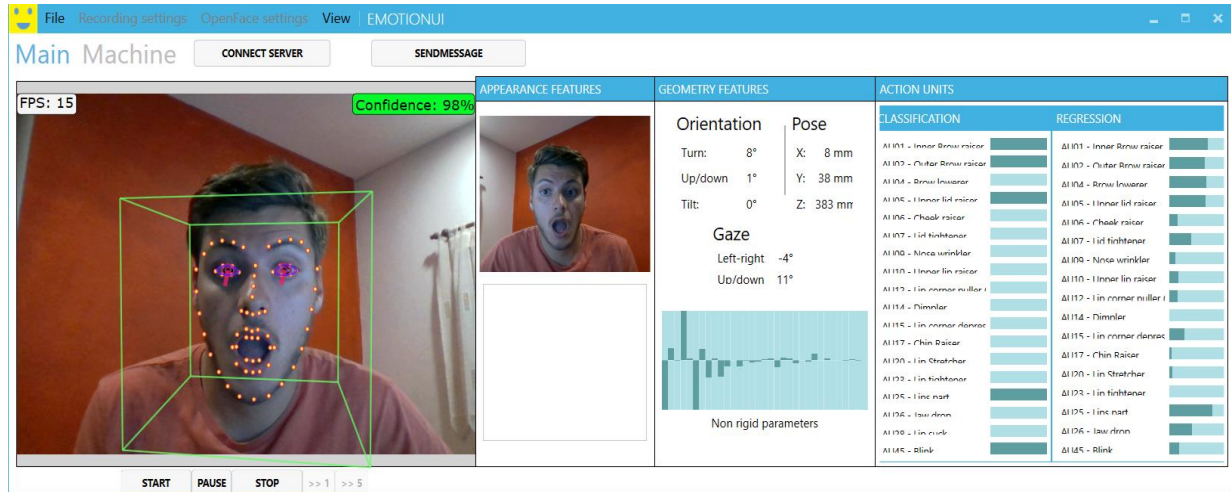


Figura 2 - Interface gráfica para utilização do OpenFace

3.3 Motor de jogo - Unity 3D

Como em todos os jogos acontece, a parte gráfica de uma aplicação deve ser apelativa. A criação de menus, modos de jogos e programação de jogos torna-se mais fácil com ajuda de *software* otimizado para isso. Neste projeto a parte gráfica do jogo e a criação da sua jogabilidade foi desenvolvida no motor de jogo *Unity 3D*. Foi escolhido este *software* por ser um motor de jogos de acesso gratuito e por possuir bibliotecas otimizadas para a criação de jogos 2D e 3D [32]. Para além disso possui também suporte para suportar avatares do *Mixamo*[33] com os diferentes movimentos e com a criação de objetos para a manipulação da face do avatar. Esta ferramenta, poderosa e muito utilizada atualmente, foi criada pela *Unity Technologies* e tem como base a programação orientada de objetos utilizando como linguagem o *C#*. Através de *scripts* e objetos criados é feito o controlo de todos os movimentos de uma personagem ou elemento do jogo, construindo assim jogos cada vez mais realistas nos seus movimentos. Toda a estrutura do jogo e sua dinâmica foram implementadas como auxílio desta ferramenta sendo no fim exportado o jogo para o sistema Windows permitindo que este possa ser colocado em qualquer computador sem necessitar do editor de jogo para o executar.

Na Figura 3 é possível visualizar a interface gráfica da ferramenta em questão.

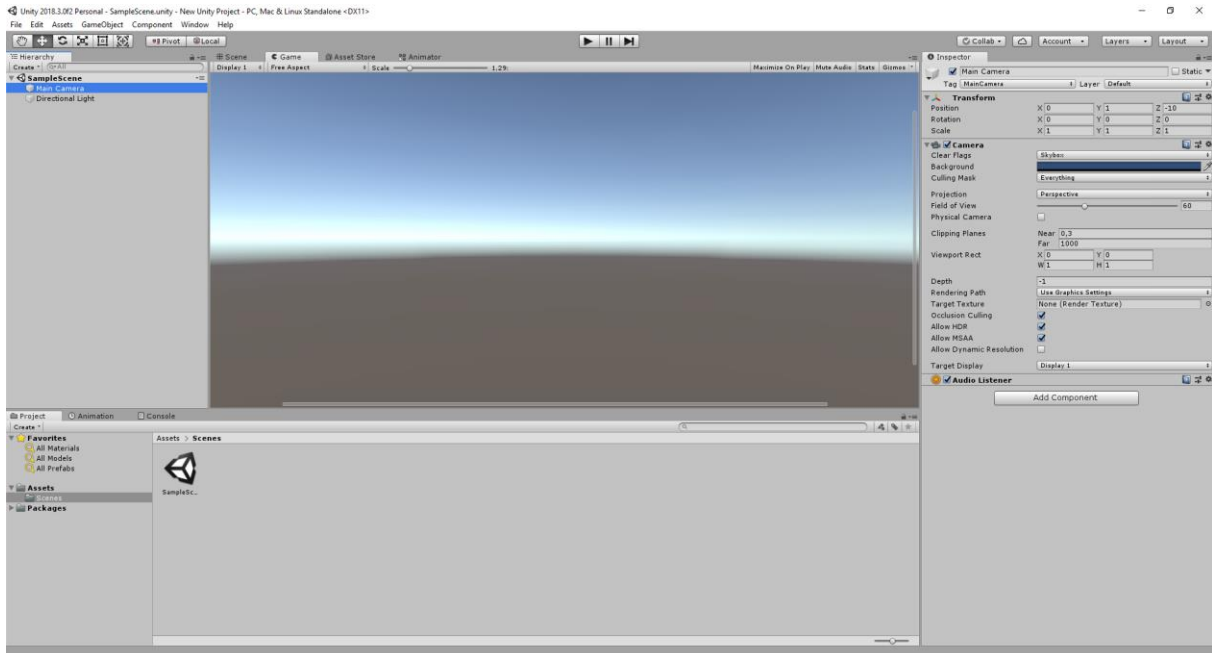


Figura 3 - Interface gráfica Unity 3D

3.4 Adobe Fuse – Criação de avatares humanoides

Como definido inicialmente, o jogo tem, como elo de ligação entre os desafios a propor e a criança, um mediador que é um avatar humanoide. Por esse motivo foi então necessário escolher uma ferramenta que permitisse o desenvolvimento de um avatar. Assim para o desenvolvimento do avatar ou modelo 3D foi escolhido o *Adobe Fuse*. Este *software* permite criar personagens rapidamente usando uma biblioteca de conteúdo 3D de alta qualidade, incluindo rostos, corpos, roupas e texturas. Com o *Adobe Fuse* é possível definir o tamanho das personagens, aumentar ou diminuir diferentes músculos bem com alterar o aspeto visual tais como cabelo, roupas e calçado, por exemplo. Na Figura 44 pode ver-se o aspeto gráfico desta ferramenta.

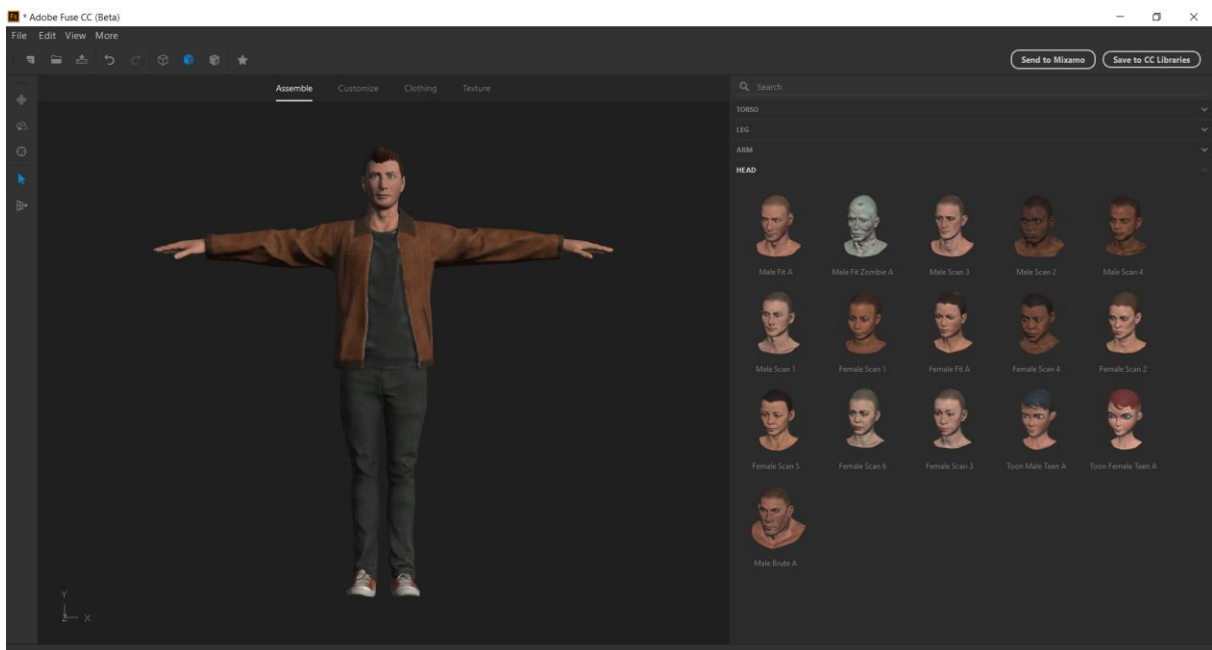


Figura 4 - Adobe Fuse - Criação de Avatares

Foi escolhido este *software* por permitir a fácil criação de um avatar e por permitir também a exportação do modelo 3D para o *Mixamo* (ferramenta apresentada a seguir) para serem feitas animações de alguns movimentos do corpo e a manipulação dos diferentes músculos da cara. Assim, é possível criar expressões faciais associadas a emoções que serão replicadas pela criança.

3.5 Mixamo – Criação de animações

O *Mixamo* é um *software* disponível *online*, capaz de criar movimentos de personagens 3D utilizando métodos de *machine learning* [33]. Para criar os movimentos basta carregar o modelo 3D das personagens e utilizar qualquer um dos movimentos disponíveis tais como caminhar, saltar e correr. Para além dos movimentos corporais é possível também criar suporte para movimentos faciais sendo depois possível a manipulação da face da personagem. Isto permite mais tarde criar expressões faciais utilizadas nos diferentes níveis criados. É possível fazer o *download* do modelo 3D em formatos como o “*FBX*” e “*DAE-Collada*”. O *Mixamo* oferece também a possibilidade de exportar em “*FBX*” destinado para a inclusão no *Unity 3D*. Na Figura 5 pode ver-se o aspeto gráfico do *Mixamo*. Do lado esquerdo tem-se alguns dos movimentos disponíveis e do lado direito pode ver-se a personagem a executar o movimento escolhido. De referir que esta ferramenta também pertence à *Adobe* e por esse motivo a animação da face de uma personagem é só possível se o modelo 3D for carregado diretamente do *Adobe Fuse*. Como neste projeto era necessário criar animações da face do avatar foram estudadas algumas soluções pelo que a utilização destas duas ferramentas da *Adobe* foi a solução mais viável.

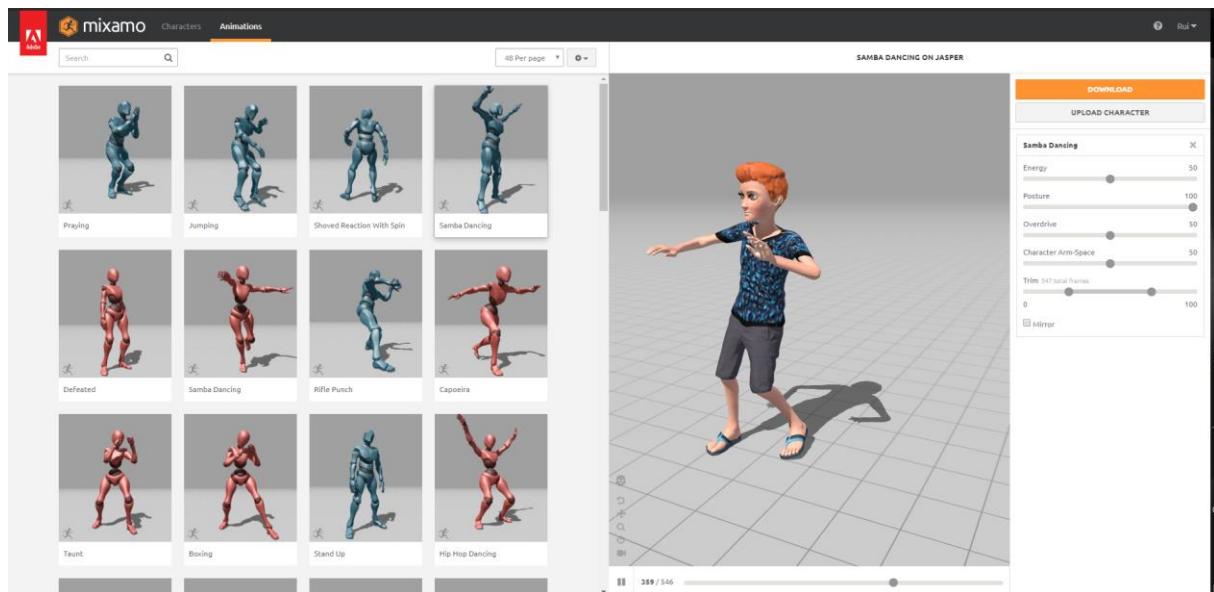


Figura 5 - Mixamo

3.6 TCP-IP – protocolo de comunicação

Este projeto envolve duas ferramentas principais distintas, que necessitam de comunicar entre si para trocar informação. Essas ferramentas são o *Unity 3D* e o *Openface*, ferramentas estas apresentadas nos capítulos anteriores. Como não foi possível a integração de uma ferramenta na outra sentiu-se a necessidade de utilizar um protocolo de comunicação para que a troca de informação fosse possível. Como tal foi escolhido o protocolo *TCP-IP* para essa comunicação. Este protocolo possui a capacidade bidirecional onde o cliente pode ler e escrever mensagens. Neste projeto o *Unity 3D* foi definido como servidor e o *Openface* como cliente sendo que apenas este envia informação para que o *Unity 3D* receba e utiliza para fazer o espelho de movimentos faciais. Na Figura 6 apresenta-se um esquema geral que resume o processo de comunicação deste protocolo.

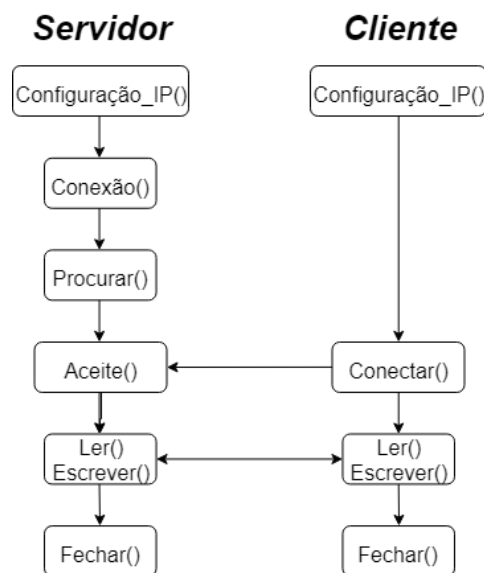


Figura 6 - Comunicação TCP-IP

Este protocolo foi utilizado para enviar mensagens por parte do *Openface* onde se encontram os valores das diferentes *AU* e os ângulos de rotação do pescoço e da cabeça. Para além disso foi também utilizado para enviar a informação sobre qual a emoção detetada num determinado momento. Esta troca de informação é feita em tempo real para que o *Unity 3D* tenha conhecimento do estado da face do utilizador a cada ciclo de máquina.

4. DESENVOLVIMENTO DO JOGO – HiZECA

Neste capítulo está detalhada toda a parte de desenvolvimento do jogo. É feito um esclarecimento sobre o público-alvo e uma breve apresentação do projeto ZECA robô. Depois disso serão apresentados todos os objetivos do jogo e todos os detalhes de desenvolvimento, desde a criação do avatar ao desenvolvimento de cada nível. Serão apresentados fluxogramas de implementação de código desenvolvido e quais os procedimentos utilizados na criação de menus, modelos 3D e animações do avatar.

4.1 Público-alvo

As crianças com PEA tem dificuldades na comunicação com os outros e possuem dificuldades na distinção de emoções. Essa dificuldade pode causar alguns problemas no que diz respeito à interação com os pares. Como tal é necessário criar mecanismos que possam ajudar estas crianças nesse sentido. A abordagem de jogos didáticos é uma das soluções que, segundo informações de jogos desenvolvidos anteriormente, é bem aceite por parte destas crianças. Este trabalho foi desenvolvido para ser utilizado por crianças com idades entre o 6 e 10 anos.

4.2 Projeto ZECA Robô

O projeto do Zeca robô, desenvolvido na universidade do Minho, foi um dos pioneiros em Portugal no que diz respeito à utilização de robôs para a interação com crianças com PEA. Este projeto faz o uso de um robô humanoide, previamente programado para ajudar crianças com PEA a distinguir emoções [7]. Na Figura 7 pode ver-se o aspeto do robô em causa.



Figura 7 - ZECA robô

Este robô expressa emoções para que as crianças o possam observar e num momento posterior expressar também. Para além disso existe também um modo onde o robô imita alguns movimentos feitos pela criança. Por fim foi também criado um modo história onde são contadas histórias associadas a emoções. O avatar a desenvolver no jogo deste trabalho foi inspirado no robô humanoide utilizado no projeto do ZECA robô (Figura 7).

O HiZeca torna-se assim uma alternativa a utilização do robô, pois permite que qualquer criança em qualquer lugar possa ter acesso ao jogo, precisando apenas de possuir o jogo num computador. Neste trabalho ocorre assim uma replicação das atividades feitas com o robô ZECA, mas em ambiente de jogo didático e sério. Assim, com a facilidade de acesso torna-se possível um elevado número de utilizadores sendo mais uma ferramenta disponível no projeto robótica autismo (<https://robotica-autismo.dei.uminho.pt/>).

4.3 Objetivo do jogo

O *HiZeca* é um jogo que se enquadra no grupo de jogos desenvolvidos para serem utilizados principalmente por crianças com PEA. Apesar de não ser um jogo que se destina apenas a essas crianças, toda a jogabilidade, níveis e desafios foram pensados de forma a ajudar o desenvolvimento das capacidades de comunicação, capacidades essas onde estas crianças apresentam algumas dificuldades. Como referido em capítulos anteriores, estas crianças possuem algumas dificuldades no que diz respeito a atenção e concentração pelo que um dos primeiros objetivos do jogo é ser suficientemente apelativo, cativando a atenção da criança. Assim, foram tidos alguns cuidados no que diz respeito aos aspetos da parte gráfica do jogo. Foram criados o número mínimo de elementos com potencial de distração evitando que a atenção da criança recaia em pormenores que não tem relevância para o seu desenvolvimento e objetivo do jogo. Nesse sentido todos os menus e ambientes de jogo criados apresentam um *design* simples e de fácil utilização.

Neste jogo foram criados três níveis distintos com objetivos diferentes. Para além da distinção e reconhecimento de emoções é também um objetivo do *HiZeca* desafiar a criança a executar movimentos faciais por forma a poder mais tarde espelhar as suas emoções com expressões faciais. Por fim, neste jogo pretende-se também melhorar a capacidade de atenção e raciocínio da criança através de desafios onde é necessário criar uma ligação entre histórias, imagens e emoções.

4.4 ZECA Avatar – Desenvolvimento do objeto 3D da personagem principal do jogo

O desenvolvimento do modelo 3D do avatar foi conseguido através do *Adobe Fuse*. Para além do desenvolvimento foi também necessário a criação de animações do corpo e da face do avatar. Para isso foi utilizado o *Mixamo* por permitir que o objeto 3D seja carregado diretamente do *Adobe Fuse*, processo este que só é possível por estas duas ferramentas pertencerem a *Adobe*. Depois de desenvolvido o modelo 3D no *Adobe Fuse* onde foram definidos todos os membros do avatar bem como criada a estrutura do cabelo e roupas, este foi exportado para o *Photoshop* para ser feita uma alteração de cores e introdução de alguns elementos característicos do robô. Na Figura 8 pode ver-se o aspeto gráfico do avatar depois de desenvolvido.



