

UNIVERSIDADE DO ALGARVE

*INTERFACES HUMANO-COMPUTADOR COM
RECONHECIMENTO DE GESTOS E REPRESENTAÇÃO
HOLOGRÁFICA*

LUÍS CARLOS RODRIGUES DE SOUSA

Dissertação

Mestrado em Engenharia Elétrica e Eletrónica

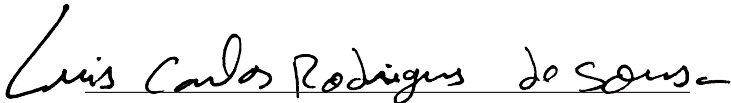
Trabalho efetuado sob a orientação de:
Professor Doutor Jânio Miguel Evangelista Ferreira Monteiro & Professor Doutor
João Miguel Fernandes Rodrigues

2015

INTERFACES HUMANO-COMPUTADOR COM RECONHECIMENTO DE GESTOS E REPRESENTAÇÃO HOLOGRÁFICA

Declaração de autoria de trabalho

Declaro ser o autor deste trabalho, que é original e inédito. Autores e trabalhos consultados estão devidamente citados no texto e constam da listagem de referências incluída.



Luís Carlos Rodrigues de Sousa

©2015, LUÍS CARLOS RODRIGUES DE SOUSA

A Universidade do Algarve tem o direito, perpétuo e sem limites geográficos, de arquivar e publicitar este trabalho através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, de o divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

Resumo

Hoje em dia, várias tecnologias estão em desenvolvimento com o intuito de permitir uma melhoria da interatividade humano-computador. No domínio do reconhecimento de instruções por exemplo, diversos dispositivos permitem determinar os gestos das mãos de utilizadores. Esta capacidade pode ser utilizada vantajosamente num grande número de aplicações, em especial naquelas em que os teclados e ecrãs táteis não são as melhores soluções. Há ainda outros dispositivos que são capazes de detetar todo o corpo do utilizador, permitindo entre outros, o reconhecimento dos gestos efetuados com o braço ou o corpo completo. Em todo o caso, a melhoria da interatividade humano-computador não se restringe ao reconhecimento de gestos. Um dos domínios que tem ganho grande relevância nos últimos anos é o da representação holográfica. Os hologramas permitem uma visualização de personagens e objetos mais próxima daquela a que estamos habituados, e por essa via, a mais natural. Nesta dissertação desenvolvem-se e combinam-se estes dois domínios. Começa-se por apresentar o estudo e desenvolvimento de vários sistemas de reconhecimento de gestos corporais, permitindo a construção de novas aplicações capazes de maximizar a interatividade com o computador. Em seguida o reconhecimento de gestos é combinado com a representação holográfica, suportando a criação e o desenvolvimento de uma rececionista, ou de uma relações públicas de uma empresa, produto ou evento. Ao combinar hologramas com reconhecimento de gestos, o sistema desenvolvido permite uma interação

dinâmica com os utilizadores, em resposta ao que lhes for solicitado, servindo de guia orientador e de fonte de informação útil. O resultado é uma instalação inovadora, capaz de produzir hologramas a partir de objetos virtuais ou de vídeos de pessoas reais que podem ser observados em toda a sua volta (360°).

Palavras-chave: Holograma, Reconhecimento de Gestos, Interfaces Humano-Computador.

Abstract

Nowadays, various technologies are under development to improve the human-computer interaction. In the field of human gesture's recognition, for instance, several devices can already be used to capture of the user's hands position. This capability can be used advantageously in many applications, particularly those in which keyboards and touch screens are not the best solutions. Besides these, there are also other devices that are capable of detecting any user's body, allowing the recognition of gestures made with the arm or the full body. Yet, the improvement of human-computer interaction extends far beyond the recognition of gestures. One of the areas that has gained great importance in recent years is the holographic representation. Holograms allow the viewing of characters and objects, in a representation that is closer to the one that we are used to, and in a more natural way. These are the two areas covered by the work done in this dissertation. It begins by presenting the study and development of various gesture recognition systems, allowing the construction of new applications that maximize interactivity with the computer. Then gesture recognition is combined with holographic representation, supporting the creation and development of a receptionist, a company's public relations, a product or event. By combining holograms with gesture recognition, the developed system allows the dynamic interaction with users, in response to what is requested, acting at the same time as an advisor guide and as a useful information source. The result is an innovative facility, capable of producing

holograms from virtual objects or real videos of people that can be seen all around it (360°).

Keywords: Hologram, Gesture Recognition, Human-Computer Interfaces.

*Aos meus pais,
por todo o esforço que passaram para eu chegar até aqui.
À minha restante família e namorada por todo o apoio e força.*

Agradecimentos

Em primeiro lugar, gostaria de agradecer ao meu orientador, Professor Doutor Jânio Monteiro e co-orientador Professor Doutor João Rodrigues por toda a ajuda, todo o esforço e disponibilidade que me facultaram para a elaboração desta dissertação.

Agradeço também aos meus colegas de projeto Ricardo Alves, Aldric Négrier e Jaime Martins por quase dois anos de trabalho como colegas e por toda a ajuda que compartilharam no projeto e artigos científicos.

Por fim, um obrigado à empresa SPIC - Creative Solutions, por todo o apoio, principalmente no design das aplicações do projeto PRHOLO.

Conteúdo

Agradecimentos	xi
Lista de Tabelas	xv
Lista de Figuras	xvi
Lista de Abreviaturas	xxi
Capítulo 1 Introdução	1
1.1 Enquadramento	4
1.2 Objetivos	5
1.3 Estrutura da dissertação	7
Capítulo 2 Estado da Arte	9
2.1 Interfaces humano-computador	9
2.2 Sistemas de reconhecimento de gestos	11
2.3 Instalações holográficas	12
Capítulo 3 Interfaces humano-computador baseadas em reconhecimento de gestos	15
3.1 Reconhecimento de gestos com Leap Motion	16
3.1.1 Aplicações	19
3.2 GyroGlove - Luva eletrónica com reconhecimento de gestos	24
3.2.1 Sensores de inércia	25
3.2.2 GyroGlove	27
3.2.3 Aplicação de aprendizagem do alfabeto da língua gestual portuguesa (LGP)	38
3.3 Testes e resultados	43
3.3.1 Reconhecimento de gestos com Leap Motion	44
3.3.2 Reconhecimento de gestos com o GyroGlove	45
Capítulo 4 Interface humano-computador aplicada a uma instalação holográfica 360°	47
4.1 Técnica <i>Pepper's Ghost</i>	48
4.2 Instalação	50
4.2.1 Especificações de hardware	59
4.3 Interface Humano-Computador	59
4.3.1 Base de dados	60

4.3.2	Sistema de Kinects e utilizadores	60
4.3.3	Sistema gerador de hologramas	60
4.3.4	Sistema de menus interativos	64
4.3.5	Reconhecimento de gestos	67
4.3.6	Reconhecimento de voz	69
4.3.7	Sistema RFID	70
4.3.8	Estatísticas	70
4.3.9	Aplicação Android	72
4.4	Resultados	73
Capítulo 5 Conclusões e trabalho futuro		77
5.1	Reconhecimento de gestos com Leap Motion	78
5.2	Sistema GyroGlove	79
5.3	Sistema holográfico	81
5.4	Lista de publicações	82
Bibliografia		85

Lista de Tabelas

3.1	Pacote de transmissão dos dados entre a luva e o PC	36
3.2	Relações entre sensores / partes da mão para comparação e detecção de gestos.	40
3.3	Tabela com dados quantitativos sobre os resultados de várias pessoas que experimentaram as aplicações.	44
3.4	Tabela com os resultados quantitativos dos testes realizados com o sistema de reconhecimento de gestos com a luva GyroGlove.	46

Lista de Figuras

3.1	Sistema de coordenadas do Leap Motion, imagem à esquerda (adaptado de citealpleap) e os seis tipos de movimentos retilíneos usando o sistema de reconhecimento de gestos, imagem à direita (adaptado de Sousa et al., 2014).	18
3.2	Na primeira linha, três diferentes zonas do modelo 3D da casa. Da esquerda para a direita, primeiro andar, rés de chão e sótão. Na segunda linha, a informação disponível de um equipamento elétrico selecionado entre configuração de novos programas, verificação de programas ativos, estatísticas e informações. Na última linha, a interface principal em utilização com o Leap Motion.	21
3.3	Interface para selecionar vários os transportes (cima) e de picking (baixo).	22
3.4	Interface usando o LeapMotion para o sistema PoolLiveAid. Utilização diretamente numa mesa de snooker à esquerda. Projeção do menu interativo numa parede à direita.	24
3.5	Uma PCB (Printed Circuit Board) de um sensor MPU-6000 (InvenSense, 2015)	27
3.6	A designação anatómica dos ossos da mão à esquerda. A localização dos sensores e módulo principal à direita.	28
3.7	Esquemático do circuito do módulo controlador e o seu respetivo desenho da PCB.	32
3.8	Esquemático do circuito do módulo recetor e o seu respetivo desenho da PCB.	34
3.9	As PCBs dos protótipos dos módulos finais do sistema, em cima. O protótipo da luva em baixo.	35
3.10	Esquema generalizado das comunicações entre os módulos principais.	36
3.11	Quatro exemplos da replicação da mão do utilizador num modelo 3D. Da esquerda para a direita e de cima para baixo. Mão inclinada para a esquerda, mão fechada, mão estendida e direita e mão semi-aberta com indicador apontado.	38
3.12	Alfabeto da Língua Gestual Portuguesa (LGP) (Wiki, 2014)	39
3.13	Relações entre sensores adjacentes e entre o eixo vertical	40
3.14	Aplicação de estimulação da aprendizagem do alfabeto da linguagem gestual portuguesa. Instruções da aplicação na imagem em cima, à esquerda. O design da aplicação na imagem de cima, à direita, e em baixo, duas imagens durante a utilização.	43