Figura 8 – ZECA avatar desenvolvido no Photoshop

Apesar de haver semelhança com o robô ZECA e de o avatar ser visualmente atrativo, este modelo 3D não pôde ser utilizado. Como era necessário adquirir suporte para a manipulação da face do avatar no *Unity*, e o *Mixamo* não permitir que esse suporte seja dado para modelos 3D que não sejam carregados diretamente do *Adobe Fuse*, então este modelo 3D não pode ser usado. Como tal foi necessário criar outro modelo 3D exclusivamente no *Adobe Fuse*.

Para isso foi utilizado o modelo 3D já criado e ajustadas as cores de roupas e acessórios como botas por exemplo. O aspeto final do avatar está visível na Figura 9.



Figura 9 - ZECA avatar versão final

Depois de desenvolvido e animado o modelo 3D o avatar foi exportado do *Mixamo* em formato “*FBX*” com suporte para a sua utilização no *Unity 3D*. De referir que o *Mixamo* permite incluir no ficheiro exportado apenas um movimento corporal e o suporte para manipulação da face do avatar. Caso em algum momento sejam necessários diferentes movimentos para o mesmo avatar, então o número de avatares exportados irá ser igual ao número de movimentos diferentes pretendidos.

4.5 Cenários de jogo

Habitualmente um jogo possui um conjunto de níveis com objetivos variados. Esses níveis servem para criar um caminho condutor a ser percorrido pelo utilizador com o intuito de se atingir um objetivo. Para além de possíveis alterações no cenário de jogo ou até mesmo das personagens, cada nível possui um grau de dificuldade que normalmente é aumentado de nível para nível. Com isso o utilizador vai treinando algumas técnicas ou modos de jogar para que possa ir passando de nível para nível tornando assim mais fácil atingir o grande objetivo estabelecido para o jogo.

No *HiZeca* foram então criados três níveis diferentes com metas ou objetivos distintos. Como o público-alvo deste jogo possui algumas características especiais foi necessário traçar uma lógica de jogo para tornar o jogo simples de utilizar e ao mesmo tempo introduzir pouco a pouco alguns desafios que aumentam de dificuldade.

O grande objetivo do jogo é a distinção, compreensão e imitação de emoções. No entanto para que isso possa ser feito é necessário que as crianças consigam produzir expressões faciais em primeiro lugar e consigam saber a que emoção está associada cada expressão. Partindo desse princípio é importante treinar alguns movimentos faciais para tornar mais fácil a expressão de emoções. Depois disso espera-se que a compreensão e leitura de expressões faciais torna-se mais fácil. Essa compreensão permite que seja feita uma associação entre expressão e emoção facilitando a interação destas crianças com os outros. No esquema da Figura 10 pode ver-se o caminho estabelecido para este jogo. No nível um tem-se apenas o treino de alguns movimentos faciais. O segundo é uma introdução a expressões faciais, treino das mesmas e associação a emoções. No terceiro e último tem-se a identificação de emoções.



Figura 10 - Linha lógica do HiZeca

Cenário 1 – Treino de movimentos

Neste primeiro cenário o propósito é treinar alguns movimentos faciais que irão ajudar mais tarde na execução de algumas expressões. Este nível serve também para ambientar o utilizador à personagem principal para que possa ser criada alguma empatia entre ambos. Os movimentos feitos pelo utilizador são replicados pelo avatar num contexto de imitação por modelo. Estamos presente uma imitação por modelo quando há replicação ou espelhamento de movimentos [34]. Neste caso o utilizador serve de modelo para o avatar seguir ou copiar. Para isso neste nível de jogo o utilizador é convidado a executar um conjunto de movimentos básicos. Como temos presente uma imitação por modelo o utilizador passa a ter o controlo do avatar através dos seus movimentos. Como tal é então pedido que ajude o avatar a conseguir efetuar corretamente os movimentos pedidos. Veja-se o esquema da Figura 11.

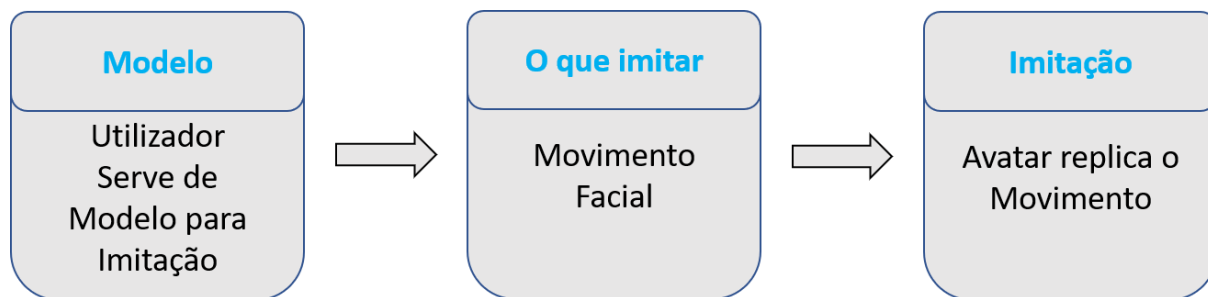


Figura 11 - Imitação por Modelo

No esquema da Figura 11 tem-se uma visão sobre o ciclo de imitação por modelo. Este método é também utilizado em *machine learning* onde uma determinada tarefa ou conjunto de tarefas são executadas várias vezes para que sejam criados padrões de decisão através de algoritmos permitindo que mais tarde a máquina consiga executar essas tarefas autonomamente e de uma forma eficaz. Na Figura 12 são apresentadas algumas imagens com o aspeto gráfico do nível 1.

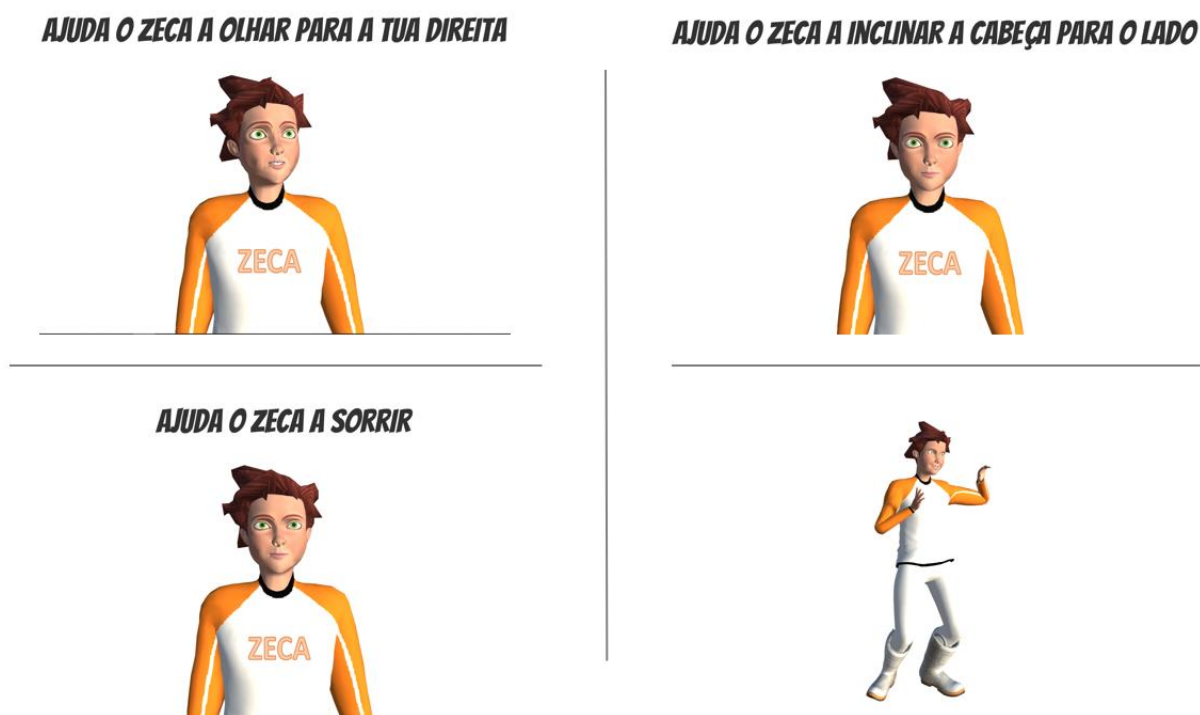


Figura 12 - Aparência gráfica de alguns momentos de jogo do cenário 1

Na Figura 12 estão representados vários momentos do jogo com diferentes movimentos nos quais o avatar imita o utilizador imitando-o facilmente. Neste cenário é apresentado o avatar que imita o utilizador em tempo real e é apresentada uma mensagem com o pedido de execução do respetivo movimento. Caso esse movimento seja executado corretamente pelo utilizador e por sua vez pelo avatar é apresentado um momento de comemoração por parte do avatar baseado numa dança. Assim, neste nível a introdução ao estilo de jogo e a apresentação da personagem principal (ZECA Avatar) torna-se mais fácil. A dificuldade do nível é baixa apesar de serem já introduzidos alguns movimentos característicos de algumas emoções. Este nível introdutório também permite uma aproximação entre a criança e o avatar tendo como objetivo também a criação de uma ligação emotiva entre ambos tornando o jogo mais interessante para o utilizador.

Para este cenário, foram predefinidos treze movimentos faciais diferentes incluindo alguns movimentos relacionados com a posição da cabeça. Na Tabela 4 encontra-se esse conjunto de movimentos básicos. Alguns destes movimentos faciais estão presentes em expressões faciais que representam determinadas emoções. Na Tabela 4 pode ver-se também essa ligação entre movimentos e emoções.

Tabela 4 - Lista de movimentos cenário 1

<i>Movimento</i>	<i>AU</i>	<i>Presente em Emoção</i>	<i>Emoção</i>
<i>Abrir a boca</i>	AU25 e AU26	Sim	Surpresa
<i>Baixar cabeça</i>	AU54	Não	—
<i>Levantar cabeça</i>	AU53	Não	—
<i>Inclinar cabeça Esquerda</i>	AU55	Não	—
<i>Inclinar cabeça Direita</i>	AU56	Não	—
<i>Olhar para a Esquerda</i>	AU51	Não	—
<i>Olhar para a Direita</i>	AU52	Não	—
<i>Levantar Sobrancelhas</i>	AU02	Sim	Surpresa
<i>Sorrir</i>	AU12	Sim	Felicidade
<i>Piscar Olhos</i>	AU45	Não	—
<i>Franzir Sobrolho</i>	AU04	Sim	Raiva
<i>Inclinar cabeça lado</i>	AU55 e AU56	Não	—
<i>Olhar para o lado</i>	AU51 e AU52	Não	—

Os vários movimentos vão sendo apresentados por uma ordem aleatória para evitar que o utilizador ao longo do tempo e com a possibilidade de várias utilizações do *HiZeca* execute os movimentos de uma forma mecânica e não de uma forma pensada.

Para o espelhamento da face do utilizador foi necessário usar os valores das *AU* fornecidas pelo *OpenFace*. Neste nível foram usados os valores do *Openface* no modo de regressão. Este modo classifica as *AU* entre os valores de 0 e 5, onde zero significa que a *AU* não está presente e cinco que o respetivo músculo está numa distensão ou contração elevada dependendo da *AU*. No motor de jogo *Unity 3D* a manipulação dos músculos da face está definida para variar de 0 a 100. Foi então necessário fazer um escalonamento dos valores de 0 - 5 para 0 - 100. Foi utilizada a seguinte equação (equação 1) de conversão:

$$N = \frac{W - \min}{\max - \min} \times (B - A) + A \quad (1)$$

Onde:

A - Valor mínimo do *OpenFace*

B - Valor máximo do *OpenFace*

W - Valor a converter

Apesar de idealmente os valores do *OpenFace* variarem de 0 a 5, isto não acontece em ambiente real podendo variar com o local onde está a ser feita a leitura dos valores das *AU*. Como tal, e para tornar o sistema mais otimizado foram feitos testes ao *OpenFace* e criada uma base de dados com todos os valores recolhidos. Para a criação desta base de dados foram utilizados 32 adultos com idades entre os 18 e 30 anos e 11 crianças com idades entre os 6 e 9 anos. Os testes com as crianças foram feitos em ambiente escolar e os testes com adultos foram feitos em laboratório [35]. De referir que esta base de dados foi criada noutra projeto tendo sido usada no presente projeto também [35].

Todos os dados relativos ao percurso feito neste nível são guardados num ficheiro de texto para que os pais, professores, educadores ou outros responsáveis pela criança possam analisar o desempenho da criança a nível de respostas certas, erradas, tempos de resposta ou até mesmo tempos limites ultrapassados. Na Figura 13 podemos ver o tipo de dados gravados e como são estruturados na sua gravação. São definidas três colunas com o nome do movimento, resposta e tempo de resposta. Foi implementado um tempo limite de resposta para cada movimento a fim de evitar que o jogador fique bloqueado num movimento e não possa avançar. Nesses casos passando esse tempo o jogo avança para o próximo movimento e regista que o tempo limite foi atingido (*Timeout*).

NIVEL 1

Movimento	Resposta Correta	Tempo de Resposta
Baixar Cabeça	Sim	00:00:02
Levantar a Cabeça	Sim	00:00:03
Piscar olhos	Sim	00:00:09
Inclinar a Cabeça Esq	Sim	00:00:03
Inclinar a Cabeça lad	Sim	00:00:12
Sorrir	Não	TimeOut
Inclinar a Cabeça Dir	Sim	00:00:09
Abrir a boca	Não	TimeOut
Lev/ Sobrancelha	Sim	00:00:02
Olhar para a Esq	Sim	00:00:02
Olhar para a Lad	Sim	00:00:06
Franzir sobrolho	Sim	00:00:04
Olhar para a Dir	Sim	00:00:02

Figura 13 - Dados gravados cenário 1

Cenário 2 – Treino de Expressões faciais

Neste segundo cenário é feita uma introdução às emoções. É neste nível que são treinadas várias expressões faciais associadas a diversas emoções. Para isso é pedido ao utilizador que execute algumas expressões. Para facilitar, o avatar executa repetidamente a expressão facial associada à emoção em questão. Contrariamente ao cenário anterior, aqui o avatar serve como modelo para a criança imitar. Estamos assim num contexto de imitação diferida. Segundo o psicólogo *Jean Piaget* a imitação diferida acontece quando o ser humano tem um comportamento ou um conjunto de comportamentos, que já observou no passado, no presente [36]. O avatar tem aqui um papel importante pois serve de ajuda para esclarecer o utilizador qual a expressão certa a fazer perante cada emoção. No esquema da Figura 14 pode ver-se o fio condutor deste cenário.

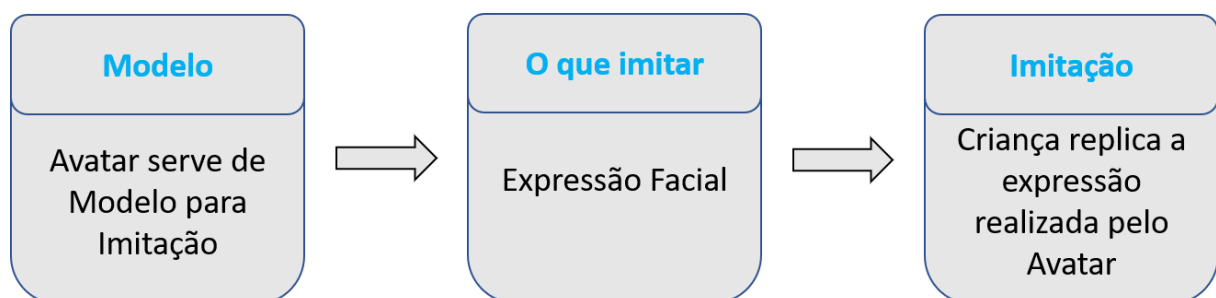


Figura 14 - Imitação diferida

Como referido atrás, neste contexto o avatar é o modelo a imitar. A criança, em caso de dúvida ou confusão na distinção da expressão facial correta pode observar o avatar e imitá-lo.

Neste cenário são treinadas cinco expressões faciais associadas a seis diferentes emoções: felicidade, tristeza, raiva, nojo, medo e surpresa. Para cada emoção, a criança terá de executar cinco vezes a expressão facial respetiva. À imagem do nível anterior a ordem de aparecimento das emoções é aleatória para evitar a memorização da ordem caso o nível seja utilizado várias vezes pelo mesmo utilizador. À medida que a criança vai efetuando a expressão correta, a imagem do avatar vai-se afastando tornando-se mais pequena, até desaparecer (Figura 15).

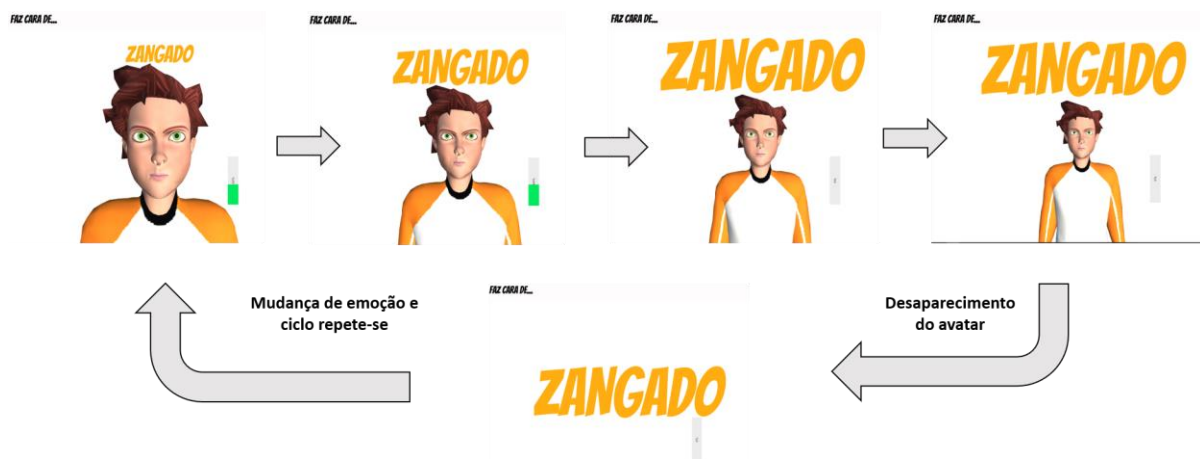


Figura 15 - Ciclo de jogo Cenário 2

Na Figura 15 pode ver-se o ciclo de funcionamento do cenário 2. Para as cinco emoções o tamanho do avatar vai sendo diminuído e a etiqueta com o nome da respetiva emoção aumenta até ao último ponto onde apenas permanece a etiqueta. Com este estilo de jogo pretende-se que no início ocorra um momento de aprendizagem ou associação de expressões a emoções e no último estado a criança possa ser capaz de efetuar a expressão facial sem o apoio visual do avatar.

Neste nível o *Openface* envia constantemente uma mensagem para o *Unity* com a expressão atual da face do utilizador. Para a validação de cada expressão facial o utilizador terá de permanecer um tempo predefinido efetuando permanentemente essa mesma expressão. Uma barra de controlo foi colocada estrategicamente (Figura 16) para que o utilizador controle exatamente se a sua expressão está correta e o tempo de resposta.



Figura 16 - Barra de controlo de resposta

Todos os dados relativos ao percurso feito neste nível são também guardados no ficheiro de texto à semelhança do cenário anterior. Na Figura 17 pode ver-se o tipo de dados gravados e como são estruturados na sua gravação. São definidas três colunas com o nome da emoção, resposta e tempo de resposta. Foi também estabelecido um tempo limite de resposta para cada emoção a fim de evitar o jogador ficar bloqueado na expressão de alguma emoção. Em cada emoção temos a gravação dos cinco passos intermédios desde o início, onde estão presentes a etiqueta e o avatar, até ao fim, onde permanece apenas a etiqueta identificadora da emoção. Com este tipo de gravação, os responsáveis pela criança podem perceber se o desempenho da criança aumenta ou não consoante os tempos de resposta. Para além disso podemos também identificar qual a emoção em que a criança teve mais facilidade de expressar e qual a emoção ou expressão que demorou mais tempo a ser executada.