4.1	Técnica <i>Pepper's Ghost</i> . A imagem do objeto está colocada num plano horizontal enquanto uma película refletora transparente fica inclinada 45° sobre ela.	49
4.2	Modelos 3D de exemplos e soluções da estrutura holográfica 360° replicando a técnica <i>Pepper's Ghost</i> de uma forma circular. Nas duas imagens de cima pode-se observar a configuração de 4 e 8 faces, à esquerda e direita respetivamente. Estes são dois exemplos onde as faces estão desenhadas em forma de triângulo. Na imagem do meio à esquerda, a solução final para a estrutura holográfica aplicando a técnica <i>Pepper's Ghost</i> com 8 faces. Nas restantes figuras, as dimensões da da estrutura. A cor azul/roxo apenas tem como objetivo a melhor visualização das placas acrílicas. Na verdade, elas são transparentes.	52
4.3	Modelos 3D de peças de alumínio para fortalecer a estrutura holográfica. De cima para baixo e esquerda para a direita, junção de cima e baixo (partes azuis), espaçador entre faces (parte amarela), perfil de alumínio em forma de U (vermelho), perfil de alumínio em forma de L (verde) e estrutura montada e fixa. As cores apenas contribuem para a melhor visualização dos contornos, na verdade as peças são pretas para refletir a luz o menos possível.	54
4.4	Sequência da construção da estrutura base. De cima para baixo e esquerda para a direita, fixação do ecrã de projeção pela retaguarda, divisórias para bloquear luz vinda de outras perspetivas (interferência entre perspetivas), estrutura em forma de "fatia de bolo" com o mesmo propósito anterior e as suas dimensões à direita, posicionamento das câmaras Kinect e estrutura base. As restantes figuras são a estrutura vista de várias perspetivas e as suas dimensões.	55
4.5	Testes de reflexão de luz nos modelos 3D da instalação holográfica. De cima para baixo e da esquerda para a direita, as luzes de testes, observação diretamente a uma face e na esquina com correção de interferências e a mesma observação à face e esquina sem correção de interferências.	57
4.6	Ângulos do posicionamento entre cada Kinect à esquerda e visualização dos campos de visão das 8 Kinect nas duas imagens à direita.	58
4.7	Base de dados no servidor da estrutura holográfica.	61
4.8	Vários exemplos das imagens do holograma de 4 e 8 perspetivas. Da esquerda para a direita e de cima para baixo, video da cabeça de uma mulher com 4 perspetivas, editor no Unity com as câmaras a apontar para um avatar 3D, avatar amarelo com 8 perspetivas, o mesmo avatar com 4 perspetivas, carro com 8 perspetivas, um soldado e um carro.	63
4.9	Os 7 diferentes tipos de menus. Da esquerda para a direita e de cima para baixo, menu <i>Simples</i> , <i>Titulo_Texto</i> , <i>Titulo_Texto_Imagem</i> , <i>Diagonal</i> , <i>Imagem Mural</i> e <i>Selfie</i>	66
4.10	Os 3 diferentes tipos de personagens nos menus. Da esquerda para a direita, personagem 3D, videos de apresentação para cada menu e video de fundo (atrás das opções dos menus).	67
4.11	Mapas de calor de vista aérea numa instalação com 8 e 4 Kinect na imagem à esquerda e direita respetivamente.	71

4.12	A aplicação Android. Da esquerda para a direita, o menu lateral da aplicação, seleção dos menus da instalação holográfica, seleção dos menus em forma de comando, leitor de códigos QR e submissão de opiniões.	73
4.13	Vários resultados da instalação. Instalação completa na primeira linha e detalhes nas transições entre as perspetivas nas imagens das duas linhas centrais. As últimas 3 imagens da última linha representam outra instalação, baseada na descrita neste capítulo, onde as principais diferenças são o uso de 4 perspetivas/faces invés de 8 e apenas uma Kinect, colocada na perspetiva do ecrã lateral.	75

Lista de Abreviaturas

API	<i>Application Programming Interface</i>
DMP	<i>Digital Motion Processor</i>
DoF	<i>Degrees of Freedom</i>
GUI	<i>Grafical User Interface</i>
HCI	<i>Human-Computer Interaction</i>
IMU	<i>Inertial Measurement Unit</i>
JSON	<i>JavaScript Object Notation</i>
LGP	<i>Língua Gestual Portuguesa</i>
NUI	<i>Natural User Interface</i>
PCB	<i>Printed Circuit Board</i>
PRHOLO	<i>O Relações Públicas Holográfico Realista - The realistic holographic public relations</i>
SDK	<i>Software Development Kit</i>
VUI	<i>Voice User Interface</i>

1

Introdução

A evolução dos computadores durante as últimas décadas, obrigou à criação de interfaces humano-computador (HCI) para o seu controlo cada vez mais simples e naturais pelos utilizadores. Os antigos cartões perfurados, na altura essenciais para controlar um computador, rapidamente foram descontinuados com o aparecimento de novas tecnologias de interfaces, tais como o rato e teclado que ainda são usados até hoje. Paralelamente, com o aumento das vendas dos produtos de eletrónica de consumo, procurou-se sempre desenvolver novas formas de interação, mais fáceis de usar, naturais para o ser-humano e que permitissem um controlo mais intensivo e imersivo entre este e as tecnologias. Estes desenvolvimentos foram acompanhados também pela criação de novas interfaces gráficas (GUI) que melhoram a interação humano-

computador.

Os filmes de ficção científica são bons exemplos dessa evolução. As tecnologias aí apresentadas, não totalmente reais, promovem o desenvolvimento das novas formas de interação com computadores (Wells, 2002), estimulam a criatividade, que por sua vez tornam a ficção realidade.

Embora estas interfaces tenham verificado uma grande evolução nos últimos anos, as exigências deste tipo de tecnologias também aumentam. Essas expectativas levaram a que recentemente se iniciassem pesquisas no sentido de desenvolver métodos inovadores de representação holográfica e do seu respetivo controlo.

No âmbito do controlo, existem atualmente vários sensores com a capacidade de *tracking* e reconhecimento de gestos corporais tais como a Kinect (Kinect, 2014), Leap Motion (Motion, 2014), Structure Sensor (Structure, 2014), Asus Xtion (ASUSTeK, 2015), entre outros. Todos estes sensores apresentam uma grande importância para as indústrias de videojogos e ferramentas de interação utilizador-máquina. Esses sensores, quando complementados com software apropriado, têm a capacidade de detetar a estrutura corporal e/ou a mão do utilizador, replicar com precisão essa estrutura numa *3D mesh* (replicação virtual do utilizador) e efetuar a deteção de gestos. Devido ao seu preço e dimensões, podem na sua maioria ser facilmente integrados em vários locais dentro de uma casa permitindo o controlo dos mais variados tipos de aparelhos ou eletrodomésticos (desde computadores, máquinas de lavar, a fogões, etc.) sem que para tal tenhamos que recorrer aos periféricos de controlo clássicos, tais como *touchscreens*, teclados, ratos, etc., aumentando por essa via a acessibilidade e interatividade com o utilizador. Para além disso, as interfaces baseadas em gestos (sem toque físico) podem ser bastantes úteis em situações onde os periféricos normais não são de toda a melhor solução a ser usada. A título de exemplo, pode-se referir o carregamento de um camião de transporte, em que a atividade implica o uso de luvas ou acarreta ter as mãos sujas (e.g., camiões de lixo), ou um vendedor de peixe num supermercado ou na praça local, onde a necessidade de manusear o peixe e ao mesmo tempo ter acesso a

um terminal eletrônico, para fazer a conta ou pesagem, não é de todo a mais prática.

Para aplicações que requerem apenas o reconhecimento intuitivo de gestos efetuados pela mão, o Leap Motion (Motion, 2014) é uma das melhores opções devido não só ao seu baixo preço, comparativamente aos outros, mas também pela sua rapidez, fiabilidade, alta resolução e área de observação/atuação. Tem ainda como vantagem, a sua diminuta dimensão, podendo ser facilmente ocultado ou embutido numa mesa, numa parede ou num veículo.

Todos os sensores atrás referidos são baseados em câmaras a cores (RGB) e/ou de profundidade (infravermelhos) e conseqüentemente enfrentam problemas de limitações espaciais, i.e., o utilizador tem que se situar perto do dispositivo, ou na área para onde estas câmaras estão a apontar, caso contrário não funcionam corretamente. Para além da limitação espacial, na maioria dos casos, estes dispositivos não funcionam quando estão próximos de uma fonte de infravermelhos, por exemplo no ar livre ou numa sala com bastantes lâmpadas fluorescentes.