NIVEL 2		
Emoção	Resposta Correta	Tempo de Resposta
TRISTE 1	Sim	00:00:03
TRISTE 2	Sim	00:00:08
TRISTE 3	Sim	00:00:12
TRISTE 4	Sim	00:00:10
TRISTE 5	Sim	00:00:02
ZANGADO 1	Sim	00:00:03
ZANGADO 2	Sim	00:00:03
ZANGADO 3	Sim	00:00:03
ZANGADO 4	Sim	00:00:04
ZANGADO 5	Sim	00:00:09
NOJO 1	Sim	00:00:13
NOJO 2	Sim	00:00:03
NOJO 3	Sim	00:00:04
NOJO 4	Sim	00:00:02
NOJO 5	Sim	00:00:02
CONTENTE 1	Sim	00:00:09
CONTENTE 2	Sim	00:00:14
CONTENTE 3	Sim	00:00:05
CONTENTE 4	Sim	00:00:05
CONTENTE 5	Sim	00:00:02
MEDO 1	Sim	00:00:03
MEDO 2	Sim	00:00:03
MEDO 3	Sim	00:00:02
MEDO 4	Sim	00:00:03
MEDO 5	Sim	00:00:04
SURPRESO 1	Sim	00:00:03
SURPRESO 2	Sim	00:00:28
SURPRESO 3	Não	TimeOut
SURPRESO 4	Sim	00:00:02
SURPRESO 5	Não	TimeOut

Figura 17 - Dados gravados cenário 2

Cenário 3 – Modo História

Neste terceiro e último nível foi implementado o modo história. Inicialmente este modo de jogo não foi estabelecido nos objetivos principais tendo sido propostos apenas os primeiros dois níveis. No entanto depois do desenvolvimento dos cenários anteriores surgiu a oportunidade de desenvolvimento deste cenário. Como referido, neste nível foi implementado o modo história onde a criança é convidada a ouvir um conjunto de 15 histórias associadas a emoções. Estas 15 histórias foram criadas noutra projeto envolvendo o robô ZECA e o modo história[37]. As emoções utilizadas neste cenário são as cinco utilizadas no cenário anterior: felicidade, tristeza, raiva, nojo e medo. Para cada uma destas emoções existem um conjunto de três histórias diferentes. A seguir (Figura 18) apresenta-se um esquema relativo ao nível 3.

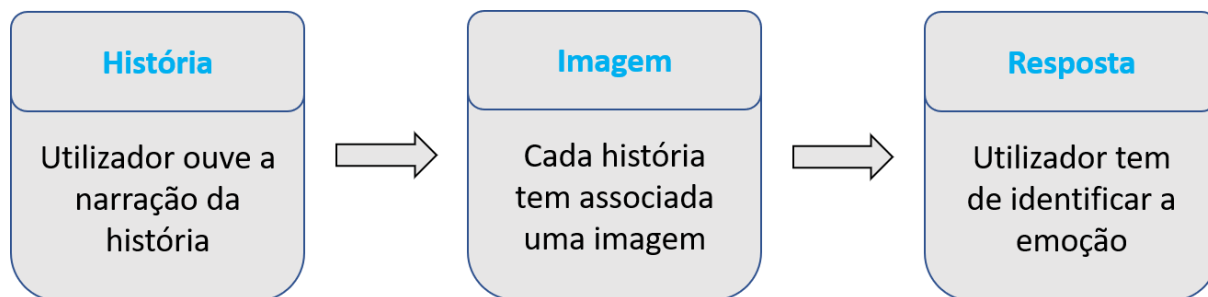


Figura 18 - Modo história

No esquema da Figura 18 pode ver-se o fio condutor de funcionamento do nível 3. Inicialmente a história é narrada e é apresentada a imagem representativa onde o ZECA avatar está presente e espelha uma determinada emoção. No final da narração da história a criança ou jogador terá de ser capaz de identificar a emoção presente.

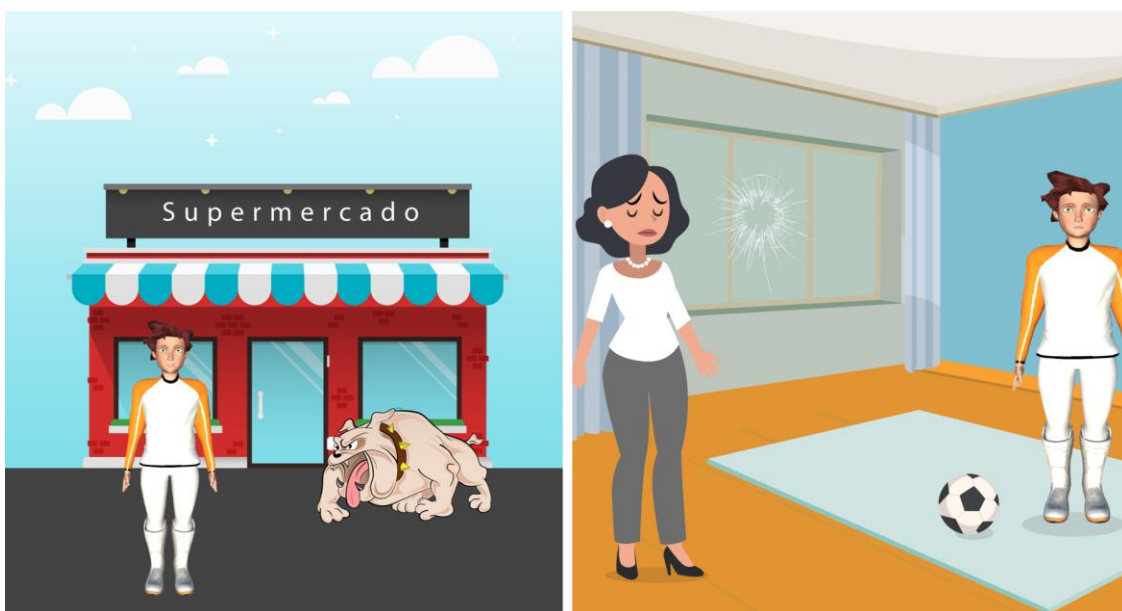


Figura 19 - Imagens exemplo do modo história

No modo história para além da narração existem também imagens (Figura 19) representativas da história e da emoção. Em todas elas o ZECA avatar está presente e espelha na sua face a expressão facial associada à respetiva ao seu estado emocional. A ordem de aparecimento das histórias é aleatória à semelhança dos cenários anteriores.

Com este modo o objetivo é fazer com que a criança ouça a narração da história, observe a imagem apresentada e em seguida identifique qual a emoção presente. Para essa identificação é apresentado o menu da Figura 20.

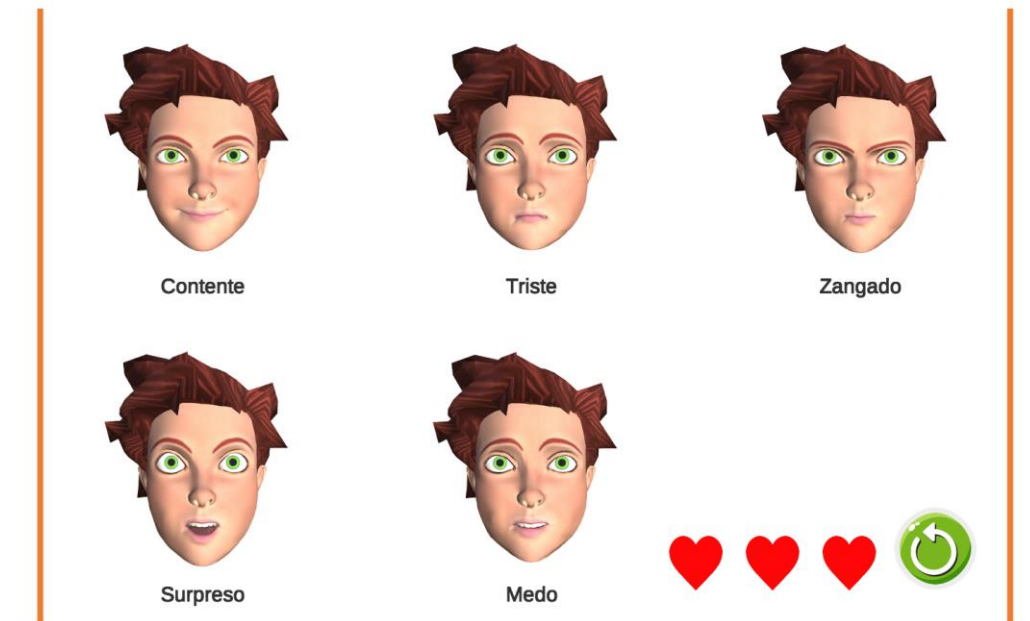


Figura 20 - Menu identificação da emoção

No menu da Figura 20 é apresentada a face do avatar expressando as cinco emoções possíveis. Para além da face é também acrescentado um rótulo em cada cara para ajudar a reconhecer o nome da emoção, em caso de dúvida. No final da narração de cada história o menu em questão é apresentado. Foi implementado um sistema de repetição de história caso a criança necessite de ouvir ou ver outra vez a história ou imagem. Essa funcionalidade pode ser acedida através da seleção do botão verde presente na Figura 20. Foi também implementado um sistema de vidas de jogo representado com a presença dos três corações presentes na Figura 20. Na eventualidade da resposta dada pelo utilizador ser uma resposta incorreta, o número de vidas é decrementado e o botão da resposta dada desaparece. Vejamos a seguinte sequência de imagens (Figura 21).

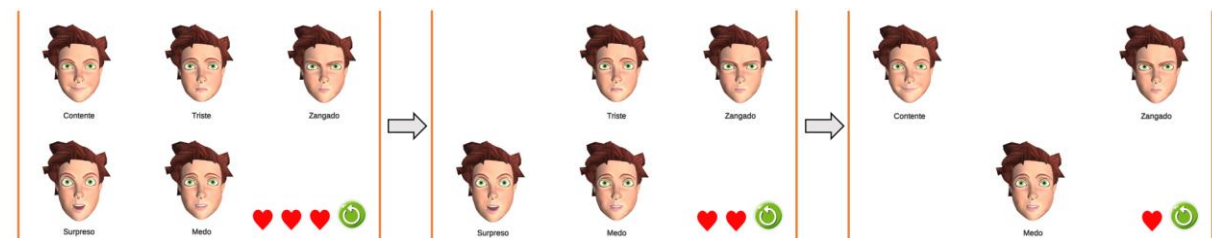


Figura 21 - Sequência de respostas erradas

No caso de serem esgotadas as três vidas possíveis o sistema passa para a história seguinte registrando que na história em questão o utilizador esgotou as vidas.

Neste nível é então feito um teste geral sobre o jogo pois é aqui que a criança pode demonstrar se a utilização do *HiZeca* teve ou não impacto na distinção e identificação de emoções. Com o conjunto de histórias tem-se uma melhor percepção sobre a capacidade de raciocínio e de associação que a criança possui acerca do reconhecimento de expressões faciais ligadas as emoções.

Neste último nível também são gravados dados importantes acerca do desempenho da criança ao longo do nível. Na Figura 22 podemos ver a estrutura de dados guardados. Neste documento de texto são gravadas três colunas com o número da história, a resposta e o tempo de resposta, á semelhança dos cenários anteriores. Neste caso podemos verificar o número de respostas erradas em cada história e se ouve alguma história onde a criança esgotou as vidas.

NIVEL 3		
História	Resposta Correta	Tempo de Resposta
Historia 2	Sim	00:00:02
Historia 4	Errado	
Historia 4	Sim	00:00:06
Historia 6	Sim	00:00:04
Historia 8	Errado	
Historia 8	Errado	
Historia 8	Errado	
Historia 10	Sim	00:00:02
Historia 12	Sim	00:00:00
Historia 14	Sim	00:00:00
Historia 1	Errado	
Historia 1	Errado	
Historia 1	Errado	
Historia 3	Sim	00:00:01
Historia 5	Sim	00:00:00
Historia 7	Sim	00:00:01
Historia 9	Sim	00:00:00
Historia 11	Sim	00:00:04
Historia 13	Sim	00:00:01
Historia 15	Sim	00:00:01

Figura 22 - Dados gravados cenário 3

5. ARQUITETURA E DESIGN DO HiZECA

No presente capítulo são apresentados os métodos e técnicas utilizados no desenvolvimento do *HiZeca*. São apresentados fluxogramas de funcionamento dos vários níveis de jogo, métodos de programação utilizados e toda a estrutura base do jogo. Para além disso são também apresentados os menus presentes no jogo e suas funções. Mostra-se também todas as funções específicas e principais do jogo e suas funcionalidades. Por fim são também exibidas todas as estruturas de objetos criados no *Unity 3D*, criação de animações e alguns detalhes sobre algumas implementações do motor de jogo utilizado.

5.1 Menus iniciais e menus comuns

No arranque do HiZeca é apresentado um menu (Figura 23) onde o utilizador tem de introduzir o seu nome. Este menu serve como identificador da criança para mais tarde ser criado o ficheiro de texto com todas as informações sobre o seu desempenho no jogo. Depois de introduzido o nome, o jogo apresenta o menu principal (Figura 24) onde o utilizador pode escolher o nível a jogar.



Figura 23 - Menu identificador do utilizador



Figura 24 - Menu inicial

No menu da Figura 24 podemos ver o aspeto gráfico do menu principal. Neste menu o utilizador tem a oportunidade de escolher o nível desejado. Não foi imposta nenhuma ordem obrigatória de jogo entre os diferentes níveis. Isso deve-se ao facto de o HiZeca ser um jogo construtivo e de apoio ao desenvolvimento de capacidades cognitivas. Nesse sentido no caso de o responsável pela criança optar por definir a ordem que achar mais oportuna pode fazê-lo. Para além disso permite também à criança repetir várias vezes o mesmo nível sem ter de passar pelos níveis anteriores. No esquema da Figura 25 apresenta-se um esquema geral de funcionamento do menu principal.

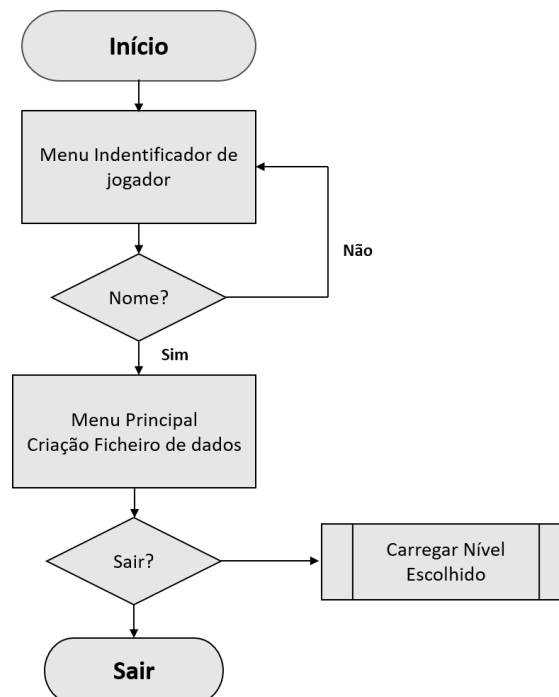


Figura 25 - Fluxograma menu principal

Para além dos menus apresentados anteriormente, foram criados menus comuns a todos os níveis. No início de cada nível é feita uma apresentação do mesmo através de uma narração. No final dessa narração a criança é solicitada para ajudar o ZECA avatar (Figura 26), tendo esta que responder “sim” ou “não”.

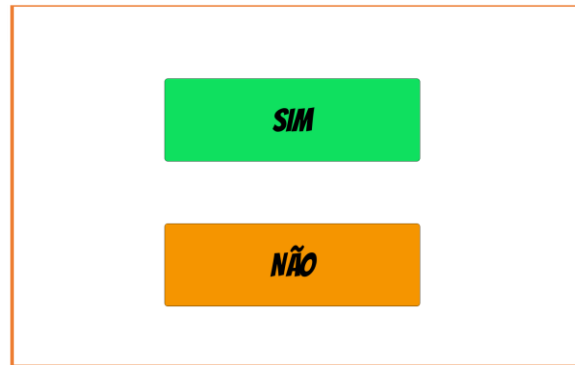


Figura 26 - Menu início de cada nível

No caso de a criança selecionar o botão “sim” o nível é iniciado. Caso contrário o jogo volta ao menu principal e a criança pode escolher o nível desejado.

À semelhança de todos os jogos, foi também desenvolvido um sistema de pausa de jogo. No caso de ser solicitada a pausa do jogo é apresentado o menu da Figura 27.



Figura 27 - Menu pausa

Neste menu são apresentadas duas possibilidades: continuar o jogo ou voltar ao menu. Se este menu for chamado, o tempo de resposta não é afetado pois este inclui apenas o tempo útil de jogo e não o tempo total. No caso de ser selecionada a opção de voltar ao menu os dados do jogo são guardados e é apresentado o menu principal.

Por fim, no final de cada nível é apresentado um menu de fim de nível onde é feita uma narração de agradecimento pela ajuda. Esse menu está presente na Figura 28.



Figura 28 - Menu fim de nível

No menu anterior são também apresentadas duas opções de escolha: jogar outra vez ou voltar ao menu. Se for escolhida a opção de jogar outra vez o jogo regista no ficheiro de dados que vai ser iniciado um novo jogo e iniciado de novo o nível em questão.

5.2 Funções comuns a todos os níveis

Foi desenvolvido um conjunto de funções genéricas utilizadas nos vários níveis para fazer a gestão do jogo. No início de cada nível é apresentado o menu da Figura 26 e foi necessário criar uma função para gerir o momento do início do jogo para controlar depois os dados guardados. Na Figura 29 encontra-se o esquema do funcionamento da função chamada no início de cada nível.

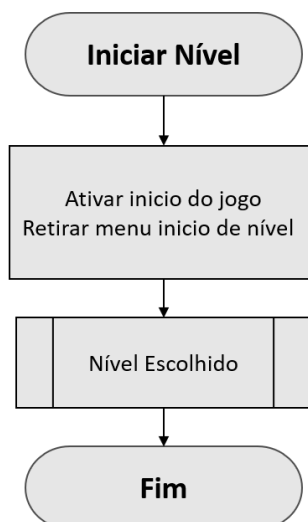


Figura 29 - Fluxograma função do início do nível

Do mesmo modo, para o menu de pausa de jogo da Figura 30 foi implementada uma função para gerir a possibilidade de continuar a jogar. Visto que para os tempos de resposta foi estabelecido que apenas contava o tempo ativo de jogo, neste menu foi necessário contabilizar o tempo utilizado no modo de pausa para não afetar os tempos de resposta.

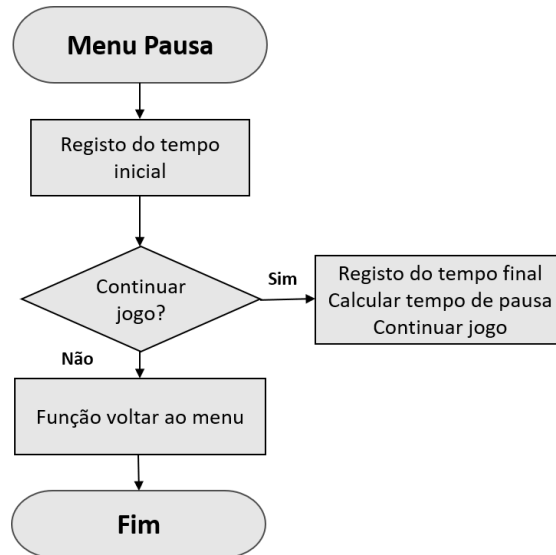


Figura 30 - Fluxograma função continuar jogo

Como podemos verificar no esquema da Figura 30, quando o jogo entra em modo pausa, o tempo atual é registado. No final, caso seja para continuar o jogo é registado outra vez o tempo atual e subtraído ao inicial. Esta função está presente em todos os níveis. No caso de o utilizador escolher a opção de voltar para o menu principal então são gravados os dados no ficheiro e carregado o menu principal.

Na Figura 31 apresenta-se o esquema relativo a função de jogar outra vez. Neste caso, no ficheiro de dados é registado que este modo foi ativo (para organização de dados), é gerada uma nova ordem aleatória (ordem dos movimentos no nível 1 por exemplo) e todas as variáveis necessárias são inicializadas a zero.

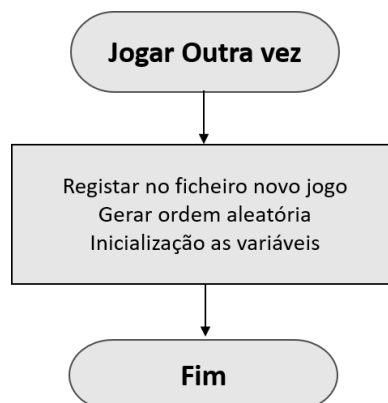


Figura 31 - Fluxograma função jogar outra vez

Como referido anteriormente os dados estão a ser guardados num ficheiro de texto. Como tal sempre que um nível é iniciado esse ficheiro é aberto e todos os dados vão sendo acrescentados aos dados que possam existir. Quando o jogo volta para o menu esse ficheiro é fechado.

Por fim quando o jogo é desligado é criado um ficheiro com o nome do jogador e a hora de início. Em seguida são copiados todos os dados do ficheiro *standard* para o ficheiro criado. Com isto os responsáveis pela criança podem identificar facilmente qual o ficheiro que contém os dados da sua criança. Na Figura 32 temos um esquema elucidativo do mecanismo de fecho do jogo.

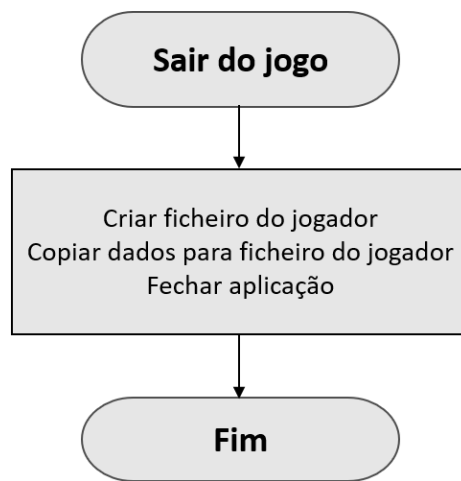


Figura 32 - Fluxograma função sair do jogo

5.3 Cenário 1

Neste subcapítulo apresentar-se a parte técnica de desenvolvimento do nível 1. Serão apresentados alguns dos métodos utilizados bem como esquemas de algumas funções específicas do Unity 3D, entre outras funcionalidades utilizadas. É apresentado também o fluxograma de funcionamento geral deste nível.

Animação do Avatar

A atualização da face do avatar é feita em tempo real onde podem ocorrer grandes variações de movimentos feitos pelo utilizador. Para além disso a taxa de atualização da face do avatar ocorre a cada ciclo de jogo criando imperfeições visualmente perceptíveis. Por forma a suavizar os movimentos foi utilizado o filtro de *smoothing* [38]. Este filtro calcula pontos intermédios através do cálculo da média entre pontos distantes evitando que haja saltos numéricos muito grandes (Figura 33).

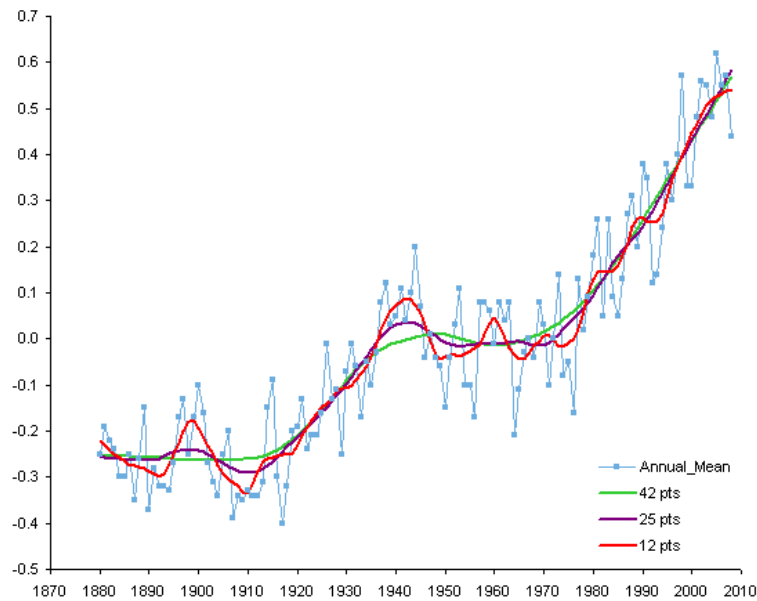


Figura 33 - Filtro de Smoothing [38]

Na Figura 33 temos um exemplo da aplicação do filtro de *Smoothing* com variação do número de pontos intermédios. Para o valor de 12 pontos o filtro apresenta algumas oscilações significativas, provocadas pela variação elevada de pontos do sinal que está a ser filtrado. À medida que vamos aumentando o número de pontos intermédios vamos diminuindo as oscilações criando um sinal mais estável e com oscilações mais suaves. Este filtro é muito utilizado para a redução do ruído de sinais.

Neste projeto este filtro não serviu para redução de ruído, mas sim para a suavização de movimentos. O *Unity 3D* oferece uma função predefinida deste filtro:

- *SmoothDamp()*

Para utilizar esta função temos de passar como parâmetros apenas o valor inicial, o final e o tempo que o filtro demora desde o tempo inicial até ao final. A função apresenta ainda outros parâmetros. Assim com a ajuda desta função a atualização da face do avatar é feita com os valores extraídos do filtro evitando variações bruscas que são perceptíveis visualmente.

Arquitetura e Design

Para ser possível a manipulação da face do avatar foi necessário perceber quais as funções existentes para esse fim. Neste cenário foram utilizadas essencialmente três funções disponibilizadas pelo *Unity 3D*.

Mas antes de falar nessas funções é importante perceber como está criado todo o suporte no *Unity 3D* para permitir a animação do avatar.

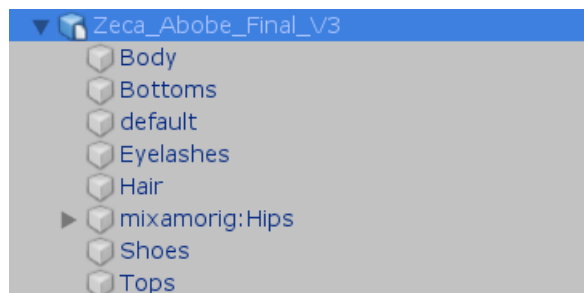


Figura 34 - Árvore de objetos do avatar

Na Figura 34 pode ver-se como está estruturado o modelo 3D do avatar no *Unity 3D*. Para todos os avatares é criada uma árvore de objetos. É através dessa árvore de objetos que percebemos como é constituído cada avatar e quais os objetos que temos disponíveis para o animar. No entanto, para este jogo, o que mais importa é o “*mixamorig:Hips*” e o “*Body*”. O “*mixamorig:Hips*” é gerado no *Mixamo* e é ele que permite aceder ao corpo do avatar para manipular a cabeça, pernas, e braços tal como podemos ver na Figura 35.



Figura 35 - Estrutura de objetos do *Mixamorig:Hips*

No separador “*Body*” é onde pode aceder-se a parte das *blendshapes*. Os objetos 3D são compostos por malhas que é o que define a sua estrutura. *Blendshapes* refere-se a malha presente na face do avatar e é através da deformação desta malha que conseguimos obter as expressões faciais e todos os movimentos da cara.

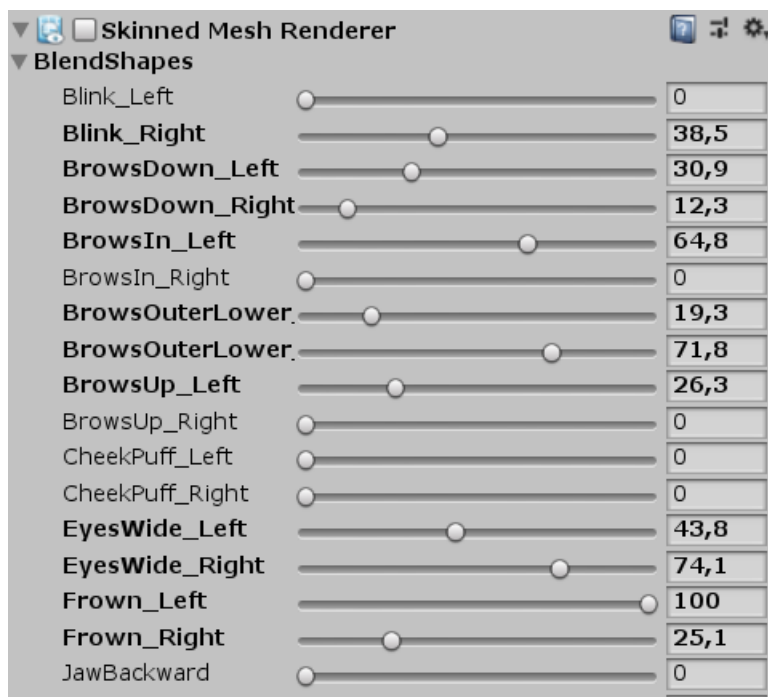


Figura 36 - BlendShapes Unity 3D

No separador da Figura 36 tem-se acesso a 29 músculos faciais diferentes. Estes movimentos podem variar de 0 a 100 como já referido noutros capítulos. Com a variação destes valores ao longo do tempo criamos movimentos realistas perceptíveis.

Para aceder a estes valores foram utilizadas funções disponíveis no *Unity 3D*:

- `blendshapesRenderer.sharedMesh.GetBlendShapeIndex()`
- `blendshapesRenderer.SetBlendShapeWeight();`
- `blendshapesRenderer.sharedMesh.GetBlendShapeIndex()`

Através deste conjunto de funções pode aceder-se a cada um dos músculos com seu índice e manipular o seu valor constantemente, criando movimento. Com a função “set” escrevemos o valor. Com a função “get” podemos ler o seu valor atual.

Em seguida será explicado como foi feito o envio dos dados do *Openface* para o *Unity 3D* e como foi feito o tratamento dessa informação.

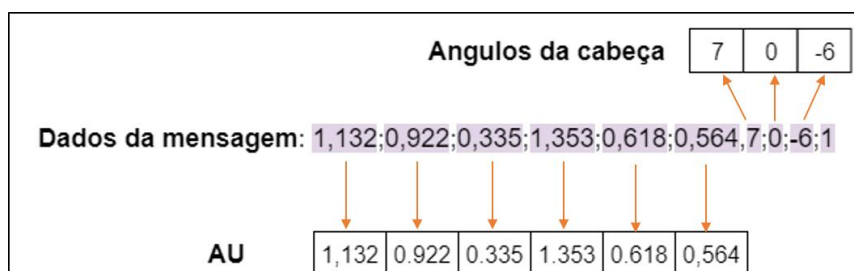


Figura 37 - String de dados enviados

Todos os valores das diferentes AU retirados e processados pelo *Openface* são enviados sob a forma de uma mensagem (Figura 37). Nessa mensagem estão presentes os valores das 18 *AU* diferentes, os ângulos da cabeça e um dígito de confirmação da presença de uma face. Um algoritmo desenvolvido é responsável pela separação dos diferentes valores, guardando-os em vetores diferentes para que possam ser usados mais tarde. Foram criados dois vetores diferentes. Um deles para guardar os dados das *AU* e o outro os ângulos. Esses valores são utilizados para fazer a atualização da face do avatar em tempo real permitindo que o avatar imite o utilizador. Na Figura 38 apresenta-se um fluxograma geral de funcionamento do cenário 1.

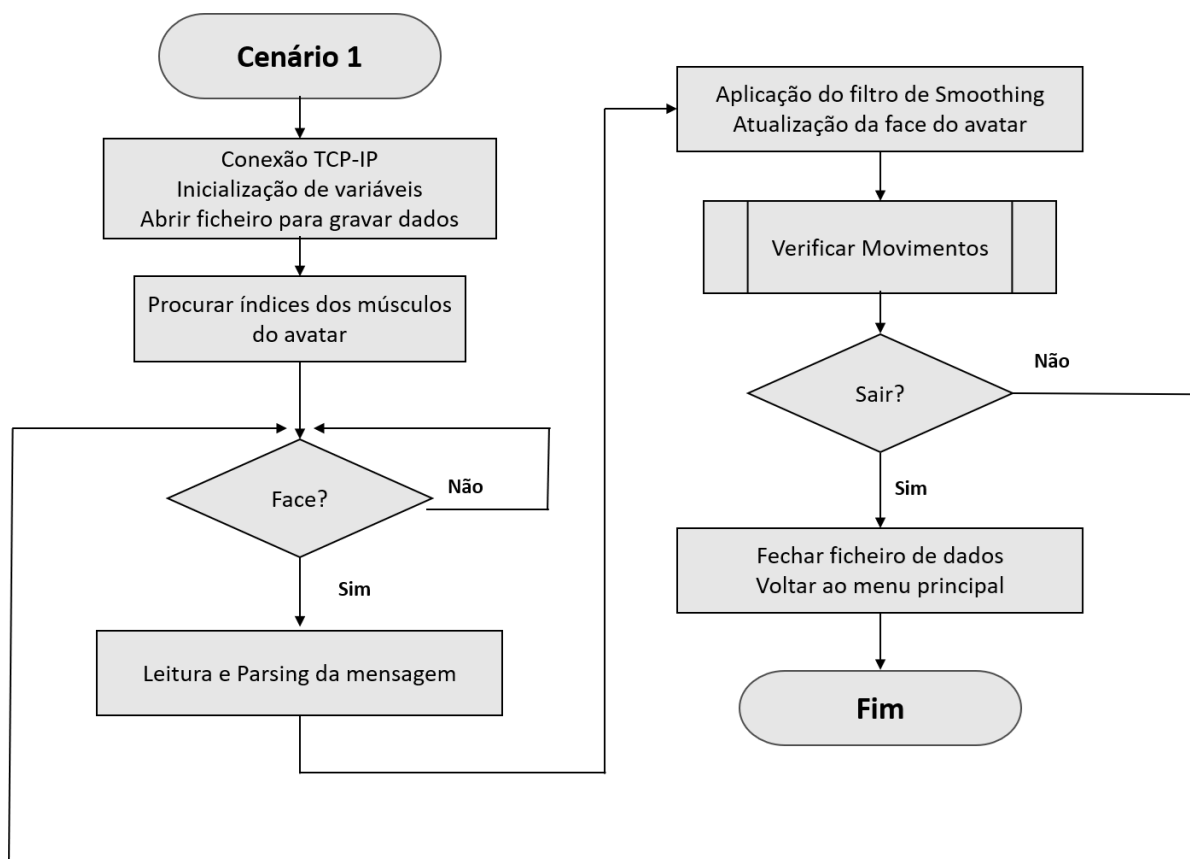


Figura 38 - Fluxograma geral Cenário 1

No início deste nível é feita a conexão entre o *Openface* e o *Unity3D*. Esta conexão é feita através do protocolo *TCP-IP*. Todas as variáveis necessárias são inicializadas e é feito também uma procura de todos os objetos necessários presentes no *Unity 3D*. Depois da conexão feita e da inicialização de variáveis é feita uma verificação da presença ou não da face. Se nenhuma face estiver a ser detetada não é ativa a atualização da face do avatar através dos movimentos do utilizador, nem é verificado

qualquer tipo de movimento. Caso contrário então é feito o *parsing* dos dados enviados sob a forma de *string* e é feita a replicação de movimentos no avatar. Uma sub-rotina (Figura 39) é chamada e é aí que são lançados todos os movimentos e é feita toda a verificação de movimentos certos. Por fim verifica-se o fim de nível.

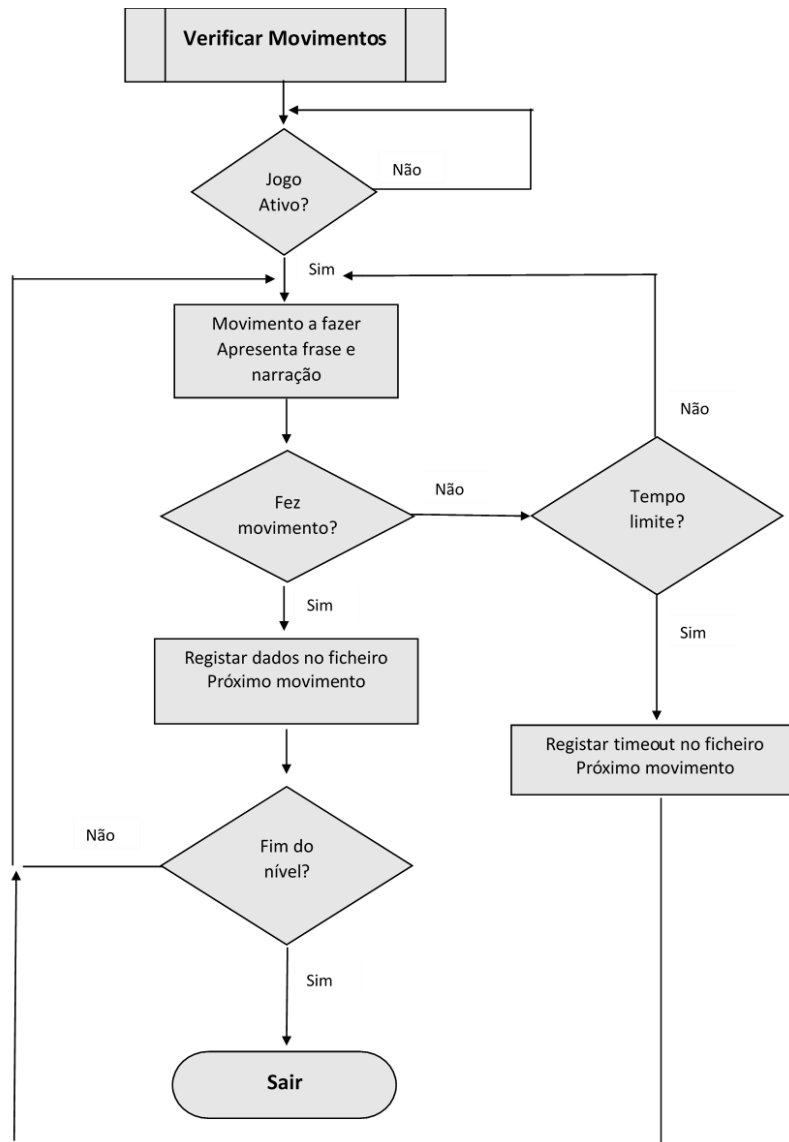


Figura 39 – Sub-rotina para verificação dos movimentos

No esquema da Figura 39 temos um fluxograma da função responsável pela verificação dos movimentos. Depois de ativo o jogo é lançado o primeiro movimento. A cada ciclo é verificado esse movimento através da posição atual do avatar. Se for detetado o movimento correto é registado o tempo de resposta e apresentado o seguinte movimento. Caso contrário é verificado o tempo limite imposto para cada movimento. Se esse tempo for ultrapassado é registado no ficheiro de dados essa informação e passa-

se para o próximo movimento. No final em caso de estarmos no final do nível o jogo volta ao menu principal.

Para a narração de cada frase foi necessário criar um ficheiro de áudio com uma gravação de voz. Na apresentação de cada movimento é reproduzido esse ficheiro de áudio. Na Figura 40 temos a lista completa de todas as falas disponíveis neste nível.