Para os contextos onde os sensores baseados em captação de imagens podem não ser a melhor solução, existe a possibilidade de utilizar um sistema de *tracking* e reconhecimento de gestos, que utiliza sensores denominados de *Inertial Measurement Units* (IMU). Estes sensores quando colocados na mão e dedos do utilizador em posições específicas, permitem captar a rotação 3D dos mesmos.

A melhoria da interatividade humano-computador não se restringe no entanto ao reconhecimento de gestos. Um dos domínios que tem ganho grande relevância nos últimos anos é o da representação holográfica. Os hologramas permitem uma visualização de personagens e objetos mais próxima daquela a que estamos habituados, e por essa via mais natural. Para as empresas esta tecnologia pode ser utilizada vantajosamente. Conseguir apresentar novos produtos, manter clientes regulares e cativar novos, embora sejam os principais objetivos de uma empresa, nem sempre é tarefa fácil. O primeiro contacto de um cliente com a empresa passa pelo seu *website*, mas por vezes a falta ou pouca informação poderá deixar questões em aberto. Por este

motivo, por vezes também se faz do uso de relações públicas, persuadindo o mercado e os clientes dos novos produtos da empresa de uma forma mais pessoal e específica. Embora esta última seja uma solução razoavelmente eficaz, existem soluções tecnológicas capazes de atingir melhores resultados e de forma diferente, captando o cliente pela originalidade e pela eficácia.

Os sistemas holográficos são uma opção com carácter inovador que uma empresa possui para conseguir enfrentar o mercado. Os hologramas são dos melhores efeitos especiais que um filme de ficção científica pode conter, cativando clientes na indústria do entretenimento e por isso desde à muito tempo que empresas tentem trazer esses hologramas ficcionais para o mundo real, em produtos reais.

Pelos pontos apresentados acima, nesta dissertação começa-se por apresentar o estudo e desenvolvimento de vários sistemas de reconhecimento de gestos corporais, permitindo a construção de novas aplicações capazes de maximizar a interatividade com o computador. Em seguida o reconhecimento de gestos é combinado com a representação holográfica, suportando a criação e o desenvolvimento de uma rececionista, ou de uma relações públicas de uma empresa, produto ou evento. Ao combinar hologramas com reconhecimento de gestos, o sistema desenvolvido permite uma interação dinâmica com os utilizadores, em resposta ao que lhes for solicitado, servindo de guia orientador e de fonte de informação útil. O resultado é uma instalação inovadora, capaz de produzir hologramas a partir de objetos virtuais ou de vídeos de pessoas reais que podem ser observados em toda a sua volta (360°).

1.1 Enquadramento

Parte desta dissertação foi desenvolvida através de "PRHOLO: O Relações Públicas Holográfico Realista", um projeto financiado pelo QREN I&DT (nº 33845) que tem como promotor a empresa SPIC - Creative Solutions. Este projeto teve como objetivo, o desenvolvimento de novas experiências holográficas que visam a ter funções de re-

cecionista, de relações públicas de uma empresa, produto ou evento que interagirá dinamicamente com os utilizadores conforme o que lhes for solicitado, servindo de guia orientador e de fonte de informação útil. Baseado neste projeto, também foram desenvolvidos, como complemento, várias novas formas de interação com o sistema holográfico acima referido, como também interfaces humano-computador aplicadas a situações específicas que necessitam de novas formas de controlo, mais fáceis e cómodas.

1.2 Objetivos

Os principais objetivos desta dissertação são o desenvolvimento de i) novas formas de controlo e interfaces humano-computador para situações onde dispositivos convencionais não são os adequados, ii) uma instalação holográfica inovadora, a implementar no âmbito do projeto PRHolo, e iii) uma interface com reconhecimento de gestos para o controlo da instalação holográfica.

Especificando os principais objetivos de forma mais detalhada:

- **a) Instalação holográfica 360º:** Sendo este um dos objetivos do projeto PRHolo, pretende-se a construção de uma estrutura holográfica que compreende as seguintes funcionalidades:
 - Módulo que apresente um holograma volumétrico numa estrutura invisível. O holograma não pode estar dependente de estruturas sólidas opacas na zona onde o mesmo é apresentado de forma a tornar o holograma o mais realístico possível.
 - Colocação de câmaras Kinect na estrutura que permitem detetar os utilizadores à volta da instalação.
 - Sistema de captação de som para fazer o reconhecimento de voz dos utilizadores, permitindo aos mesmos interagir com a instalação. Este sistema

também terá que reconhecer a direção do som.

- Sistema de reconhecimento de gestos para permitir o controlo e interação por gestos corporais com a instalação.
 - Sistema RFID que permite os utilizadores se apresentarem ao holograma, causando uma ação/resposta por parte do holograma.
 - Uma interface gráfica com conteúdos em formas de textos, imagens e vídeos.
 - Gerador de dados estatísticos sobre a posição, número de gestos, tempos de utilização, entre outros, de todas os utilizadores que interajam com a instalação.
 - Produção de um mapa de calor que mostre visualmente, ao longo de um espaço de tempo, os locais à volta da instalação nos quais as pessoas mais frequentaram.
 - Construção de uma base de dados que grave todos os dados estatísticos, imagens, textos, vídeos, hologramas e configurações necessárias.
- **b) Sistemas de interação humano-computador:** Desenvolvimento de novos sistemas que consigam substituir os dispositivos convencionais para situações onde os mesmos não são os mais indicados. No decorrer desta dissertação, o desenvolvimento destes sistemas será realizado em duas componentes:
 - Aplicações com interfaces naturais do utilizador usando o sensor Leap Motion, uma câmara especial capaz de detetar gestos produzidos com as mãos dos utilizadores.
 - Uma luva eletrónica com capacidades de reconhecimento de gestos e interação com uma aplicação de aprendizagem de linguagem gestual.

1.3 Estrutura da dissertação

Esta dissertação encontra-se dividida em 5 capítulos, no presente capítulo foi introduzido o tema da dissertação, a sua contextualização e objetivos. O capítulo 2, referente ao estado da arte, contém uma revisão às tecnologias já existentes no mercado, que suportam os desenvolvimentos efetuados no contexto desta dissertação.

O capítulo 3 apresenta o desenvolvimento de várias interfaces humano-computador divididas em duas secções. Na primeira descreve-se o desenvolvimento de sistemas de interação humano-computador baseadas em problemas reais, utilizando um sensor de reconhecimento de gestos. Na segunda secção, e em complemento às tecnologias existentes, é apresentado o desenvolvimento de uma tecnologia interativa baseada numa luva eletrónica com capacidade de reconhecimento de gestos. Será descrito nesta segunda secção todo o desenvolvimento efetuado tanto em termos de hardware como de software e no final do mesmo, uma aplicação com utilidades educativas suportada pela tecnologia criada.

No capítulo 4 será apresentado o desenvolvimento de um holograma 360° (PRHOLO: O Relações Públicas Holográfico Realista) e interfaces interativas que podem ser controladas pelo utilizador. O capítulo está dividido em 4 partes principais: i) o desenvolvimento de todo o hardware da estrutura holográfica, ii) um programa com uma interface apresentada num ecrã lateral na instalação para a visualização de conteúdos e controlo, iii) outro programa capaz de gerar as figuras holográficas e iv) uma aplicação Android complementar à interface de controlo do holograma e respetivos conteúdos.

Por fim, no capítulo 5, são apresentadas as conclusões e as propostas de trabalho futuro.

2

Estado da Arte

Este capítulo apresenta uma análise dos avanços tecnológicos das tecnologias de interação dos humanos com o computador, de reconhecimento de gestos usando essas tecnologias, de instalações holográficas e de formas de interação com hologramas existentes no mercado.

2.1 Interfaces humano-computador

Uma interface natural de utilizador (*Natural User Interface* ou NUI) é um termo usado para interfaces humano-computador (*Human-Computer Interaction* ou HCI) onde a interface é invisível ou se torna invisível após sucessivos níveis de imersão pelo utiliza-

dor. Tipicamente baseia-se na natureza, ou em elementos naturais humanos.

Nas últimas décadas, as formas de interação com o computador envolveram (e envolvem) botões, monitores, cartões perfurados, teclados, ratos, ecrãs sensíveis ao toque...etc. A evolução de todas estas tecnologias tem sido tão grande que os filmes de ficção científica as incluem e tentem prever o seu futuro tecnológico. Por exemplo em Kubrick (1968), um computador preenchido de inteligência artificial (*HAL 9000*) e os humanos mantêm diálogos complexos. Noutros exemplos, em Wachowski (1999) e Spielberg (2002) utilizam-se interfaces tecnológicas futuristas para a época em que esses filmes foram realizados.

Atualmente, a maioria das interfaces são do tipo GUI (*Graphical User Interface*) ou VUI (*Voice User Interface*). Como os próprios nomes indicam, nas primeiras os utilizadores interagem com interfaces gráficas, enquanto nas segundas é usado o reconhecimento de voz para enviar ordens de controlo aos dispositivos eletrónicos e até o reconhecimento de emoções (Rybka and Janicki, 2013). Mais recentemente, várias novas formas de interação (Dumas et al., 2009) permitiram que as pessoas interagissem com máquinas, computadores de uma forma que não era possível anteriormente com outros paradigmas.