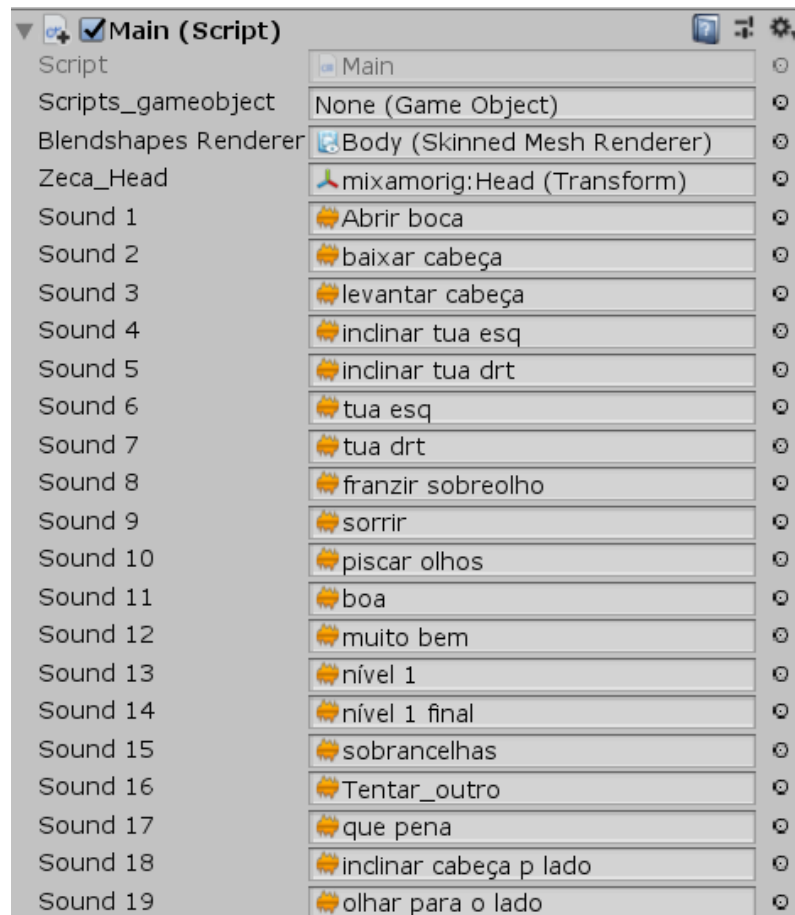


Figura 40 - Lista de falas cenário 1

A função utilizada para reproduzir estas falas foi a `source.PlayOneShot()`. Esta função, disponível nas bibliotecas do *Unity 3D*, reproduz uma vez o som recebido como parâmetro. Para a música de fundo foi criado um ficheiro de áudio separado e utilizada a função `source.Play()` configurada para repetir o som caso este chegue ao fim.

Em seguida temos a árvore de objetos existentes para garantir o funcionamento deste nível.

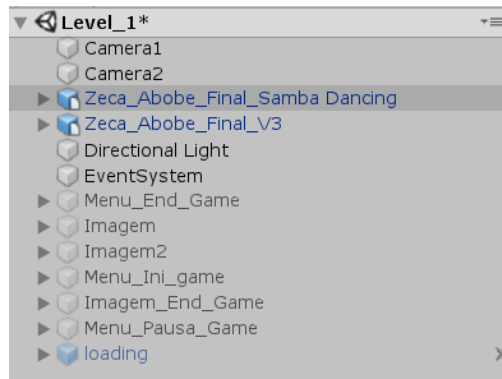


Figura 41 - Arvore de objetos cenário 1

Na Figura 41 temos então a listagem de todos os objetos que foram necessários para criar a jogabilidade e ambiente gráfico do cenário 1. Ao longo do nível os vários objetos vão sendo ativos ou não conforme a performance do utilizador e o desenrolar do próprio jogo.

5.4 Cenário 2

Neste subcapítulo apresenta-se a parte técnica de desenvolvimento do nível dois. São apresentados alguns fluxogramas representativos de como foi implementado este nível. Para além disso será explicado como foram criadas as animações das expressões faciais reproduzidas pelo avatar e como foi feita a leitura dos dados enviados pelo OpenFace.

Animações

Neste nível foi necessário criar seis animações diferentes respetivas às seis emoções presentes. Como tal foi utilizada a ferramenta de criação de animações disponível no *Unity 3D*.

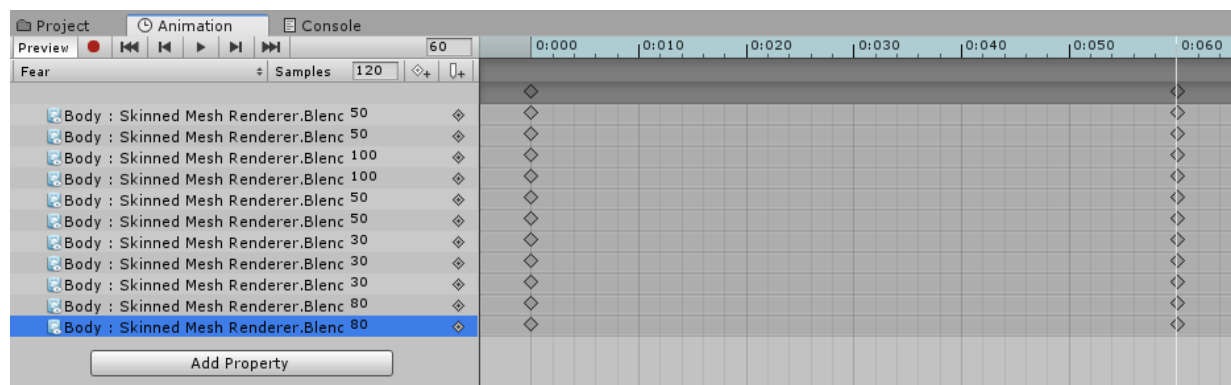


Figura 42 - Animação da expressão facial de medo

Estas animações são criadas através da alteração do valor dos respetivos músculos ao longo do tempo. No caso da expressão facial de medo (Figura 42) foram utilizados 11 músculos diferentes. Para criar os movimentos, todos os músculos são iniciados a zero e terminam num valor predefinido ao fim de um segundo, tal como podemos ver na Figura 42. Depois disso o valor é decrementado até atingir o valor inicial demorando também um segundo. Cada animação tem um tempo total de dois segundos sendo o primeiro para expressar a emoção e o segundo para voltar a expressão neutra. Estas animações foram criadas no *Unity 3D* por serem necessários poucos movimentos. Vejamos a Figura 43.

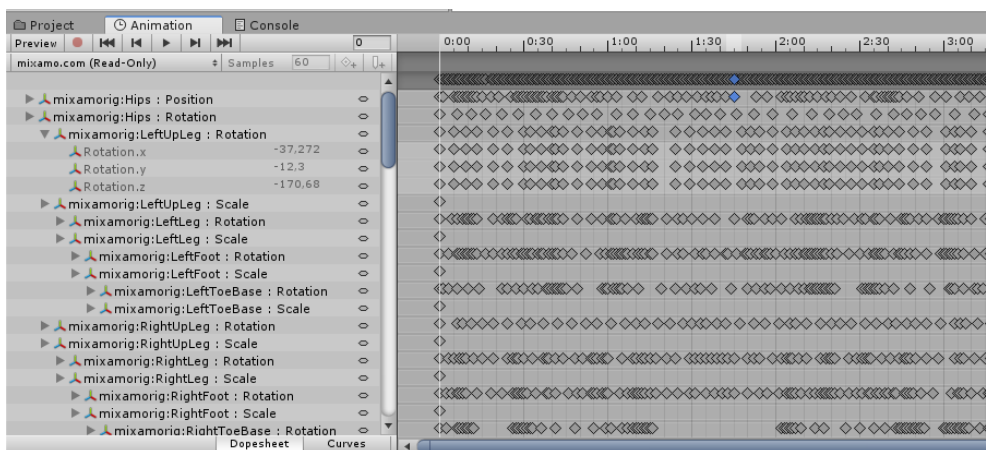


Figura 43 - Animação criada pelo Mixamo para dança do avatar

Na Figura 43 podemos verificar a animação gerada pelo *Mixamo* para a dança do avatar. Tal como é visível, para além de o número de músculos a movimentar ser elevado (não estando todos visíveis na imagem), é necessário definir muitos valores ao longo do tempo para obter movimentos realistas. Por esse motivo no caso de ser necessário criar animações envolvendo muitos movimentos e músculos recomenda-se a utilização de *software* auxiliares. No entanto, neste projeto as animações do corpo foram geradas pelo *Mixamo* tendo sido criadas apenas as seis animações responsáveis pelas seis expressões faciais pretendidas. Para o controlo das animações foram criadas variáveis para ativar e desativar. Para cada animação foi criada uma variável que manipulada por código ativa a execução da animação sempre que necessário. Quando for necessário a alteração da animação a executar pelo avatar basta apenas desativar a animação atual e ativar a variável da animação pretendida. Na Figura 44 temos o bloco de controlo das animações criadas. O bloco central a laranja verifica a ativação de animações. Caso alguma animação seja requerida este bloco ativa essa mesma animação estando esta definida para repetir a animação, até ser desativada. No momento em que a variável de controlo da animação em reprodução é desativada a animação é parada e o bloco fica à espera de uma nova ativação.

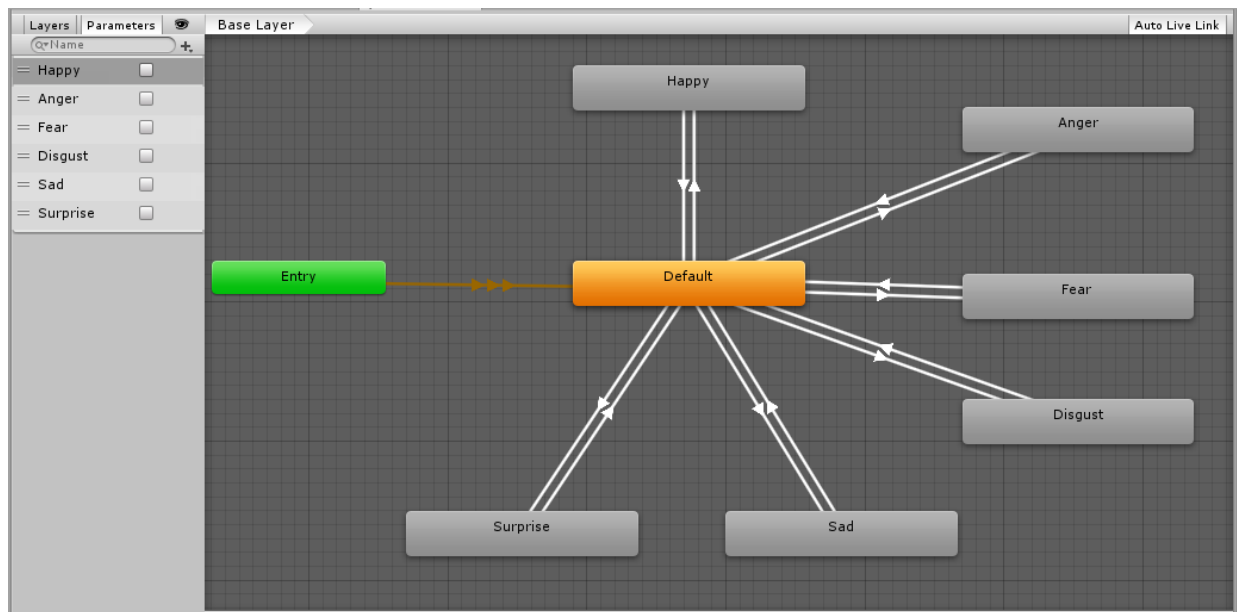


Figura 44 - Bloco de controlo das animações criadas

Arquitetura e Design

Na Figura 45 podemos verificar a árvore de objetos criados para o desenvolvimento do cenário 2. Como podemos verificar existem alguns objetos nomeadamente menus e avatares que foram utilizados no cenário 1 e importados para este cenário.



Figura 45 - Arvore de objetos cenário 2

No início do nível é feita uma inicialização de todas as variáveis, a abertura do ficheiro onde serão gravados os dados relativos ao desempenho da criança e é feita a conexão entre o *OpenFace* e o *Unity3D*. Em seguida é apresentado o menu inicial de nível. Depois de terminar a narração com a apresentação do que se pretende, a criança escolhe se quer jogar ou não. No caso positivo é gerada uma sequência

aleatória e é chamada a sub-rotina responsável pelo lançamento e verificação de expressões. No final do nível o ficheiro de dados é fechado e o jogo volta ao menu principal. Em seguida podemos ver o fluxograma de funcionamento geral do cenário 2 (Figura 46).

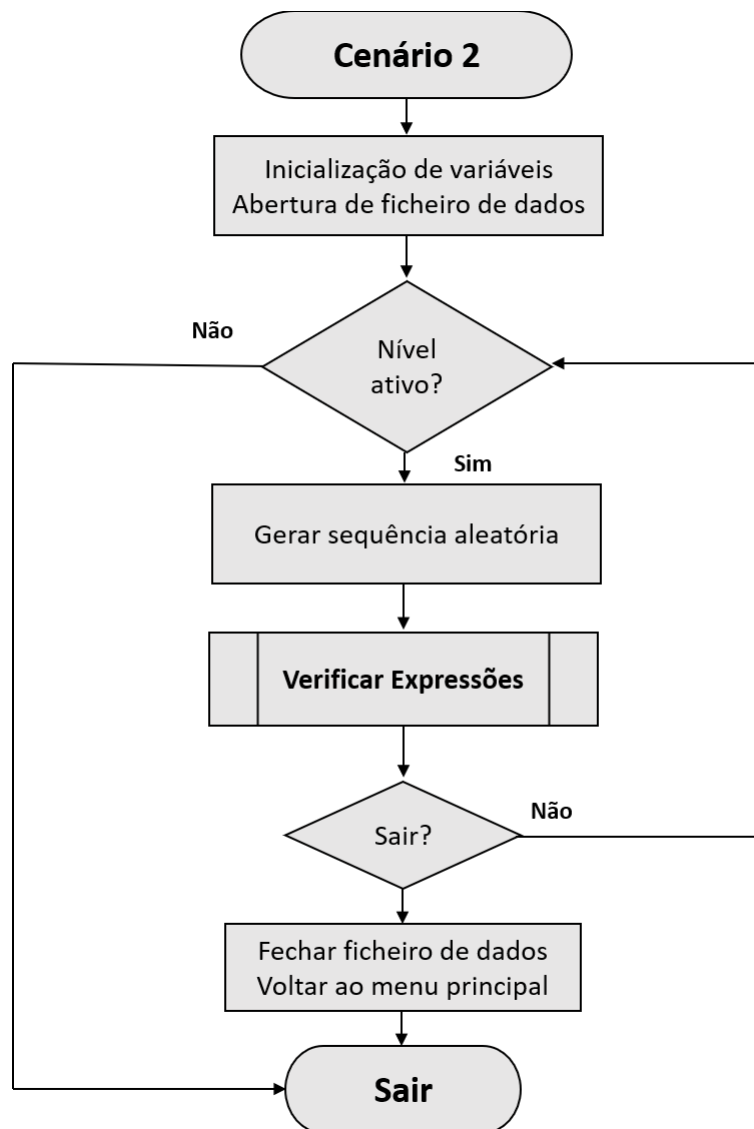


Figura 46 - Fluxograma geral Cenário 2

Antes de explicar o funcionamento da sub-rotina de verificação de expressões explicar-se-á que tipo de informação é enviada do *OpenFace* para o *Unity 3D*. O projeto do *Openface* responsável pela deteção de expressões faciais foi desenvolvido anteriormente no projeto *ZECA*. Neste projeto são detetadas 7 expressões diferentes incluindo a neutra. A cada ciclo de processamento do *OpenFace*, é enviado para

o *Unity 3D* um caracter identificador da expressão que está a ser detetada no momento. Na Figura 47 temos a receção dos dados enviados em tempo real.



Figura 47 - Dados recebidos no Unity 3D

Seleção de emoções

A numeração exemplificada no subcapítulo anterior serve como entrada para a sub-rotina que gere a verificação de expressões. No início da sub-rotina é verificada a expressão detetada consoante os números da seguinte Tabela 5. Consoante a deteção da seguinte numeração torna-se mais fácil a verificação da expressão sendo apenas necessário verificar se o número é o correto.

Tabela 5 - Numeração das emoções enviadas para o Unity 3D

INDICE	EMOÇÃO
0	Contente
1	Triste
2	Medo
3	Zangado
4	Surpresa
5	Nojo
6	Neutro

A seguir apresenta-se um fluxograma de funcionamento da sub-rotina de verificação de expressões (Figura 48). Quando esta sub-rotina é chamada, é feita uma inicialização das variáveis locais tais como variáveis de tempo por exemplo. Em seguida é narrada a frase que faz o pedido para a execução da expressão relacionada com a emoção em causa. Depois disso é verificada a cada ciclo se a expressão facial atual do utilizador é a pedida. No caso positivo uma variável auxiliar é incrementada até atingir um valor predefinido que dita que o utilizador permaneceu um determinado tempo a expressar a emoção

pedida. Depois disso é gravado no ficheiro de texto o tempo e ocorre um ajuste da posição do avatar e da etiqueta (avatar fica mais pequeno e a etiqueta maior). Repete-se este ciclo 5 vezes para cada emoção. Em seguida é verificado se o ciclo percorrido para cada emoção está terminado e no caso positivo a expressão a detetar é atualizada para a seguinte. Ao longo do ciclo é também verificado o tempo limite para executar a expressão pedida.

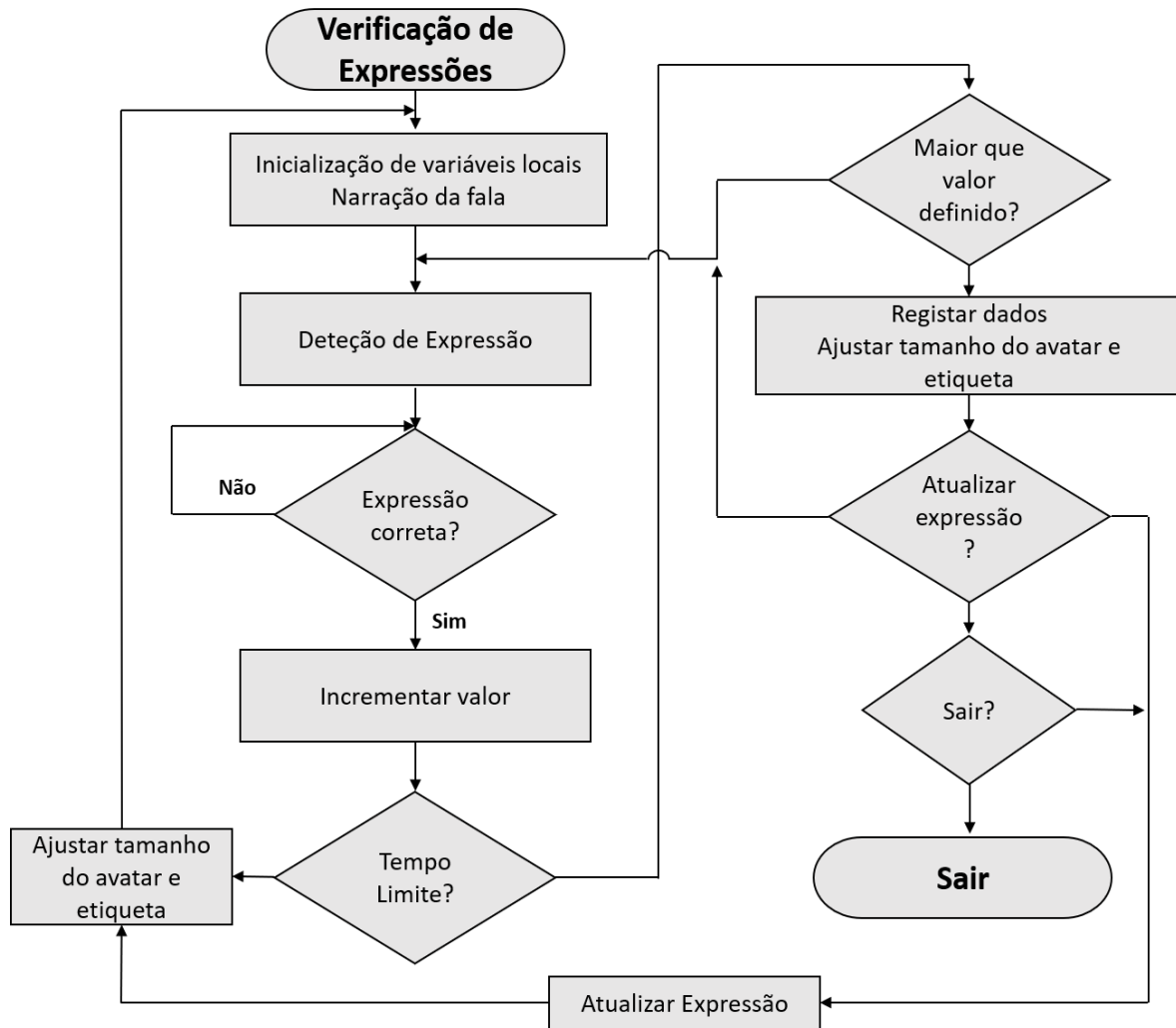


Figura 48 - Fluxograma Sub-Rotina verificação de expressões

5.5 Cenário 3

No terceiro e último cenário de jogo existem 15 histórias diferentes. O conjunto dessas 15 histórias estão disponíveis no anexo 1. Para cada história foi necessário criar dois objetos, um para controlar a narração da história, através de um ficheiro de áudio, e o outro para apresentar a imagem. No início deste cenário são inicializadas todas as variáveis e é feita a abertura do ficheiro de dados para escrita. Em seguida é gerada a ordem aleatória de histórias e apresentada a história. No final da narração da história é apresentado o menu de respostas. No caso de resposta dada ser a certa passamos para a próxima

história caso contrário é feita a atualização de vidas e retirado o botão do menu. No fim, se estiver no final do nível o ficheiro é fechado e o jogo volta ao menu principal se o utilizador não quiser jogar outra vez. A seguir (Figura 49) pode ver-se um fluxograma resumo sobre a implementação no nível 3.

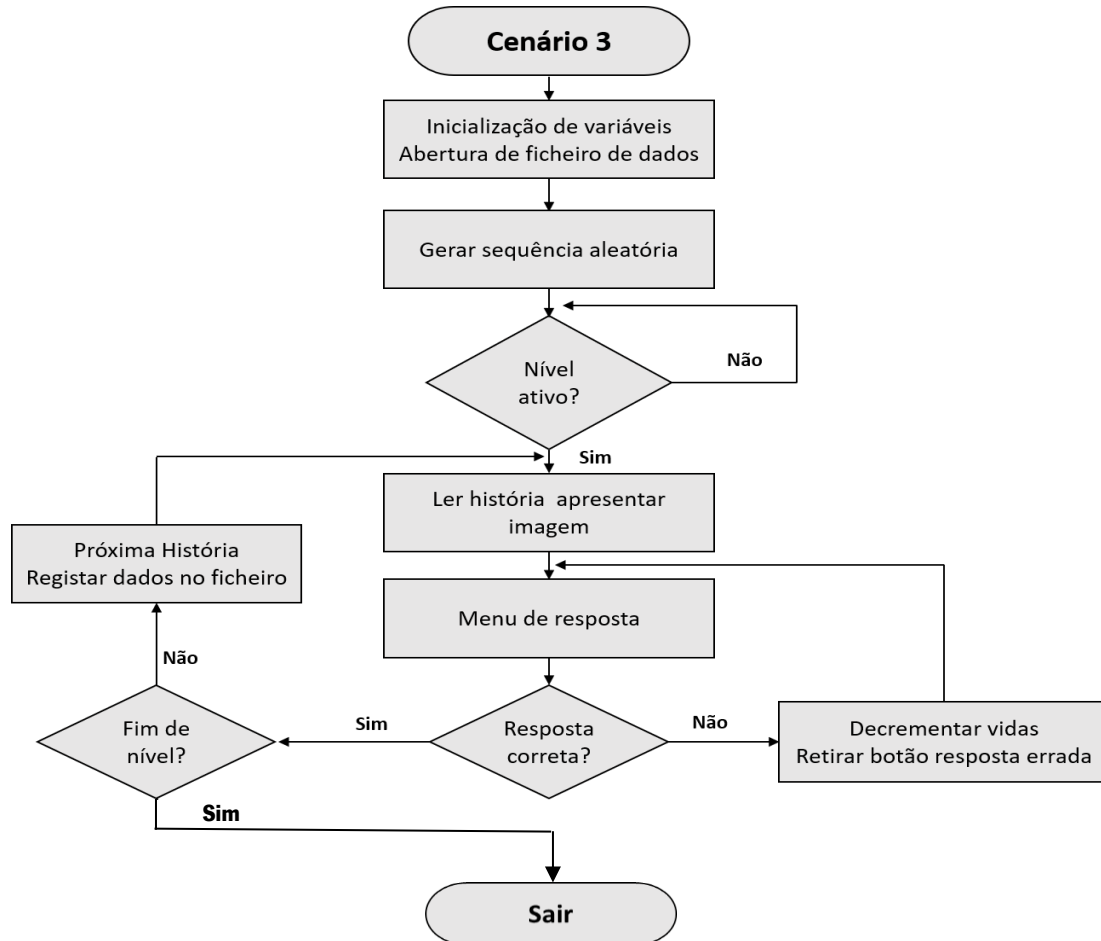


Figura 49 - Fluxograma geral cenário 3

5.6 HiZeca – Junção de módulos

Todos os cenários atrás referidos foram desenvolvidos em projetos separados. Depois de todos terminados foi feita uma junção de todos obtendo assim a versão final do jogo. Para se fazer a integração de todos os projetos, basta copiar a pasta de cada projeto para a pasta do jogo tal como podemos verificar na Figura 50.

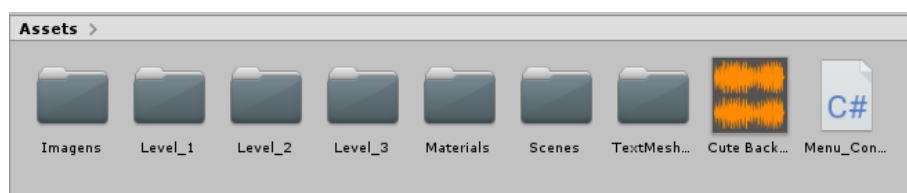


Figura 50 - Assets do HiZeca

Depois disso é necessário ter todos os ficheiros das cenas de cada cenário na pasta de cenas, para que o menu principal as possa chamar.



Figura 51 - Cenas disponíveis para o jogo

Por fim foi criado um script c# para fazer a gestão do menu principal a fim de carregar os diferentes níveis chamando as respetivas cenas. Para isso e como podemos ver no excerto de código da Figura 52 para cada nível basta carregar a respetiva cena através da função "*SceneManager.LoadScene*".

```
public void level_1()
{
    if (segundo > 2)
    {
        musica_fundo.Stop();
        loading.gameObject.SetActive(true);
        SceneManager.LoadScene(Level_1);
    }
}
```

Figura 52 - Excerto de código responsável por abrir o nível 1

6. RESULTADOS

Neste capítulo apresenta-se os resultados obtidos através dos testes realizados para verificar a robustez e fiabilidade do jogo. Foram feitos testes laboratoriais envolvendo adultos e testes em ambiente escolar com crianças tipicamente desenvolvidas e com crianças com PEA. Para além disso apresenta-se também a forma como foram realizados os testes a nível logístico.

6.1 Testes Laboratoriais

Para testar a robustez do jogo foram feitos inicialmente testes em ambiente laboratorial. Como tal foram escolhidos adultos com características físicas distintas e idades entre os 15 e 55 anos. Estes testes laboratoriais tiveram em média uma duração de 10 minutos cada um sendo alguns deles realizados num laboratório da Universidade do Minho. Cada adulto esteve sentado estrategicamente a fim de permanecer com a face ao nível da câmara ou Webcam. Na Figura 53 pode ver-se a disposição do adulto para a realização dos testes.

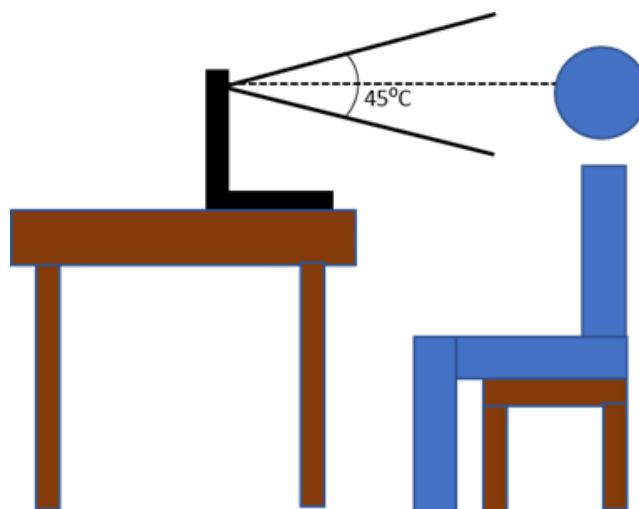


Figura 53 - Posicionamento para realização de testes laboratoriais

Cada adulto foi então colocado numa posição favorável para a deteção da sua face garantindo que o ângulo de ação da camara não fosse superior a 45 graus. A distância entre o adulto e a camara não excedeu os 50 centímetros. Como tal no início de cada teste a altura da camara foi ajustada consoante a fisionomia do adulto em teste. Esse ajuste foi feito com a ajuda do posicionamento do avatar. Como o

HiZeca possui um jogo de imitação, então para o ajustar a altura da câmara foram utilizados os valores de inclinação fornecidos pelo *Openface* e espelhados pelo ZECA avatar. Foram então feitos testes com nove indivíduos. Na Tabela 6 apresentam-se algumas características relevantes dos oito adultos.

Tabela 6 - Lista de adultos utilizados nos testes laboratoriais

Individuo	Idade	Sexo
A	22	Masculino
B	23	Masculino
C	23	Masculino
D	24	Masculino
E	55	Feminino
F	52	Masculino
G	23	Masculino
H	15	Masculino
I	47	Feminino

Durante cada um dos testes foram registadas todas as informações relevantes acerca do desempenho da cada adulto, das dificuldades sentidas e de alguns erros de funcionamento do jogo. Todas essas informações foram avaliadas posteriormente tendo sido corrigidos os erros e feitos os ajustes necessários para a melhor performance do jogo. Para além disso os ficheiros de dados gravados para cada adulto também ajudaram a detetar e perceber quais as dificuldades mais sentidas e qual o desempenho de cada um em cada momento do jogo sendo que o objetivo destes testes laboratoriais não era avaliar a performance de cada jogador, mas sim a robustez e funcionalidade ou jogabilidade do jogo e deteção de possíveis falhas. Por fim, no final de cada teste os valores gravados foram comparados com as informações retiradas e escritas, com o auxílio do jogador, para ser avaliada também a veracidade da informação guardada por parte do HiZeca.

Relativamente aos resultados obtidos para o nível 1, foram identificadas algumas dificuldades na deteção de alguns movimentos bem como o elevado valor limite imposto para a validação de cada movimento. Na *Tabela 7* apresentam-se os resultados dos testes laboratoriais.

Tabela 7 - Tabela resumo com resultados dos testes laboratoriais nível 1

Movimento	Deteção	Timeout	Tempo médio(s)
<i>Abrir a boca</i>	77,78%	2	14
<i>Baixar cabeça</i>	100%	0	2,22
<i>Levantar cabeça</i>	100%	0	4,22
<i>Inclinar cabeça esquerda</i>	100%	0	3,78
<i>Inclinar cabeça direita</i>	100%	0	6,11
<i>Olhar para a esquerda</i>	100%	0	9
<i>Olhar para a direita</i>	100%	0	4
<i>Levantar sobrancelhas</i>	100%	0	3
<i>Sorrir</i>	88,89%	1	4,50
<i>Piscar olhos</i>	88,89%	1	13,88
<i>Franzir sobrolho</i>	100%	0	2,67
<i>Inclinar cabeça lado</i>	100%	0	6
<i>Olhar para o lado</i>	100%	0	3,68

Através dos dados da *Tabela 7* pode observar-se que ocorreram algumas falhas na deteção de alguns movimentos. O movimento de abrir a boca foi o que apresentou uma taxa de deteção mais baixa ficando nos 77,78%. Este valor deve-se ao facto de o limite imposto para a validação do movimento estar um pouco acima do valor que cada adulto conseguia atingir. Por esse motivo apesar a deteção do movimento ser feita corretamente, em alguns casos foi necessário um esforço adicional por parte do adulto para que o movimento fosse validado. Por esse motivo o tempo médio de deteção também acaba por ser atingido tendo sido registado no valor de 14 segundos. Para este movimento o tempo máximo registado foi de 37 segundos e o mais baixo de 2 segundos. Em contrapartida, a deteção e validação deste movimento não foi apenas afetado pelo argumento anterior. Uma outra dificuldade sentida relaciona-se com adultos que possuem barbas. Nestes casos, a deteção dos limites da boca por parte do *OpenFace* foi seriamente afetada. A deteção do lábio superior e inferior da boca do adulto não era feita corretamente o que por sua vez afetava a replicação do movimento por parte do avatar. Como a validação do movimento é feita através dos movimentos do avatar então nestes casos particulares o tempo máximo de resposta foi ultrapassado tendo sido registados dois *timeout*. Um outro movimento afetado pelo motivo atrás referido foi o sorrir. Neste caso, apesar de ter sido registado apenas um *timeout*, o que levou a

uma taxa de detecção de 88,89%, a presença de barba dificultou a ação do *OpenFace* que por sua vez afeta os movimentos replicados.

A detecção do movimento de piscar os olhos é também um caso onde a taxa de detecção não foi máxima. Neste caso foi registado apenas um caso onde a validação deste movimento não foi possível tendo sido registado um *timeout*. Apesar de alguns adultos utilizarem óculos, esta falha não ocorreu nesses casos. Para esta dificuldade não foi encontrada nenhuma explicação óbvia. No entanto, visto ser um movimento que o ser humano executa inconscientemente e com uma elevada velocidade, então esses fatores podem estar na causa da dificuldade de detecção. Como tal este movimento apresenta o tempo médio de detecção também elevado atingindo os 13,88 segundos. Sendo assim, para este primeiro nível os movimentos onde ocorram mais dificuldades foi o de abrir a boca e piscar olhos. Em terceiro lugar aparece o movimento de olhar para a esquerda apresentando um tempo médio de nove segundos. Os restantes movimentos apresentaram uma taxa de detecção que atingiu o valor máximo possível (100%) tendo sido verificadas algumas dificuldades na detecção de movimentos aleatórios de adulto para adulto.

Relativamente aos resultados obtidos para o nível 2, foram detetadas algumas dificuldades na validação de algumas emoções. Neste segundo nível foi possível testar mais uma vez o projeto realizado anteriormente, projeto este responsável pela detecção de emoções a partir do *OpenFace*. A seguir apresenta-se uma tabela resumo (*Tabela 8*) relativamente aos testes laboratoriais.

Tabela 8 - Tabela resumo com resultados dos testes laboratoriais nível 2

Emoção	Validação	Tempo médio(s)	Tempo máximo(s)
<i>Contente</i>	100%	3,22	14
<i>Triste</i>	97,78%	3,55	12
<i>Surpresa</i>	95,56%	4,77	29
<i>Zangado</i>	100%	4,76	22
<i>Medo</i>	100%	4,22	24
<i>Nojo</i>	97,78%	6,34	30

Observando a *Tabela 8* demonstra-se que a validação de cada emoção obteve taxas de validação elevadas. Os dados disponíveis na tabela acima foram conseguidos a partir do conjunto de todos os testes realizados. Cada emoção foi testada cinco vezes por cada um dos nove adultos dando no final um total de 45 validações.

Para a emoção contente foi registado o valor de 100% de validação. No entanto apesar do tempo médio não ser muito alto (3,22 segundos) registou-se um tempo máximo de 14 segundos. Este tempo foi registado por um adulto que possui barba. À imagem do nível 1, a presença de barba torna o processamento de imagem mais complicado por parte do *Openface*. Já para a emoção triste, a taxa de validação foi de 97,78%. Este valor deve-se a um caso isolado e pontual tendo sido descartadas todas as possíveis razões. A emoção que apresentou o valor de validação mais baixo foi a emoção surpresa com 95,56%. Neste caso foram registados dois momentos de falha de validação. Mais uma vez esta falha ocorreu devido a fisionomia do adulto tendo sido registadas as duas falhas durante o teste do mesmo adulto que apresentava barba. Referir também que a emoção onde foi detetada mais dificuldade de expressão foi a emoção de nojo. Para além do tempo médio ser o mais alto (6,34 segundos) foi nesta emoção que se registou o tempo máximo mais alto (30 segundos). Apesar de a diferença em relação a outras emoções não ser muita (superior apenas em 1 segundo em relação a emoção de surpresa), esta emoção apresenta um valor elevado devido a dificuldade de expressão e não por problemas de deteção correta da face por exemplo. Por último foi também verificado que a emoção onde foi denotada mais facilidade de expressão foi a de contente e por esse motivo esta apresenta o valor médio mais baixo de todos registando 3,22 segundos.

Os testes realizados para o nível 3 foram apenas para validação dos dados gravados. Neste nível não se justificou contabilizar tempos de respostas nem respostas certas devido ao baixo nível de dificuldade que o cenário apresenta para um adulto. Como tal foi apenas pedido ao adulto para escolher respostas erradas em alguns casos ou então para voltar a ouvir a história. No fim toda essa informação foi comparada com a informação registada por escrito durante o teste a fim de validar a gravação de dados. Foi também verificado que em alguns casos onde os adultos apresentavam dificuldade na validação da expressão rapidamente percebiam com expressar a emoção diminuído os tempos de resposta a cada execução da mesma emoção.

Todos os dados registados, tempos médios calculados e taxas de validação foram calculados através dos registos dos nove adultos. Contudo existe ainda um caso de um adulto onde os dados registados não foram tidos em conta. Esse adulto apresentava uma característica distintiva, a cor de pele negra. Os dados registados não foram contabilizados nos cálculos da *Tabela 8* devido á impossibilidade de este adulto não finalizar o nível 1 e 2. Este facto deve-se a grandes dificuldades no processamento de imagem por parte do *Openface*. Apesar de várias repetições não foi possível terminar nenhum dos níveis tendo estes dados sido descartados. No entanto, servem para concluir que no caso de o jogador possuir cor

de pele negra pode sentir alguma dificuldade relacionada com a deteção da sua face por parte do *Openface*.

Assim, no geral, os testes laboratoriais realizados obtiveram um nível de sucesso aceitável tendo sido importantes para serem efetuados ajustes e correções a fim de o jogo se tornar mais estável para a realização dos testes em ambiente escolar.

6.2 Testes em Ambiente Escolar

Depois de testado o jogo com adultos e corrigidas todas as falhas encontradas o HiZeca foi então testado em ambiente escolar. Estes testes foram realizados na escola EB1 de Gualtar em Braga. Os testes foram realizados com crianças tipicamente desenvolvidas e com crianças com PEA.

Crianças tipicamente desenvolvidas

O HiZeca foi testado inicialmente com crianças tipicamente desenvolvidas e em seguida com crianças com PEA. Estas crianças estavam numa faixa etária entre os 6 e 8 anos de idade. Nestes testes em ambiente escolar, os objetivos foram:

- Verificar a aceitação da ferramenta por parte de crianças entre os 6 e 10 anos
- Teste da facilidade de utilização, robustez e desempenho do jogo
- Verificação de condições favoráveis para o processamento de imagem
- Testar fiabilidade dos dados adquiridos com auxílio da webcam
- Testes com crianças tipicamente desenvolvidas e depois com criança com PEA

Para isso estes testes foram feitos segundo as mesmas condições de posicionamento da criança face ao computador (à semelhança dos testes laboratoriais Figura 53 - Posicionamento para realização de testes laboratoriais). Estas sessões de testes tiveram uma duração média de 15 minutos por criança. Durante a fase de testes as crianças foram incentivadas a jogar os três níveis diferentes. Na Tabela 9 apresentam-se o conjunto das cinco crianças utilizadas para os testes.