Um dos recentes, novos paradigmas de interação humano-computador surgiu com a chegada de vários tipos de sensores 3D como o Leap Motion (Motion, 2014), a câmara Kinect (Kinect, 2014), o Structure Sensor (Structure, 2014) ou o Asus Xtion (ASUS-TeKense, 2015). Estes sensores permitem, através das suas APIs (*Application Programming Interface*), reconhecer gestos (humanos) permitindo o controlo de aplicações e manipulação de objetos num mundo virtual ou a interação com aplicações de realidade aumentada sem necessidade de tocar fisicamente num teclado ou rato. Na sua vasta maioria, estes sensores começaram a ser usados nas indústrias de jogos e entretenimento, mas a sua capacidade vai muito além disso, incluindo o reconhecimento de gestos corporais e *tracking* para instalações artísticas interativas (Alves et al., 2014), aplicações para ajuda de idosos ou pessoas portadoras de deficiência (Chung et al.,

2014), pintar sem necessidade de tocar em nenhum material físico (Sutton (2013)) ou aplicações de linguagem gestual (Potter et al., 2013), entre muitas outras possibilidades.

2.2 Sistemas de reconhecimento de gestos

O Leap Motion, referido na secção anterior, é um sensor capaz de reconhecer as mãos do utilizador. Este sensor é bastante pequeno, possuindo uma dimensão de apenas 8 cm (centímetros) de comprimento, 3 cm de largura e 1,25 cm de altura. De uma forma semelhante aos sensores apresentados anteriormente (e.g., Kinect e Asus Xtion), para conseguir detetar as mãos do utilizador, este sensor emite para a sua frente feixes de luz de infravermelhos a partir de três fontes. Um objeto que esteja próximo do sensor reflete alguma desta luz de volta, acabando por ser detetado por uma ou ambas as câmaras de infravermelhos colocadas dentro do sensor. Desta forma, e fazendo uso de visão computacional das imagens recebidas pelas duas câmaras, o sensor consegue detetar esses objetos e perceber que tipo de objetos são. Como o objetivo dos criadores do sistema Leap Motion foi o de reconhecer as mãos do utilizador de forma a poder-se tocar um piano virtual, a maioria do seu esforço foi colocado na tarefa de conseguir reconhecer com grande exatidão e rapidez as posições, rotações, tamanho, entre outros, de toda a mão e dedos. Além das características referidas anteriormente, o sensor possui uma API capaz de fornecer ao programador todos os dados que o sensor consegue detetar (i.e., posições das mãos e dedos), propriedades (velocidades, distâncias, etc.), ou um pouco mais avançado, que tipos de gestos o utilizador efetuou (i.e., movimentos retilíneos numa determinada direção com a mão estendida, *swipe*, clique com o dedo ou um círculo com o dedo).

Existem outros dispositivos e/ou tecnologias e técnicas diferentes capazes de detetar as mãos e reconhecer gestos. Chen et al. (2003) e Yoon et al. (2001) são alguns exemplos que definem métodos, através de visão computacional, capazes de reconhecer

as mãos e fazer reconhecimento de gestos através de uma câmara de vídeo normal. Usando tecnologia mais recente, Ren et al. (2011, 2013) e Li (2012) fazem uso de uma câmara 3D, a Microsoft Kinect para atingir o mesmo objetivo.

Existem outras técnicas que não fazem uso de nenhum tipo de câmara nem de algoritmos de visão computacional para extrair os mesmos dados das soluções anteriores. Dois destes casos, Patel et al. (2011) e Kessler et al. (1995), são exemplos que usam luva eletrônica especial contendo sensores em forma de fitas de resistências flexíveis em todos os 5 dedos da mão. A resistividade do sensor (resistência) altera-se consoante o quanto curvado o mesmo se encontra, e dessa forma, os autores conseguem extrair (embora de uma forma algo limitada) a rotação e conseqüentemente a posição de cada dedo em relação à palma da mão, conseguindo assim criar uma nova interface humano-computador.

2.3 Instalações holográficas

Várias tecnologias, aplicações e/ou instalações têm sido desenvolvidas ao longo dos anos usando a holografia, por exemplo Antonio et al. (2013) e Mihaylova et al. (2013). Uma das técnicas mais conhecidas e populares é a técnica de *Pepper's Ghost* pela sua simplicidade em criar a ilusão de um holograma apenas colocando um vidro/acrílico transparente inclinado de 45° sobre uma imagem num ecrã.

Esta técnica simples, foi primeiramente usada por John Henry Pepper, um mágico que num dos seus espetáculos, usou-a para lutar contra um fantasma (holograma) em pleno palco em 1862. Figueiredo et al. (2014) usa esta técnica aplicando-a no ensino para visualizar partes mecânicas em 3D. A técnica também tem aplicações nas áreas da música com a imortalidade do cantor Tupac que subitamente aparece no palco a cantar ao lado de Snoop Dogg (Rennie, 2015) e também Flyway (2015a), um concerto usando dois músicos em forma de holograma. Para criar estas ilusões, foi colocado um projetor entre o público e o palco, projetando os objetos/imagens que se quer apresentar no

holograma, numa película refletora no chão. Uma película transparente está colocada em todo o palco, inclinada 45° sobre o público refletindo a luz vinda da projeção no chão e criando a ilusão de um holograma do conteúdo que se pretende.

Relações públicas virtuais também são algumas aplicações. Uma delas, Flyway (2015b), usa uma assistente em tamanho real na forma de um holograma. Contudo, nenhuma destas aplicações conseguem apresentar uma imagem holográfica em todo o seu redor (holograma 360°) e com a capacidade de interação pelos utilizadores através de gestos e comandos de voz.

D'Strict (2014) é uma aplicação onde o holograma é colocado dentro de uma pequena caixa e pode ser controlado por gestos mas que apenas contém uma perspetiva. Active8-3D (2015) já possui uma dimensão maior mas apenas é observável (sem distorções) quando o utilizador se coloca mesmo na sua frente, possuindo assim uma interação muito limitada. Cheoptics 360 é um holograma com mais do que uma perspetiva, usando a técnica de *Pepper's Ghost* em forma de pirâmide, criando assim um holograma com 4 perspetivas e visível a 360° (Vizoo, 2015), mas não possui interação.

Todas estas tecnologias possuem todas entre 1-4 perspetivas e quase todas elas possuem limitações nas soluções de interação com os utilizadores.

Para mais detalhes sobre a técnica *Pepper's Ghost* descrita, ver secção 4.2.

3

Interfaces humano-computador baseadas em reconhecimento de gestos

Neste capítulo, analisam-se e implementam-se várias soluções/sistemas de reconhecimento de gestos. Começa-se por apresentar e desenvolver um método de deteção de vários tipos de gestos efetuados com as mãos, recorrendo ao Leap Motion. Com base neste sensor, implementam-se várias interfaces humano-computador que permitem a substituição dos tradicionais periféricos de computador em situações onde estes últimos não são uma solução cómoda.

Em comparação e em alternativa ao dispositivo Leap Motion, disponível comercialmente, também será apresentado o desenvolvimento de uma luva eletrónica capaz de

reconhecer gestos manuais através de sensores 3D eletrónicos. Todos os passos do desenvolvimentos desta luva são descritos em detalhe, desde a eletrónica introduzida na luva até ao software capaz de traduzir os dados recolhidos da luva em deteções de gestos. No final do capítulo será ainda apresentada uma aplicação com finalidades educativas através do reconhecimento de gestos de linguagem gestual com o objetivo de facilitar a sua aprendizagem.

3.1 Reconhecimento de gestos com Leap Motion

O Leap Motion possui uma API capaz de detetar vários gestos com as mãos. A configuração do dispositivo permite fazer o reconhecimento de vários tipos de gestos tais como reconhecer um movimento circular com um dedo, movimento retilíneo com a mão com os dedos estendidos, movimento para baixo com um dedo (como teclar) e a abrir ou fechar a mão. Outros exemplos são o movimento de lado a lado com a mão (*swipe*) como também reconhecimento de ferramentas na mão, como por exemplo um lápis. O dispositivo funciona muito bem quando a mão é colocada próxima do mesmo e tem a grande vantagem de adquirir todos os dados das mãos a 60 *frames* por segundo.

O Leap Motion faz medições de distâncias em milímetros, tempos em micro-segundos, velocidade em milímetros/segundo e ângulos em radianos. Na altura de escrita deste capítulo, a API suporta várias linguagens de programação tais como C++, C#, Objective-C, Java, Python, JavaScript e as duas plataformas de jogos UnrealEngine e Unity3D.