Tabela 9 - Lista de crianças utilizadas nos testes em ambiente escolar

Criança	Idade	Sexo	PEA
<i>A</i>	<i>6</i>	<i>Feminino</i>	<i>Não</i>
<i>B</i>	<i>6</i>	<i>Masculino</i>	<i>Não</i>
<i>B</i>	<i>6</i>	<i>Masculino</i>	<i>Não</i>
<i>D</i>	<i>8</i>	<i>Feminino</i>	<i>Sim</i>
<i>E</i>	<i>7</i>	<i>Masculino</i>	<i>Sim</i>

Todas as informações relevantes foram registadas teste a teste. Para além disso não foram gravados nenhuns vídeos ou fotografias a fim de manter a anonimidade das crianças em teste. Estes testes foram feitos com o consentimento dos pais e professores. Para além das informações registadas por escrito foram também gravados os dados do ficheiro de texto sobre a performance de cada criança. A presença das professoras da referida escola também permitiu obter um *feedback* destas profissionais que trabalham diariamente com estas crianças.

Em relação aos testes feitos para o nível 1 com crianças tipicamente desenvolvidas (três crianças apenas) foram registados os valores presentes na Tabela 10.

Tabela 10 - Tabela resumo com valores médios dos resultados dos testes em ambiente escolar nível 1 com crianças tipicamente desenvolvidas

Movimento	Deteção	Timeout	Tempo médio(s)	Tempo máximo(s)
<i>Abrir a boca</i>	<i>66,67%</i>	<i>1</i>	<i>5,5</i>	<i>7</i>
<i>Baixar cabeça</i>	<i>100%</i>	<i>0</i>	<i>5</i>	<i>10</i>
<i>Levantar cabeça</i>	<i>100%</i>	<i>0</i>	<i>9,67</i>	<i>13</i>
<i>Inclinar cabeça esquerda</i>	<i>100%</i>	<i>0</i>	<i>15</i>	<i>21</i>
<i>Inclinar cabeça direita</i>	<i>100%</i>	<i>0</i>	<i>11,33</i>	<i>23</i>
<i>Olhar para a esquerda</i>	<i>100%</i>	<i>0</i>	<i>4,67</i>	<i>6</i>
<i>Olhar para a direita</i>	<i>100%</i>	<i>0</i>	<i>9,67</i>	<i>16</i>
<i>Levantar sobrancelhas</i>	<i>66,67%</i>	<i>1</i>	<i>8,5</i>	<i>15</i>
<i>Sorrir</i>	<i>33,33%</i>	<i>2</i>	<i>5</i>	<i>5</i>
<i>Piscar olhos</i>	<i>100%</i>	<i>0</i>	<i>10,33</i>	<i>20</i>
<i>Franzir sobrolho</i>	<i>100%</i>	<i>0</i>	<i>16</i>	<i>39</i>
<i>Inclinar cabeça lado</i>	<i>100%</i>	<i>0</i>	<i>5,33</i>	<i>8</i>
<i>Olhar para o lado</i>	<i>100%</i>	<i>0</i>	<i>7,33</i>	<i>11</i>

Foram registados alguns valores que mereceram atenção especial. Para o movimento de abrir a boca voltou a detetar-se uma pequena falha na validação deste movimento. Foi registado apenas um *timeout* o que levou uma taxa de deteção de 66,67%. Apesar de os limites de validação deste movimento terem sido ajustados previamente não foi o suficiente pelo que o seu valor terá de ser ajustado futuramente. No entanto, não foi este o movimento que apresentou uma taxa de deteção mais baixa. Para o movimento de sorrir foram registados dois *timeout* tendo por esse motivo uma taxa de validação mais baixa no valor

de 33,33%. Para o movimento de levantar as sobancelhas verificou-se também que os limites de validação estariam altos levando a que ocorresse um *timeout* para este movimento. Em relação aos tempos médios de resposta o movimento que apresentou um maior valor foi o movimento de franzir o sobrolho. Este movimento apresenta uma taxa de validação de 100%, no entanto possui um tempo médio alto devido á dificuldade executar o movimento por parte das crianças. Como seria de esperar foi neste movimento também que se registou um valor máximo de resposta de 39 segundos. De referir que os valores dos tempos médios calculados da tabela anterior não incluem os *timeout* que ocorreram. É importante perceber também que apesar dos valores registados serem valores reais o reduzido número de crianças em teste leva a que as possíveis oscilações de valores de respostas de cada criança afete muito os valores médios ou taxas de validação.

Em relação aos testes feitos para o nível 2 com crianças tipicamente desenvolvidas (três crianças apenas) foram registados os valores presentes na Tabela 11.

Tabela 11 - Tabela resumo com valores médios dos resultados dos testes ambiente escolar nível 2 com crianças tipicamente desenvolvidas

Emoção	Validação	Tempo médio (s)	Tempo máximo (s)
<i>Contente</i>	<i>100%</i>	<i>6,80</i>	<i>19</i>
<i>Surpresa</i>	<i>93,33%</i>	<i>6,56</i>	<i>22</i>
<i>Medo</i>	<i>100%</i>	<i>6,80</i>	<i>28</i>
<i>Triste</i>	<i>80%</i>	<i>6,70</i>	<i>18</i>
<i>Zangado</i>	<i>86,87%</i>	<i>6,85</i>	<i>20</i>
<i>Nojo</i>	<i>80%</i>	<i>8,08</i>	<i>27</i>

Os dados disponíveis na tabela acima foram conseguidos a partir do conjunto de todos os testes realizados. Cada emoção foi testada cinco vezes por cada uma das três crianças dando no final um total de 15 validações por emoção.

Neste segundo nível as emoções de contente e medo foram as que registaram uma taxa de validação máxima de 100%. Para estas mesmas emoções os tempos médios de resposta foram exatamente iguais fixando-se nos 6,8 segundos. Já para a emoção de nojo o valor registado foi o mais baixo de todos sendo de 80%. Mesmo antes da realização dos testes laboratoriais e em ambiente escolar que era sabido que esta emoção causava mais dificuldades na expressão. Isto deve-se ao facto de ser uma emoção que mesmo sendo usual não é uma das que está mais presente no dia a dia, daí a dificuldade em expressá-la. A emoção que apresentou um tempo máximo de resposta foi a emoção do medo. Este valor foi

explicado por ter sido sentido por parte da criança alguns nervosismos neste segundo nível acabando por demonstrar algum desconforto no que diz respeito a expressar as diferentes emoções. Em relação aos tempos médios para a emoção de surpresa foi registado o valor mais baixo de 6,56 segundos. Em contrapartida para a emoção de nojo registou-se o valor mais alto de 8,08 segundos. Neste segundo nível foi notado que algumas crianças apresentavam alguma dificuldade em expressar emoções. Esta dificuldade não se prendeu com a falta de conhecimento sobre as próprias emoções, mas sim pelo facto de se sentirem observadas o que acabava por a deixar desconfortáveis. Foi detetado que quando expressavam as emoções inconscientemente a sua validação era mais fácil.

Em seguida apresenta-se a Tabela 12 com os dados relativos ao desempenho das crianças tipicamente desenvolvidas no nível 3 (modo história).

Tabela 12 - Tabela resumo com resultados dos testes ambiente escolar nível 3 com crianças tipicamente desenvolvidas

História	Respostas erradas	Repetições	Tempo médio(s)	Tempo máximo(s)
<i>História 1</i>	<i>0</i>	<i>0</i>	<i>11</i>	<i>22</i>
<i>História 2</i>	<i>1</i>	<i>1</i>	<i>7</i>	<i>12</i>
<i>História 3</i>	<i>0</i>	<i>0</i>	<i>13,67</i>	<i>17</i>
<i>História 4</i>	<i>2</i>	<i>1</i>	<i>11,5</i>	<i>14</i>
<i>História 5</i>	<i>0</i>	<i>0</i>	<i>6,33</i>	<i>11</i>
<i>História 6</i>	<i>0</i>	<i>1</i>	<i>7,33</i>	<i>16</i>
<i>História 7</i>	<i>0</i>	<i>0</i>	<i>7,33</i>	<i>15</i>
<i>História 8</i>	<i>0</i>	<i>0</i>	<i>2,33</i>	<i>4</i>
<i>História 9</i>	<i>0</i>	<i>1</i>	<i>2,67</i>	<i>6</i>
<i>História 10</i>	<i>0</i>	<i>0</i>	<i>8,33</i>	<i>18</i>
<i>História 11</i>	<i>2</i>	<i>0</i>	<i>7</i>	<i>9</i>
<i>História 12</i>	<i>0</i>	<i>0</i>	<i>3</i>	<i>5</i>
<i>História 13</i>	<i>1</i>	<i>0</i>	<i>9,67</i>	<i>13</i>
<i>História 14</i>	<i>0</i>	<i>1</i>	<i>5,67</i>	<i>8</i>
<i>História 15</i>	<i>1</i>	<i>0</i>	<i>4,67</i>	<i>7</i>

Na Tabela 12 foram registadas algumas informações pertinentes para a avaliação do desempenho das crianças ao longo do nível. Neste terceiro nível foram registadas algumas respostas erradas em algumas histórias bem como algumas repetições que aconteceram quando as crianças sentiam a necessidade de

voltar a ouvir e ver a história para dar a sua resposta final. Para as histórias quatro e onze foram registadas duas respostas erradas sendo este o maior número de respostas erradas em todos os testes. A história que apresentou um maior valor médio de resposta foi a história 3 relacionada com a emoção de zangado com o valor de 13,67 segundos. O valor máximo de resposta foi registado para a história 10. Neste nível notou-se a capacidade de raciocínio das crianças para associar rapidamente as emoções a história em causa. O facto de existir uma imagem onde o avatar aparece expressando a emoção pretendida foi também uma grande ajuda pois algumas crianças não sabiam ainda ler e as suas respostas eram dadas através da associação da face do avatar com a imagem presente nos diferentes botões do menu de seleção de resposta deste nível. A funcionalidade de repetição de histórias também foi uma ajuda em casos onde a criança ficava confusa. Nestas situações depois da repetição da história a resposta dada pela criança era espontânea e acertada. Dos três níveis testados, o terceiro e último nível foi aquele que apresentou melhores resultados de *performance* e de aceitação por parte destas crianças tipicamente desenvolvidas. A oportunidade de utilização do rato para jogar acaba por ser um elemento cativante para as crianças. Isso conclui-se por ter sido notado que estas estavam mais interessadas e empenhadas neste último nível.

Crianças com PEA

Foi apenas possível a realização destes testes com duas. No entanto estas crianças apresentavam características bem diferentes. A criança D foi a primeira a ser subtida aos testes. Esta criança apresenta um grau de autonomia elevado. Consegue falar fluidamente e já consegue ler o que ajuda bastante pois em todos os níveis para além do som existem elementos escritos. Na Figura 54 pode ver-se a prestação da criança D no nível 1.

NIVEL 1		
Movimento	Resposta Correta	Tempo de Resposta
Baixar Cabeça	Sim	00:00:10
Levantar a Cabeça	Sim	00:00:11
Piscar olhos	Sim	00:00:03
Inclinar a Cabeça Esq	Sim	00:00:28
Inclinar a Cabeça Lad	Não	TimeOut
Sorrir	Sim	00:00:06
Inclinar a Cabeça Dir	Sim	00:00:28
Abrir a boca	Sim	00:00:08
Lev/ Sobrancelha	Não	TimeOut
Olhar para a Esq	Sim	00:00:20
Olhar para a Lad	Sim	00:00:04
Franzir sobrolho	Sim	00:00:02
Olhar para a Dir	Sim	00:00:06

Figura 54 - Resultados nível 1 criança D

Como se pode observar, neste nível 1 esta criança apenas não conseguiu executar corretamente dois movimentos (inclinar a cabeça para o lado e levantar as sobrancelhas). Para o movimento de levantar as sobrancelhas esta criança não conseguiu executar o movimento devido a dificuldades de controlo de cada movimento isolado da face. Por esse motivo o tempo máximo de execução do movimento foi ultrapassado. Uma outra dificuldade sentida foi a distinção entre a direita e esquerda. Daí, a partir da observação dos dados da Figura 54 consegue perceber-se que os movimentos que envolvem a distinção entre direita e esquerda apresentam tempos de resposta elevados tal como acontece para o movimento de inclinar a cabeça para a direita registando um valor de 28 segundos de resposta ou execução do movimento. Apesar dessas dificuldades o desempenho desta criança teve um *feedback* positivo tendo sido o objetivo deste nível cumprido neste caso.

Para o nível 2, pode ver-se os resultados obtidos acerca do desempenho da criança na Figura 55.

NIVEL 2		
Emoção	Resposta Correta	Tempo de Resposta
CONTENTE 1	Sim	00:00:05
CONTENTE 2	Sim	00:00:10
CONTENTE 3	Sim	00:00:06
CONTENTE 4	Sim	00:00:03
CONTENTE 5	Sim	00:00:03
TRISTE 1	Sim	00:00:05
TRISTE 2	Sim	00:00:02
TRISTE 3	Sim	00:00:05
TRISTE 4	Sim	00:00:02
TRISTE 5	Sim	00:00:02
NOJO 1	Sim	00:00:28
NOJO 2	Sim	00:00:02
NOJO 3	Sim	00:00:03
NOJO 4	Sim	00:00:03
NOJO 5	Sim	00:00:03
MEDO 1	Sim	00:00:04
MEDO 2	Sim	00:00:04
MEDO 3	Sim	00:00:06
MEDO 4	Sim	00:00:11
MEDO 5	Sim	00:00:08
SURPRESO 1	Sim	00:00:04
SURPRESO 2	Sim	00:00:14
SURPRESO 3	Sim	00:00:03
SURPRESO 4	Sim	00:00:02
SURPRESO 5	Sim	00:00:03
ZANGADO 1	Sim	00:00:04
ZANGADO 2	Sim	00:00:05
ZANGADO 3	Sim	00:00:02
ZANGADO 4	Sim	00:00:02
ZANGADO 5	Sim	00:00:02

Figura 55 - Resultados nível 2 criança D

No segundo nível, onde são testadas as capacidades de expressar emoções o desempenho da criança foi surpreendente. Para além de conseguir expressar todas as emoções corretamente, efetuou cada expressão em tempos aceitáveis. Para a expressão de zangado por exemplo os tempos de resposta foram baixos (nunca ultrapassou os 5 segundos) o que traduz uma facilidade na expressão da emoção. Á

exceção da emoção de medo, todas as emoções apresentaram valores baixos nos últimos dois momentos de expressão. No penúltimo momento, apesar de existir o apoio do avatar este encontra-se reduzido no seu tamanho sendo complicado detetar a sua expressão. No último momento, apenas prevalece o nome da emoção e a criança em causa conseguiu expressar cada emoção em tempo recorde. Um outro facto registado foi os valores obtidos para a emoção nojo. Se observarmos os tempos para esta emoção, no primeiro momento que é pedida a execução da expressão associada ao nojo a criança teve algumas dificuldades comprovadas pelos 28 segundos de tempo de resposta. Em seguida todos os momentos seguintes desta emoção registaram valores baixos o que significa que num primeiro momento houve um tempo de aprendizagem e em seguida a criança já era capaz de expressar facilmente a emoção. Assim para este segundo nível onde foi já introduzido a distinção de emoções e capacidade de as expressar esta criança teve um bom desempenho e aprendizagem.

Para o terceiro e último nível foi necessário em primeiro lugar explicar a criança como utilizar o rato do computador pois esta nunca tinha utilizado. Depois desse momento de aprendizagem, foi então testado o modo história. Na Figura 56 temos os dados registados sobre o modo história (nível 3) da criança D.

NIVEL 3		
História	Resposta Correta	Tempo de Resposta
Historia 3	Repetição	
Historia 3	Sim	00:00:33
Historia 6	Sim	00:00:17
Historia 9	Repetição	
Historia 9	Sim	00:00:04
Historia 12	Sim	00:00:10
Historia 15	Sim	00:00:11
Historia 14	Sim	00:00:32
Historia 11	Sim	00:00:09
Historia 8	Sim	00:00:07
Historia 5	Errado	
Historia 5	Repetição	
Historia 5	Sim	00:00:13
Historia 2	Sim	00:00:11
Historia 1	Errado	
Historia 1	Sim	00:00:59
Historia 4	Sim	00:00:07
Historia 7	Sim	00:00:15
Historia 10	Sim	00:00:12
Historia 13	Sim	00:00:25

Figura 56 - Resultados nível 3 criança D

Neste terceiro nível não foi registado nenhuma história onde o tempo máximo de resposta fosse atingido ou então o número de vidas esgotado. No entanto a partir dos dados gravados pode ver-se que ocorreram três momentos onde a repetição da história foi necessária para a criança esclarecer qual a emoção certa. Vejamos por exemplo o caso da história cinco onde inicialmente a criança selecionou uma resposta errada e em seguida, depois de repetir a história, foi capaz de selecionar a emoção certa. À imagem do

que aconteceu com as crianças tipicamente desenvolvidas, este terceiro nível foi o que suscitou mais interesse para a criança. Como esta criança dizia tudo o que estava a pensar foi possível perceber que depois de ouvir a história e ter observado a face do avatar, associava a expressão e o contexto da história à emoção certa à exceção das histórias um e cinco. Como referido atrás, o facto desta criança saber ler também ajudou em caso de dúvida na relação à expressão ou imagem presente. Em cada botão esta associava a emoção através do nome do botão.

A segunda criança (criança E) apresentou outras características e dificuldades em relação à criança D. Esta criança tem muitas dificuldades de concentração e era uma criança muito ativa e mexida. Desde logo essas complicações provocaram uma maior dificuldade na realização dos testes. Em relação ao nível 1, a validação dos movimentos foi afetada pelas constantes movimentações da criança o que provocava dificuldades na deteção da face da criança. Por esse motivo o teste do nível um não foi terminado tendo sido testados apenas alguns movimentos tal como podemos ver na seguinte figura (Figura 57).

NIVEL 1		
Movimento	Resposta Correta	Tempo de Resposta
Baixar Cabeça	Sim	00:00:23
Levantar a Cabeça	Sim	00:00:26
Inclinar a Cabeça Lad	Não	TimeOut
Inclinar a Cabeça Dir	Não	TimeOut
Sorrir	Sim	00:00:39
Inclinar a Cabeça Esq	Não	TimeOut
Franzir sobrolho	Sim	00:00:04

Figura 57 - Resultados nível 1 criança E

Os testes do nível 2 obtiveram melhores resultados. Neste nível, inicialmente a criança conseguiu expressar algumas emoções tal como se pode ver na Figura 58. Para a emoção de surpresa, a criança associou o movimento de abrir a boca a esta emoção tendo obtido tempos de resposta aceitáveis.

NIVEL 2

Emoção	Resposta Correta	Tempo de Resposta
NOJO 1	Sim	00:00:08
NOJO 2	Sim	00:00:03
NOJO 3	Sim	00:00:15
NOJO 4	Sim	00:00:12
NOJO 5	Sim	00:00:20
ZANGADO 1	Sim	00:00:28
ZANGADO 2	Não	TimeOut
ZANGADO 3	Sim	00:00:19
ZANGADO 4	Sim	00:00:19
ZANGADO 5	Sim	00:00:30
MEDO 1	Sim	00:00:06
MEDO 2	Sim	00:00:03
MEDO 3	Sim	00:00:05
MEDO 4	Sim	00:00:07
MEDO 5	Sim	00:00:07
CONTENTE 1	Não	TimeOut
CONTENTE 2	Não	TimeOut
CONTENTE 3	Não	TimeOut
CONTENTE 4	Não	TimeOut
CONTENTE 5	Não	TimeOut
SURPRESO 1	Sim	00:00:05
SURPRESO 2	Sim	00:00:02
SURPRESO 3	Sim	00:00:04
SURPRESO 4	Sim	00:00:16
SURPRESO 5	Sim	00:00:03
TRISTE 1	Sim	00:00:15
TRISTE 2	Sim	00:00:16
TRISTE 3	Sim	00:00:02
TRISTE 4	Sim	00:00:05
TRISTE 5	Sim	00:00:05

Figura 58 - Resultados nível 2 criança E

No entanto depois disso a criança começou a associar o facto de dizer a expressão em voz alta ao progresso de carregamento da barra presente neste nível, indicadora da correta execução da expressão pretendida. A partir desse momento, apesar de muito esforço por parte da professora presente no teste, a criança não se focou no essencial permanecendo em movimentos permanentes e dizendo sempre o nome da emoção que era pedida. Uma outra complicação prendia-se com o facto desta criança ter momentos onde fixava a sua atenção e concentração em elementos externos ao jogo o que provocava uma desatenção total sobre o jogo. Para o terceiro e último nível o comportamento desta criança mudou significativamente. No início do nível foi possível verificar que a postura da criança foi mudando (Figura 59).