Como referido atrás, o Leap Motion consegue detetar vários tipos de gestos. Para para cada tipo de gestos existem algumas propriedades configuráveis na sua API com a finalidade de melhorar a sua deteção. Por exemplo no movimento do círculo com o dedo, existem duas propriedades configuráveis: o raio mínimo e o arco mínimo, sendo que por omissão são usados os valores de 5mm (milímetros) e $1,5\pi$ radianos respetivamente. Para o caso do *swipe*, podem ser seleccionáveis a distância mínima e

a velocidade mínima (por omissão são usados os valores de 150mm e 1000mm/s , respectivamente). Estes são alguns tipos de reconhecimento de gestos que a API fornece ao programador de uma forma bastante fácil e intuitiva. Se se juntarem os dados da posição e orientação dos dedos e mão, como também as velocidades dos movimentos, é possível detetar outros tipos de gestos mais avançados ou complexos. Para o desenvolvimento de uma interface simples, podem usar-se aqueles gestos que contêm o mínimo de movimentos e são mais intuitivos (simples), como por exemplo os *swipes*. Para estes tipos de gestos, a API também fornece a informação do vetor direção 3D que ajuda a diferenciar vários tipos de *swipe*, como por exemplo, se o movimento foi feito de cima para baixo, da esquerda para a direita, ou até por exemplo na diagonal. Este vetor 3D contém valores que variam entre $-1,0$ até $1,0$ de acordo com o sistema de coordenadas cartesianas ortogonal padrão, conforme representado na Fig. 3.1, esquerda. A origem do sistema de coordenadas encontra-se no centro do topo do dispositivo, sendo a sua “frente” o lado com a luz verde. O eixo x situa-se horizontalmente ao longo do dispositivo com os valores a incrementarem do lado esquerdo para o lado direito. O eixo z também está colocado na horizontal, perpendicular ao eixo x e com os valores a incrementarem na direção da parte frontal do dispositivo. Finalmente, o eixo y está colocado na vertical com os valores a incrementarem para cima.

Como o gesto de *swipe* pode ser qualquer um que cumpra os requisitos de velocidade e distância mínima, eles podem ter qualquer direção. Para uma aplicação de interface pode-se usar a informação do vetor direção para conseguir distinguir vários tipos de *swipes*. Para maximizar o número de diferentes tipos de movimentos independentes reconhecidos, estes podem dividir-se em seis categorias, sendo que três deles são o oposto dos outros três, como também se pode observar na Fig. 3.1, direita:

- a) Para cima e para baixo: movimento retilíneo de baixo para cima ou de cima para baixo, respetivamente.
- b) Para a esquerda e para a direita: movimento retilíneo de baixo para cima ou de

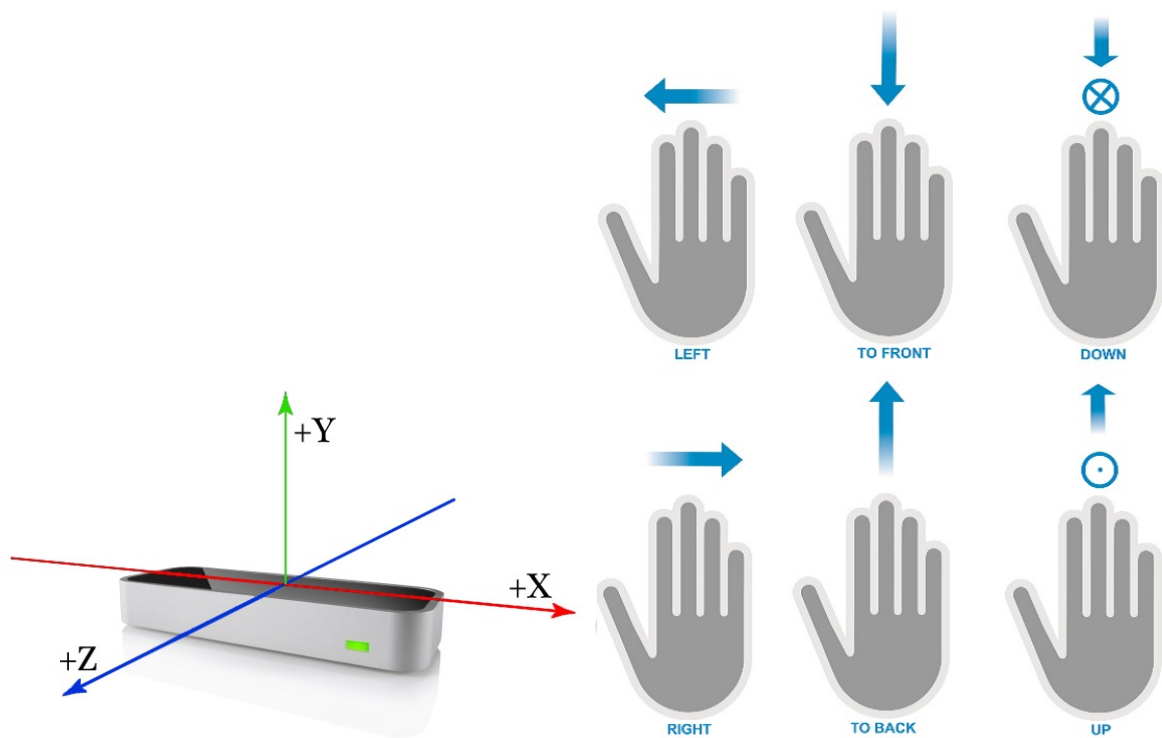


Figura 3.1: Sistema de coordenadas do Leap Motion, imagem à esquerda (adaptado de citealpleap) e os seis tipos de movimentos retilíneos usando o sistema de reconhecimento de gestos, imagem à direita (adaptado de Sousa et al., 2014).

cima para baixo, respetivamente.

- c) Para a esquerda e para a direita: movimento retilíneo de baixo para cima ou de cima para baixo, respetivamente.

No primeiro caso, a), um movimento de baixo para cima ou de cima para baixo depende maioritariamente do eixo vertical (y). Se o vetor direção possuir um valor de $y \approx +1$, com todos os outros valores próximos de zero, então será considerado que foi detetado o movimento de baixo para cima. Pelo contrário, se o vetor direção possuir um valor de $y \approx -1$, mais uma vez com todos os outros valores próximos de zero, será considerado que o utilizador efetuou o movimento de cima para baixo. Como é quase impossível um utilizador conseguir fazer um gesto de *swipe* com o vetor de direção exatamente $x = 0 \wedge y = \pm 1 \wedge z = 0$, é necessário selecionar um conjunto de valores para diferenciar entre os seis tipos de gestos. Para qualquer gesto que contenha um vetor direção que concorde com a condição $y \leq -0,5 \wedge |x| \leq 0,5 \wedge |z| \leq 0,5$,

será considerado um *swipe* de cima para baixo. Pelo contrário, se o vetor direção concordar com a condição $y \geq 0,5 \wedge |x| \leq 0,5 \wedge |z| \leq 0,5$, será considerado um *swipe* de baixo para cima. No segundo caso, b), o eixo que requer atenção maioritária é o z. Como se pode observar pela Fig. 3.1 esquerda, o vetor direção com um valor no eixo z de aproximadamente 1 será considerado um *swipe* de trás para a frente, caso contrário, da frente para trás. Semelhante a a), se um vetor direção concordar com a condição $z \leq -0,5 \wedge |x| \leq 0,5 \wedge |y| \leq 0,5$, será considerado um *swipe* de frente para trás. Pelo contrário, se o vetor direção concordar com a condição $z \geq 0,5 \wedge |x| \leq 0,5 \wedge |y| \leq 0,5$, será considerado um *swipe* de trás para a frente. No terceiro e último caso, c), qualquer *swipe* que possua um vetor direção que concorde com a condição $x \leq -0,5 \wedge |y| \leq 0,5 \wedge |z| \leq 0,5$, será considerado um *swipe* da direita para a esquerda. Pelo contrário, se o vetor direção concordar com a condição $x \geq 0,5 \wedge |y| \leq 0,5 \wedge |z| \leq 0,5$, será considerado um *swipe* da esquerda para a direita.

3.1.1 Aplicações

Nesta secção serão apresentados três protótipos utilizando o Leap Motion, desenvolvidos para aplicações onde uma interface com o Leap Motion é uma mais valia. O primeiro aborda o controlo de aparelhos / eletrodomésticos inteligentes e eficiência energética em edifícios inteligentes. No segundo caso, uma aplicação para controlo de *picking* e *loading* de veículos de transporte de mercadorias onde os trabalhadores têm dificuldades associadas com o manuseamento de dispositivos de controlo de um PC, (tais como o teclado e o rato) e no último caso, uma aplicação para o controlo do sistema PoolLiveAid (Alves et al., 2013).

Controlo de dispositivos e eletrodomésticos numa casa inteligente

Como em muitos outros domínios, a fácil interação humano-computador suportada em soluções intuitivas e simples, capazes de efetuar a monitorização e controlo traz claras vantagens na configuração de dispositivos elétricos, tais como eletrodomésti-

cos. No domínio das redes energéticas inteligentes, está-se neste momento a criar uma gama de novos dispositivos de controlo inteligente de edifícios que permite ajustar o perfil de consumo das habitações ou empresas à produção decorrente de energias renováveis ou aos tarifários energéticos (Monteiro et al., 2014a,b). Para que tal se torne possível, num futuro próximo teremos em nossa casa uma rede de equipamentos capazes de comunicar entre si, e de se coordenar com vista a reduzir o custo dos consumos energéticos, respondendo à interação do utilizador. Neste cenário, um dos domínios importantes é justamente o da interação com os utilizadores da casa. Para que se consiga melhorar a experiência de interação do utilizador, em seguida descreve-se o desenvolvimento de um sistema de reconhecimento de gestos 3D utilizando o Leap Motion que permite navegar pela planta de uma casa ou edifício e monitorizar e controlar um conjunto de equipamentos elétricos nela presentes (Sousa et al., 2014). Para se poder implementar a navegação, começou-se por construir um modelo 3D da casa, usando o programa Sweet Home 3D (SweetHome, 2015). A casa apresentada contém três andares criando um total de três diferentes áreas, como se pode observar na Fig. 3.2, 1ª linha. Cada um destes andares contém vários tipos de equipamentos elétricos tais como máquinas de lavar loiça, roupa, luzes etc. A partir do modelo tridimensional, a restante interface gráfica foi desenvolvida em Unity 3D (Unity, 2014), que por entre outros motivos, é uma excelente ferramenta de integração e manipulação de objetos 3D. Para se conseguir navegar pelo modelo da casa, a interface foi dividida em três menus:

- (1) Menu principal, onde o utilizador pode navegar entre os vários andares.
- (2) Menu por andar, onde o utilizador pode navegar pelos vários equipamentos desse andar.
- (3) Menu de equipamento, onde ao selecionar um determinado equipamento pode-se observar informações de estatísticas de consumos e programas do mesmo.

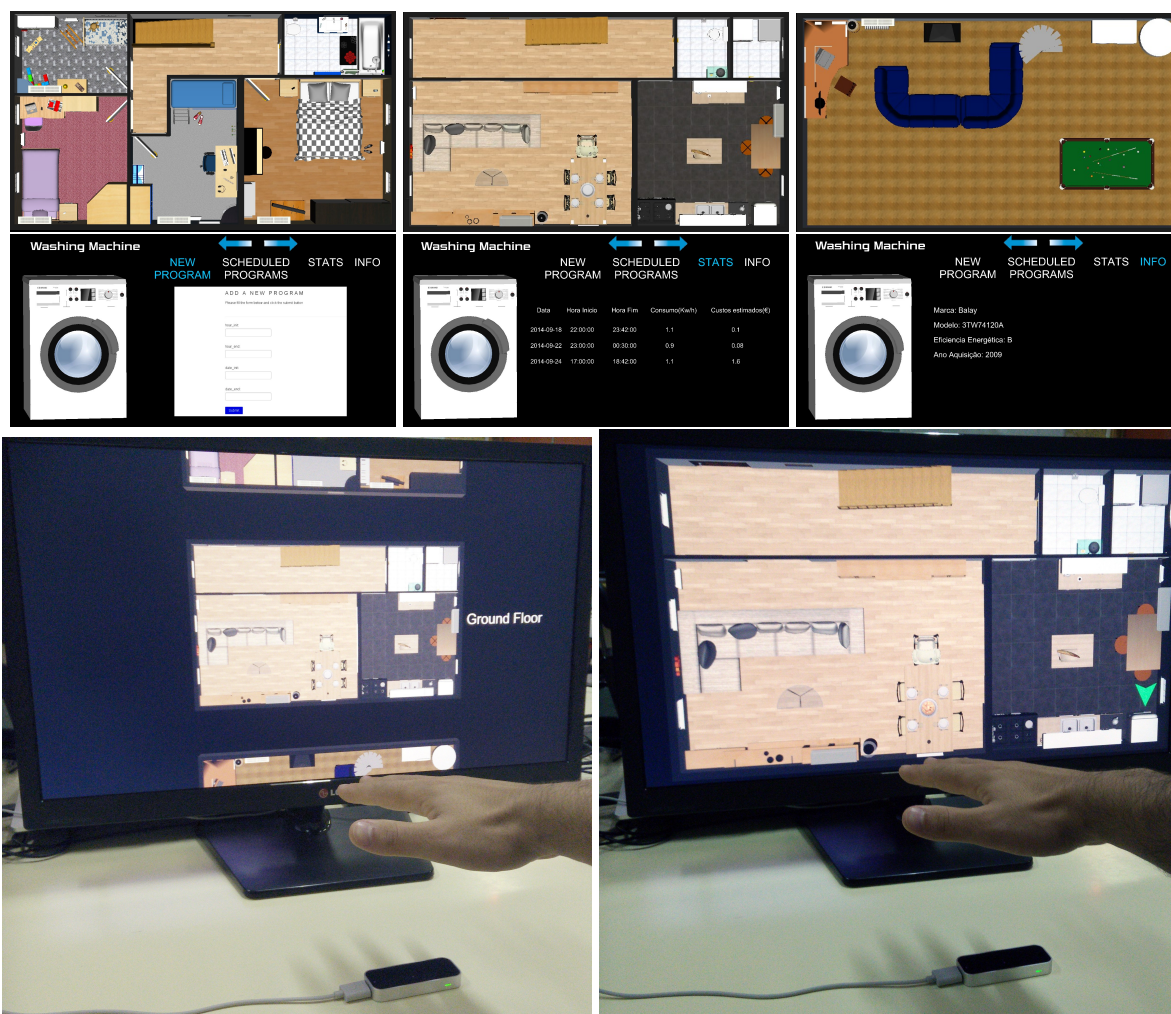


Figura 3.2: Na primeira linha, três diferentes zonas do modelo 3D da casa. Da esquerda para a direita, primeiro andar, rés de chão e sótão. Na segunda linha, a informação disponível de um equipamento elétrico selecionado entre configuração de novos programas, verificação de programas ativos, estatísticas e informações. Na última linha, a interface principal em utilização com o Leap Motion.

Usando o reconhecimento de gestos do Leap Motion, se o utilizador efetuar o *swipe* para a frente e para atrás, pode navegar entre os vários andares (Fig. 3.2, 1ª linha). Para selecionar um andar ou selecionar um equipamento é necessário efetuar o *swipe* para baixo, enquanto para desselecionar, voltar ao menu anterior, é usado o *swipe* para cima. Para navegar entre os vários equipamentos de um andar ou entre as várias informações de um determinado equipamento pode-se usar o *swipe* para a esquerda ou direita. Na Fig. 3.2, 3ª linha, pode-se observar o sistema em utilização enquanto se utiliza o menu principal, e na Fig. 3.2, 2ª linha, a informação disponível ao selecionar

um determinado equipamento.

Aplicação para controlo de *picking* e *loading* de veículos

No mercado dos transportes de mercadoria existe a necessidade de proceder ao *picking* e *loading* de veículos. Em certos casos, o utilizador que está a realizar o processo



Figura 3.3: Interface para seleccionar vários os transportes (cima) e de picking (baixo).

possui luvas ou sujidade nas mãos, o que pode tornar difícil a interação com um computador através dos periféricos clássicos, como sejam eles o teclado e rato. Por este motivo foi desenvolvida uma segunda solução de uma interface que permite ao utilizador interagir com a interface de gestão logística de forma mais simples e cómoda

(Cardoso et al., 2015). Deste modo, a aplicação está encarregue de obter as informações necessárias da base de dados da empresa de transportes e de a disponibilizar numa lista interativa com todas as informações necessárias da carga dos veículos. Assim, no menu principal apresentado na Fig. 3.3, em cima, é disponibilizada a lista de veículos a serem carregados incluindo: informações de matrícula, estado atual de carga, mapa do percurso, distância do percurso e horários. Como na aplicação anterior, esta interface também possui o reconhecimento de gestos usando o sensor Leap Motion. Se o utilizador efetuar o *swipe* para a frente ou para trás consegue navegar entre os vários veículos. Para selecionar o transporte respetivo basta efetuar *swipe* para baixo e a aplicação entra no segundo menu (Fig. 3.3 em baixo). Este sub-menu contém uma lista de todos os produtos e encomendas, que são necessários carregar. Quando um item da lista é carregado no veículo, o utilizador pode selecionar esse item (da mesma forma que foi feito no menu principal) e ao efetuar o *swipe* para baixo, indicar que o mesmo foi carregado no transporte.

Aplicação de controlo do sistema PoolLiveAid

Uma outra aplicação desta tecnologia inclui as aplicações lúdicas como o PoolLiveAid (Alves et al., 2013). Nesta aplicação, os utilizadores podem navegar num menu interativo e escolher várias opções do sistema PoolLiveAid, tais como observar os resultados anteriores, partilhar na Internet o resultado e começar um novo jogo. Neste caso em particular, o sistema de reconhecimento de gestos com Leap Motion pode ser colocado onde for mais conveniente para o utilizador. Como conceito, o sistema foi usado sobre a mesa (Fig. 3.4 em cima) ou então projetado numa parede (Fig. 3.4 em baixo). Tal como os exemplos anteriores, para o utilizador navegar entre as opções, basta efetuar o *swipe* para a frente e para trás. Para selecionar a opção pretendida basta efetuar o gesto de cima para baixo.