NIVEL 3

História	Resposta Correta	Tempo de Resposta
Historia 1	Errado	
Historia 1	Errado	
Historia 1	Repetição	
Historia 1	Sim	00:00:17
Historia 4	Errado	
Historia 4	Repetição	
Historia 4	Errado	
Historia 4	Sim	00:00:51
Historia 13	Sim	00:00:09
Historia 2	Sim	00:00:12
Historia 5	Repetição	
Historia 5	Sim	00:01:17
Historia 3	Sim	00:00:12
Historia 6	Sim	00:00:07
Historia 14	Repetição	
Historia 14	Sim	00:00:02
Historia 7	Sim	00:00:05
Historia 10	Sim	00:00:05
Historia 8	Repetição	
Historia 8	Sim	00:00:12
Historia 15	Sim	00:00:03
Historia 11	Sim	00:00:04
Historia 9	Repetição	
Historia 9	Sim	00:00:23
Historia 12	Sim	00:00:06

Figura 59 - Resultados nível 3 criança E

Neste caso, a criança em questão não conseguia selecionar a resposta correta sozinho devido a utilização do rato. Como tal, esta indicava qual a resposta que achava correta sendo esta selecionada por o autor desta dissertação.

Para a primeira história apresentada foram registadas duas respostas erradas. Em seguida a criança foi incentivada a voltar a ouvir e ver a história e em seguida conseguiu responder corretamente. Nas seguintes histórias a atenção e concentração da criança aumentou significativamente. Para algumas, foi necessário a repetição da história. Nestes casos a criança conseguia entender qual a emoção. Depois disso e a partir do momento que sabia qual a emoção certa facilmente conseguia associar essa emoção aos botões disponíveis no menu de seleção de resposta. Em relação aos tempos de resposta, estes foram melhores do que o esperado. A aceitação deste terceiro nível foi também notada e o desempenho da criança surpreendente.

Assim depois de concluído todos os testes com crianças tipicamente desenvolvidas e com crianças com PEA pode afirmar-se que o *feedback* é positivo. A aceitação do HiZeca foi boa em todos os níveis. No entanto referir que o nível 3 foi o nível que suscitou mais interesse por parte de todas as crianças.

7. CONCLUSÕES E TRABALHOS FUTUROS

O trabalho desenvolvido nesta dissertação está inserido no projeto Robótica Autismo (<http://robotica-autismo.dei.uminho.pt/>) onde se desenvolvem ferramentas capazes de interagir com crianças com PEA ajudando no desenvolvimento das suas capacidades cognitivas e sociais. Estas crianças possuem dificuldades em relacionar-se com os outros e em compreender estados emocionais. No projeto Robótica Autismo foi utilizado e testado um robô humanoide como mediador nas atividades de promoção da comunicação social e na identificação e reconhecimento de emoções.

Nos últimos anos o número de jogos capazes de auxiliar na melhoria de dificuldades de crianças com PEA tem aumentado significativamente, fruto do desenvolvimento tecnológico e social. Torna-se relevante a criação de novas formas de abordagem aos problemas da sociedade e a criação de soluções que se enquadrem no contexto social onde serão inseridas. A aceitação de jogos didáticos e sérios torna-se mais fácil quando é desenvolvido um jogo apelativo e cativante onde o conteúdo do jogo e seus objetivos estão camuflados pela parte gráfica conseguindo assim criar entretenimento e aprendizagem construtivos. Este tipo de jogos são uma alternativa a utilização de outras ferramentas como é o caso dos robôs humanoides usados em vários projetos relacionados com o autismo.

Existem inúmeros jogos de espelhamento de movimentos do corpo humano, que podem ser usados na área da saúde para treino e fisioterapia por exemplo, mas existem poucos onde esse espelhamento engloba a face humana. No caso das crianças com PEA torna-se mais importante o treino de expressões faciais e a associação dessas expressões a emoções. Em alguns jogos esse treino é feito a partir de associação de imagens, sons e até mesmo vídeos. A imitação de expressões ou replicação delas em muitos casos é posta de lado o que leva a uma aprendizagem apenas visual e não física. O reconhecimento de emoções é importante, mas o treino de expressões faciais não pode ser esquecido. Neste trabalho foi desenvolvido um jogo associado á distinção e compreensão de emoções á imagem do robô referido anteriormente. Para isso implementou-se três níveis com objetivos distintos. Pretende-se com este jogo que a criança treine alguns movimentos faciais presentes em algumas expressões e execute ou expresse emoções. Com o treino dessas expressões e movimentos a tarefa de distinção e compreensão de emoções torna-se mais intuitiva e fácil.

O HiZeca reúne um conjunto de três níveis onde o jogador é posto a prova, tendo a oportunidade de jogar três modos de jogo distintos e igualmente importantes. Foi implementada uma ordem lógica de níveis com o aumento gradual de dificuldade tendo como objetivo principal a compreensão de emoções. No nível um são treinados movimentos faciais distintos. O ZECA avatar replica os movimentos do jogador em tempo real imitando-o em todos os seus movimentos. É pedido ao jogador para efetuar variados

movimentos que são validados pelo jogo. No nível dois o jogador expressa emoções a partir de expressões faciais que são validadas pelo software. Neste caso o avatar expressa também emoções para ajudar o jogador a entender e associar as emoções às expressões. No terceiro nível foi implementado o modo história que permite a associação de histórias a emoções. Neste nível para além de ser treinada a atenção da criança é testada também a capacidade de raciocínio e compreensão de emoções em diferentes histórias.

O sistema *FACS* desenvolvido por *Ekman e Friesen* permite uma padronização de movimentos faciais através das AU relacionando-as com expressões faciais. Com este sistema a determinação de estados emocionais torna-se mais fácil. Para além disso permite o processamento e tratamento de dados obtidos com *software* otimizado para a aquisição de informação facial de uma forma mais objetiva e robusta.

O *Openface*, utilizado neste trabalho para aquisição de dados da face do utilizador, é uma ferramenta que utiliza redes neuronais para detetar movimentos faciais e os classificar nas várias AU existentes. Com esses valores foi possível criar um nível onde o avatar imita em tempo real a face do utilizador. Permitiu também a determinação de estados emocionais que serviram para validar as expressões executadas pelo jogador durante o segundo nível do presente jogo.

Para a criação de toda a parte gráfica do jogo foi utilizado o motor de jogo Unity 3D. Este software para além de ser de acesso gratuito oferece uma interface otimizada para o desenvolvimento de jogos 2D e 3D e disponibiliza documentação muito bem estruturada tornando a sua aprendizagem mais compreensível. É possível controlar os movimentos do avatar através de esquemas ou fluxogramas sem ser necessário desenvolver muita quantidade de código. Permite também a criação de animações dos avatares. No entanto, se a animação a ser criada envolver muitos movimentos de muitos músculos ou membros de um avatar humanoide por exemplo, aconselha-se a utilização de software otimizado para isso.

O avatar foi desenvolvido no *Adobe FUSE*. Este *software* permite uma rápida assemblagem de membros do corpo humano criando rapidamente avatares realistas e personalizáveis. Permite também submeter o modelo 3D para o *software online Mixamo* onde pode associar-se movimentos predefinidos ao avatar desenvolvido. Esses movimentos associados podem depois ser carregados para o Unity 3D e reproduzidos sempre que for necessário.

Foram realizados testes em ambiente laboratorial e em ambiente escolar com crianças tipicamente desenvolvidas e crianças com PEA. Todos os testes foram feitos cumprindo as questões éticas.

Os testes em ambiente laboratorial permitiram otimizar o sistema. Os jogadores utilizados em teste permaneceram a uma distância da camara nunca superior a 50 centímetros. Foram, também garantidas

limitações em relação ao ângulo de ação da câmara. Como tal foi tido sempre em conta o alinhamento da face do jogador com a câmara a fim de garantir uma melhor leitura de todos os valores retornados pelo OpenFace. Durante os testes realizados foram detetadas algumas falhas do sistema. Para a validação do movimento de abrir a boca nos testes laboratoriais verificou-se que os limites impostos para essa validação estariam altos. A presença de barba foi também um obstáculo no que diz respeito a deteção dos diferentes movimentos da cara. Foram registados casos onde as deteções de movimentos da boca se tornavam mais complicadas. No entanto a essa dificuldade não foi dada muita atenção pelo facto do jogo desenvolvido ter como público-alvo crianças que não apresentarão essa característica física.

Os testes em ambiente escolar foram realizados com crianças entre os 6 e 8 anos de idade. Os testes em ambiente escolar permitiram concluir que avatar humanoide criado, o ZECA avatar, aumentou a aceitabilidade das crianças relativamente ao HiZeca. Durante os testes realizados essa aceitação foi visível tendo sido possível comprovar que ocorre a criação de empatia e ligação entre as crianças e o avatar. Esta empatia é talvez um dos fatores mais importantes para o sucesso deste jogo pois se a criança olhar para o avatar como um amigo isso aumentará o seu interesse pelo jogo o que por sua vez aumenta o tempo de contacto com uma ferramenta de desenvolvimento e melhorias das suas dificuldades e capacidades.

Nos testes realizados com crianças com PEA foi possível perceber que o HiZeca cativou a atenção destas crianças e o avatar desenvolvido permitiu uma melhor aceitação do jogo. A dança introduzida em cada resposta, movimento ou emoção certas é também um elemento que apesar de simples, incentivava e recompensava todo o esforço da criança. Os menus simples e intuitivos permitiram que a criança pudesse jogar autonomamente não dependendo de nenhum adulto no caso da criança D. Nestes testes também foi verificado que o nível 3 foi o nível que obteve maior grau de concentração e empenho. No nível dois a repetibilidade de execução de expressões leva as crianças a associar as emoções o que se reflete por um bom desempenho no nível seguinte (nível 3).

O nível 1 do jogo foi o nível que apresentou mais dificuldades no que diz respeito a aceitação por parte das crianças. Este jogo apresenta, não só no nível 1 como também no nível 2, uma abordagem que as crianças não estão habituadas e não se sentem tão descontraídas como nos típicos jogos de computador. O reduzido número de crianças utilizadas nos testes realizados não permite garantir a criação de conclusões representativas de todo o público-alvo, mas permitiu entender quais os aspetos a melhorar e quais as situações onde a atenção das crianças aumentou. O nível 3 foi aquele que teve uma taxa de aceitação mais elevada. Tanto nos testes realizados com crianças tipicamente desenvolvidas como com crianças com PEA este nível foi o que despertou mais a atenção dos jogadores.

Com os testes realizados com crianças tipicamente desenvolvidas e com crianças com PEA foi possível concluir este tipo de jogos tem mais aceitação se existir mais interação entre a criança e o computador (utilização do rato ou do teclado por exemplo) tornando a criança mais recetiva e focada no próprio jogo. Como trabalho futuro, deverão ser feitos mais testes em ambiente escolar a fim de se obter um maior número de dados para permitir fundamentar melhor algumas conclusões. A replicação de movimentos do avatar no nível um deverá ser melhorada e a criação de mecanismos capazes de corrigir movimentos, suavizando-os permitirá o melhor desempenho do nível. Poderão ser também impostas algumas restrições no nível um a fim de se obter uma imitação mais uniformizada. Estudar uma solução para a integração do software Openface no Unity 3D evitando a troca de informação e dependência mútua destas ferramentas. Poderá ser criado também um sistema de análise dos dados retirados ao longo do jogo para obter gráficos e tabelas acerca do desempenho de cada jogador. Para além disso implementar um sistema de gravação de dados em base de dados online pode permitir a concentração de todos os dados para possíveis estudos estatísticos.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] P. Miguel Lopes de Sousa, I. Margarida Silva Costa dos Santos, and P. Miguel Lopes de Sousa Isabel Margarida Silva Costa dos Santos, “CARACTERIZAÇÃO DA SÍNDROME AUTISTA,” 2005.
- [2] Cláudia, “Face a face com o Autismo-será a Inclusão um mito ou uma realidade?,” 2012.
- [3] H. Freitas *et al.*, “Using a Humanoid Robot as the Promoter of the Interaction with Children in the Context of Educational Games,” 2017.
- [4] Vera Mónica Loureiro Ferreira Orientadora:, “UNIVERSIDADE CATÓLICA PORTUGUESA CENTRO REGIONAL DE BRAGA Utilização das tecnologias em crianças com perturbações do espectro autista em contexto da Prática de Ensino Supervisionada,” 2013.
- [5] P. Ekman and E. L. Rosenberg, *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS), Second Edition.* 2005.
- [6] B. Charles Darwin, R. Society Hardwicke, and B. Fritz Muller, *With illustrations. 2 vols. 8vo. 24s. Murray, 1871. THE VARIATION OF ANIMALS AND PLANTS UNDER DOMESTICATION. Third Thousand. With Illustrations. 2 vols. 8vo. 28s. Murray, 1868. Effects of Crossing. With Woodcuts. 1872.*
- [7] V. Silva, F. Soares, and J. S. Esteves, “Mirroring emotion system-on-line synthesizing facial expressions on a robot face,” in *International Congress on Ultra Modern Telecommunications and Control Systems and Workshops*, 2016, vol. 2016-December, pp. 213–218.
- [8] V. Silva *et al.*, “Real-time emotions recognition system,” in *International Congress on Ultra Modern Telecommunications and Control Systems and Workshops*, 2016, vol. 2016-December, pp. 201–206.
- [9] A. S. Disorder, Q. Autism, and D. More, “Using QTrobot as a therapy assistant for children with autism Using QTrobot as a therapy assistant for children with autism,” 2018.
- [10] “Conheça o Milo! | Robots4Autism.” [Online]. Available: <https://robots4autism.com/milo/>. [Accessed: 09-Aug-2019].
- [11] D. Mazzei, N. Lazzeri, D. Hanson, and D. De Rossi, “HEFES: An Hybrid Engine for Facial Expressions Synthesis to control human-like androids and avatars,” in *Proceedings of the IEEE RAS and EMBS International Conference on Biomedical Robotics and Biomechatronics*, 2012, pp. 195–200.
- [12] *Wii Documentation.* 2009.
- [13] N. Felicidade Tomaz Braz, L. Rocha Dutra, P. E. S. Medeiros, and A. A. Scianni, “Eficácia do

- Nintendo Wii em desfechos funcionais e de saúde de indivíduos com doença de Parkinson: uma revisão sistemática," *Fisioter. Pesqui. [online]*. 2018, vol.25, n.1, pp.100-106. ISSN 2316-9117, 2018.
- [14] "PlayStation® Move motion controller Instruction Manual CECH-ZCM1E," 2010.
- [15] S. K. and C. O. G. Yolcu, "Real Time Virtual Mirror Using Kinect," *Balk. J. Electr. Comput. Eng.*, vol. Vol.2, 2014.
- [16] J. U. Md. Moniruzzaman Monir, Md. Nafis Hasan Siddique, Nahyan Ebn Hashem, Afsana Pervin Tanni, *An Efficient Magic Mirror Using Kinect*, vol. 2018, no. May. 2016.
- [17] M. E. Latoschik, J.-L. Lugin, and D. Roth, "FakeMi," in *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology - VRST '16*, 2016, pp. 73–76.
- [18] M. Hazim *et al.*, "Blend Shape Interpolation and FACS for Realistic Avatar," in *3D Research*, vol. 6, 2015, p. 6.
- [19] Livox, "Livox: para pessoas com deficiências e transtornos de aprendizagem," 2019. [Online]. Available: <http://www.livox.com.br/pt/quem-somos/#sobre>. [Accessed: 06-Dec-2018].
- [20] Tobii Dynavox logo, "Sono Flex for Communicator 5 - Tobii Dynavox," 2019. [Online]. Available: <https://www.tobiidynavox.com/software/content/sono-flex-for-communicator-5/>. [Accessed: 07-Dec-2018].
- [21] "TippyTalk - Comunicação Instantânea para Pessoas Não Verbais," 2016. [Online]. Available: <http://www.tippy-talk.com/>. [Accessed: 16-Jan-2019].
- [22] "Jogos de autismo." [Online]. Available: <http://www.autismgames.com.au/>. [Accessed: 14-Aug-2019].
- [23] S. Fridenson-Hayo *et al.*, "'Emotiplay': a serious game for learning about emotions in children with autism: results of a cross-cultural evaluation," *Eur. Child Adolesc. Psychiatry*, vol. 26, no. 8, pp. 979–992, Aug. 2017.
- [24] "Autism Emotion na App Store," 2012. [Online]. Available: <https://apps.apple.com/us/app/autism-emotion/id550027186>. [Accessed: 16-Aug-2019].
- [25] D. Deriso, J. Susskind, L. Krieger, and M. Bartlett, "Emotion Mirror: A Novel Intervention for Autism Based on Real-Time Expression Recognition," in *European Conference on Computer Vision*, 2012, pp. 671–674.
- [26] G. Littlewort *et al.*, "The computer expression recognition toolbox (CERT)," in *Face and Gesture 2011*, 2011, pp. 298–305.
- [27] M. P. Jeff Cockburn, Marni Bartlett, James Tanaka, Javier Movellan and and Robert Schultz,

- “SmileMaze: A Tutoring System in Real-Time Facial Expression Perception and Production in Children with Autism Spectrum Disorder,” *ECAG 2008 Work. Facial Bodily Expressions Control Adapt. Games*, 2008.
- [28] S. Alves *et al.*, “LIFEisGAME Prototype: A Serious Game about Emotions for Children with Autism Spectrum Disorders,” 2013.
- [29] P. Ekman and E. L. Rosenberg, “What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS), Second Edition.”
- [30] P. Santos, V. Silva, F. Soares, and A. Simões, “Facial Virtual Tracking: A System to Mirror Emotions,” in *19th EPIA Conference on Artificial Intelligence 3-6th September, Vila Real*, 2019, pp. 68–79.
- [31] Brandon Amos at Carnegie Mellon University, “OpenFace.” [Online]. Available: <https://cmusatyalab.github.io/openface/>. [Accessed: 21-Nov-2018].
- [32] Unity Technologies, “Unity - Manual: Unity User Manual (2017.4),” 2017. [Online]. Available: <https://docs.unity3d.com/2017.4/Documentation/Manual/index.html>. [Accessed: 21-Nov-2018].
- [33] “Mixamo.” [Online]. Available: <https://www.mixamo.com/#/>. [Accessed: 28-Nov-2018].
- [34] Saul McLeod, “Bandura - Social Learning Theory,” 2016.
- [35] V. Silva *et al.*, “Real-time emotions recognition system,” in *8th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, 2016, vol. 2016-Decem, pp. 201–206.
- [36] Maria Lucia Seidl de Moura and Adriana F. P. Ribas, “Imitação e desenvolvimento inicial: evidências empíricas, explicações e implicações teóricas,” *Estudos de Psicologia 2002*, 2002.
- [37] S. Costa, F. Soares, A. P. Pereira, C. Santos, and A. Hiolle, *A Pilot Study using Imitation and Storytelling Scenarios as Activities for Labelling Emotions by Children with Autism using a Humanoid Robot*. 2014.
- [38] “Filtering and Smoothing Data - MATLAB & Simulink.” [Online]. Available: <https://www.mathworks.com/help/curvefit/smoothing-data.html>. [Accessed: 21-Sep-2019].