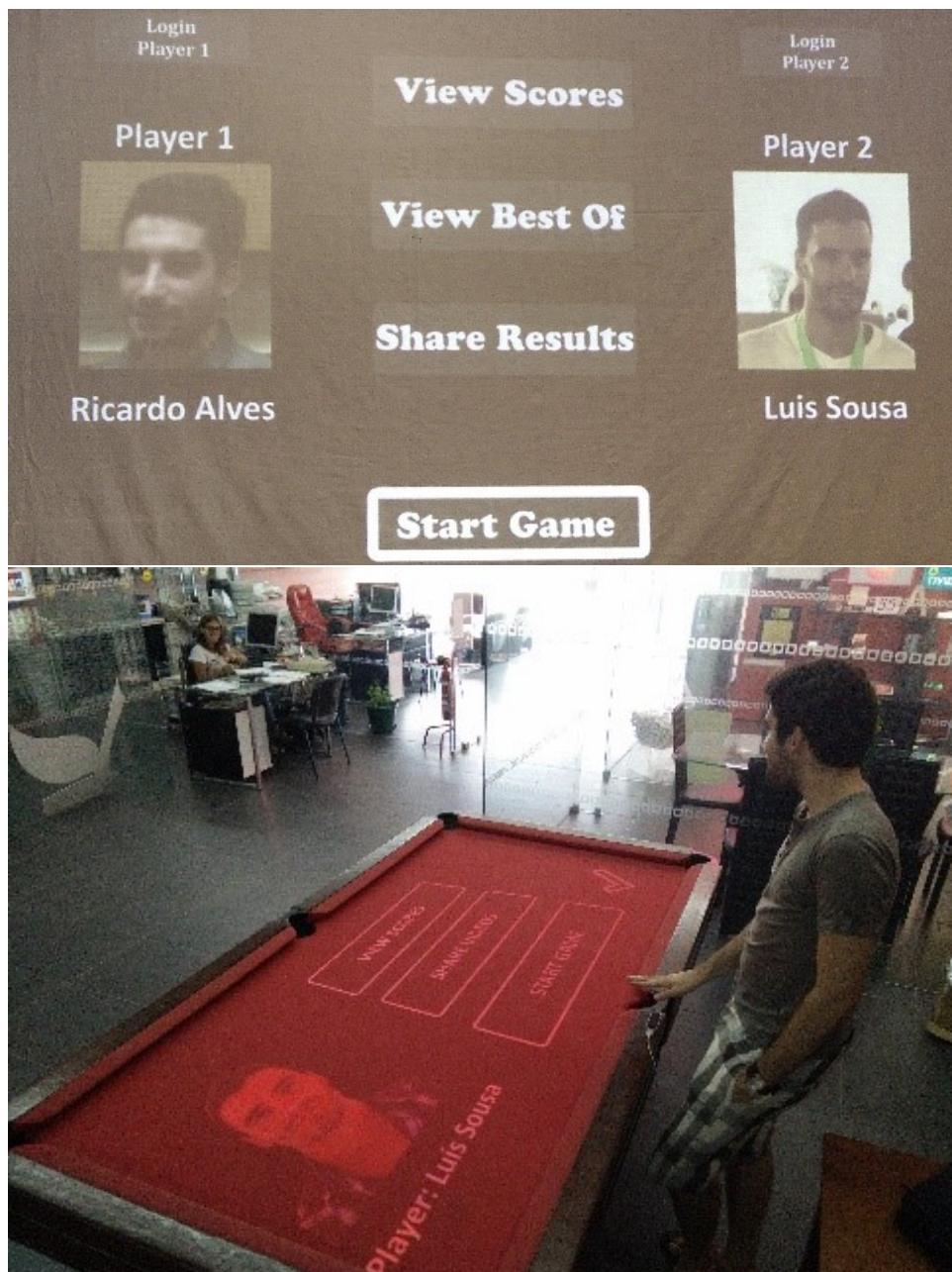


Figura 3.4: Interface usando o LeapMotion para o sistema PoolLiveAid. Utilização diretamente numa mesa de snooker à esquerda. Projeção do menu interativo numa parede à direita.

3.2 GyroGlove - Luva eletrónica com reconhecimento de gestos

Como mencionado no início deste capítulo, para além do sistema de reconhecimento de gestos com o Leap Motion, também foi desenvolvido um outro sistema, alternativo,

com o objetivo de eliminar alguns problemas oferecidos pelo Leap Motion, permitindo também o desenvolvimento de aplicações que podem ser controladas com gestos mais detalhados e específicos.

Para o desenvolvimento da luva utilizaram-se sensores de inércia que serão detalhados na seguinte secção.

3.2.1 Sensores de inércia

Um sensor *Inertial Measurement Unit* (IMU) é um dispositivo eletrónico capaz de medir vários tipos de forças de inércia. Dependendo da composição de todo o dispositivo, ele pode ser formado por vários sensores independentes tais como, giroscópio, acelerómetro, magnetómetro e menos comuns, sensor de altitude (pressão atmosférica). Os IMU utilizados na luva que iremos apresentar mais à frente contêm um giroscópio e um acelerómetro, cada um de três eixos, tornando-se assim um IMU de 6 graus de liberdade (6DoF).

O acelerómetro, tal como o nome indica, mede a aceleração (m/s^2) a que o dispositivo está sujeito em relação a um eixo especificado. Quando está parado, estático, apenas mede a força gravítica da terra ($9.8m/s^2$) no eixo descendente. O valor exato depende do local e elevação onde o dispositivo está colocado. O acelerómetro pode medir forças estáticas (gravidade) ou dinâmicas. Uma das aplicações mais usadas em telemóveis e tablets é o *tilt-sensing*. Ao medir a força da gravidade nos três eixos, consegue-se saber a orientação em relação à gravidade da terra e assim rodar o ecrã automaticamente para o benefício do utilizador. Os acelerómetros são avaliados pelas suas capacidades, entre as mais importantes encontram-se: o alcance máximo, a precisão e o número de eixos de medida. O alcance máximo é medido em forças-G. Assim, 1G indica uma aceleração de $9.8m/s^2$. Normalmente estão disponíveis as acelerações de $\pm 2G$ a $\pm 250G$ para os acelerómetros mais comuns. Quanto maior for o alcance destes valores, menor será a respetiva precisão. Os dados obtidos são normalmente compostos por números inteiros de 16 *bits* em complemento de 2 (de -32768

até +32767). Por isso, se se usar um alcance grande, p.e. $\pm 16G$, existe pouca precisão nas medições quando se pretendem medir valores baixos de aceleração. Estes dispositivos conseguem medir a aceleração entre um a três eixos, sendo que os de três eixos suportam mais funcionalidades. Embora os acelerómetros sejam bastante precisos para longos períodos de tempo, são instáveis em curtos espaços de tempo, i.e., medem valores bastante granulados em curtos espaços de tempo. Quando um acelerómetro está colocado num aparelho que se move ou treme bastante, não é possível medir com exatidão todas as acelerações. Estes componentes eletrónicos são sim, mais fiáveis, quando existem poucas variações de acelerações ao longo do tempo.

O giroscópio é um sensor capaz de medir a velocidade angular, ou seja, a velocidade com que um objeto roda num determinado eixo. É usado quando se pretende saber a orientação de um objeto em movimento. A velocidade angular é medida em graus por segundo ($^{\circ}/s$). Tal como no acelerómetro, é preciso ter em consideração algumas características dos giroscópios tais como: o alcance máximo, a precisão e o número de eixos de medida. O alcance máximo é medido em $^{\circ}/s$ e normalmente possui valores desde $\pm 30^{\circ}/s$ a $\pm 2000^{\circ}/s$. Se se escolher um valor muito baixo, são bastantes precisos mas apresentam o problemas de limitação da velocidade angular máxima (se se mover mais rápido do que é capaz, deixa de apresentar valores corretos). Por outro lado, se o valor máximo for demasiado elevado, a precisão é reduzida. Por isso é necessário ajustar o componente à aplicação pretendida. Tal como o acelerómetro, existem giroscópios capazes de medir a velocidade angular entre um a três eixos. Embora os giroscópios sejam bastante precisos na medição da velocidade angular, eles sofrem de problemas de *drift*, medem valores baixos de rotação contínua, mesmo estando imobilizados. Este problema deve-se a erros e ruído intrínsecos da fase de produção, sendo diferentes para cada módulo produzido. Ao contrário do acelerómetro, os giroscópios são muito precisos em curtos espaços de tempo, mas imprecisos em longos (devido ao *drift*), por isso complementam-se bem quando usados em conjunto. Este é o motivo de existirem componentes eletrónicos que integram tanto um acelerómetro e um

giroscópio em conjunto.

3.2.2 GyroGlove

A luva desenvolvida neste trabalho, denominada de GyroGlove (GyG) possui vários sensores *IMU* MPU-6000 (ver Fig. 3.5) (InvenSense, 2015) e um módulo controlador central que tem como função a programação e configuração dos sensores e interface entre a luva e o computador. Cada sensor MPU-6000 possui um acelerómetro e um giroscópio, ambos de 3 eixos, sendo programáveis com um alcance máximo configurável de $\pm 2G$, $\pm 4G$, $\pm 8G$, $\pm 16G$ e $\pm 250^\circ/s$, $\pm 500^\circ/s$, $\pm 1000^\circ/s$, $\pm 2000^\circ/s$ respetivamente. Como os sensores são colocados na mão e dedos do utilizador, não será expectável grandes valores de forças-G nem velocidades angulares elevadas e para manter a maior precisão possível foi decidido usar os valores de $\pm 2G$ e $\pm 500^\circ/s$, respetivamente para o acelerómetro e giroscópio.

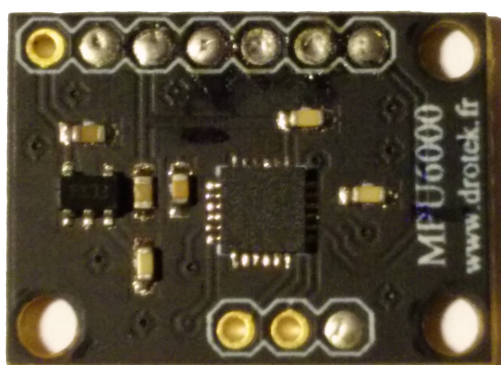


Figura 3.5: Uma PCB (Printed Circuit Board) de um sensor MPU-6000 (InvenSense, 2015)

Os sensores foram estrategicamente posicionados na luva de forma a se conseguir extrair a rotação tridimensional de cada dedo e também da mão. Para tal, os sensores foram posicionados nas zonas onde era mais importante medir, procurando minimizar o número de sensores, sem no entanto limitar a captura das rotações de toda a mão.

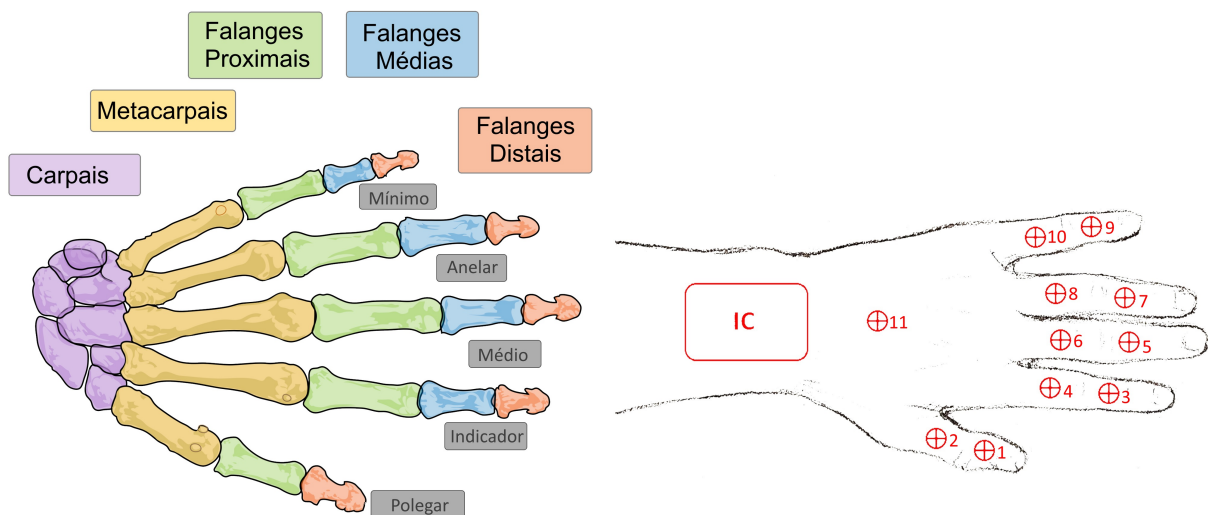


Figura 3.6: A designação anatómica dos ossos da mão à esquerda. A localização dos sensores e módulo principal à direita.

Os nomes anatómicos dos ossos da mão que se consideraram mais importantes são as falanges distais, médias e proximais (ver Fig. 3.6, esquerda). Considerando as observações anteriores, foi decidido usar onze sensores nestes locais, possibilitando a extração, sem limitações, de todos os dados rotacionais dos dedos e mão. Devido ao facto dos ossos distais serem muito curtos e pelo facto de ser fisicamente impossível (à exceção de raros casos) dobrar esta zona do dedo sem mover as adjacentes, optou-se por não colocar aí um sensor. Exceptua-se aqui o dedo polegar, pois é o único que não possui esta particularidade. Assim, no dedo polegar, para além de ter sido colocado um sensor na falange distal, também se colocou um sensor na falange proximal. Todos os outros dedos têm sensores na falange média e proximal. O 11º sensor está colocado no topo da mão, junto ao módulo principal, e serve para extrair a orientação global da mão. Este último sensor constitui a base para a correlação de todos os outros sensores, como será explicado mais para à frente. Para uma mais fácil identificação, todos os módulos estão numerados de 1 a 11. Os locais dos sensores e a numeração podem ser observados na Fig. 3.6, direita.

Estes sensores estão integrados com um sistema DMP (*Digital Motion Processor*)

(InvenSense, 2015) que processa complexos algoritmos de fusão de movimentos de 6 eixos (*6-axis Motion Fusion*). Estes algoritmos são registados pela InvenSense e a forma de funcionamento não é do conhecimento público. O cálculo destes algoritmos por parte dos módulos IMU são úteis por quatro razões, relacionadas entre si:

- (i) redução e possível eliminação dos problemas associados às imprecisões dos acelerómetros e giroscópios (pouca precisão do acelerómetro em casos de extrema vibração e o problema do *drift* no caso dos giroscópios);
- (ii) poupança de tempo em pós-processamento no módulo controlador ou no computador;
- (iii) eliminação de problemas/limitações relacionados com os ângulos Euler (explicado a seguir);
- (iv) utilização direta de ângulo na forma de quatérnios (solução para as limitações dos ângulos Euler).

Uma possibilidade para eliminar ou atenuar os problemas de imprecisões dos sensores é o uso de filtros digitais (i). O acelerómetro tem bastante precisão a longo prazo mas é impreciso a curto prazo e com o giroscópio verifica-se exatamente o oposto. Uma forma bastante fácil de atenuar este problema baseia-se no uso de um filtro complementar. Este filtro combina numa só expressão, um filtro passa baixo para o giroscópio e um filtro passa alto para o acelerómetro. Assim, sendo C o valor do complemento que pode variar de 0.0 a 1.0 e dt o tempo da amostra, pode-se calcular o ângulo final de cada eixo (An_{axis}) através de $An_{axis} = C \times (An_{axis} + Ga_{axis} \times dt) + (1 - C) \times Aa_{axis}$. Sendo Gd_{axis} o valor em graus por segundo diretamente extraído do giroscópio na sua forma original em relação a cada eixo, Ga_{axis} é o seu ângulo calculado através de $Ga_{axis} = \frac{Gd_{axis}}{G_{sens}}$. A sensibilidade do giroscópio (G_{sens}) é calculada através de $G_{sens} = \frac{2^{16}}{G_{range}} = 131.072$, sendo $G_{range} = 500 \frac{0}{s}$ o alcance máximo do giroscópio. Sendo Ad_{axis} , o valor em força-G extraído diretamente do acelerómetro, pode ser calculado o ângulo

em torno de cada eixo Aa_{axis} através de $Aa_x = \tan^{-1} \frac{Ad_y}{Ad_z} * \frac{180}{\pi}$ e $Aa_y = \tan^{-1} \frac{Ad_x}{Ad_z} * \frac{180}{\pi}$. O ângulo à volta do eixo Z (An_z) não pode ser calculado através do acelerómetro porque os cálculos do acelerómetro baseiam-se na gravidade na direcção desse eixo. Este é um fenómeno conhecido como *Gimbal Lock* (iii), um problema/limitação que ocorre quando se trabalha com ângulos Euler. Isto quer dizer que o ângulo An_z apenas pode ser calculado através dos dados do giroscópio, pelo que sofrerá do efeito de *drift* ao longo do tempo. O filtro complementar é bastante simples, rápido e oferece alguns resultados satisfatórios. Para melhores resultados também existe o filtro Kalman (Gordon et al., 2004), bastante mais complexo, dispendioso computacionalmente mas mais eficiente. Este algoritmo usa todos os valores observados ao longo de um tempo pre-determinado e estima o estado do sistema futuro baseado nos estados anteriores.

Embora o uso de filtros sejam uma boa forma de combinar os dados fornecidos pelo acelerómetro e giroscópio para obter uma melhor fiabilidade da orientação do sensor a curto e longo prazo, a implementação dos filtros, especialmente o Kalman, requer algum processamento computacional, algo que um simples micro-controlador pode ter em falta (ii). O facto de se usar ângulos Euler também induz um efeito negativo, um fenómeno, problemático, conhecido como *Gimbal Lock* (iii), a perda de um eixo, num sistema de 3 eixos. *Gimbal Lock* é um problema matemático associado aos ângulos Euler, impossível de resolver. Sem detalhar em demasia este fenómeno, devido a impossibilidades matemáticas, este é um problema real que acontece bastante em programas 3D como editores de jogos e animações. Os ângulos Euler são a forma mais fácil de obter a orientação do sensor, mas se a aplicação precisar de todos os 3 eixos, como este caso, deixa de ser a melhor opção. É devido facto que o uso do sistema DMP tem a maior importância, utilizando quatérnios, que não sofrem do mesmo problema que os ângulos Euler (iv). Os quatérnios são uma forma alternativa para representar um ângulo num espaço 3D e são descritos da seguinte forma: $Q = q_w + q_i i + q_j j + q_k k$, onde q_i, q_j e q_k são os valores de posição de um vetor diretor e q_w representa a rotação sobre esse eixo formado pelo vetor diretor. Alternativamente, o quatérnio também

pode ser expresso na sua forma matricial por 3.1.

$$Q = \begin{bmatrix} 1 - 2q_j^2 - 2q_k^2 & 2(q_i q_j - q_k q_w) & 2(q_i q_k - q_j q_w) \\ 2(q_i q_j - q_k q_w) & 1 - 2q_i^2 - 2q_k^2 & 2(q_j q_k - q_i q_w) \\ 2(q_i q_k - q_j q_w) & 2(q_j q_k - q_i q_w) & 1 - 2q_i^2 - 2q_j^2 \end{bmatrix} \quad (3.1)$$

Hardware

Todo o sistema é constituído por três partes principais: onze sensores MPU-6000, um controlador principal e o recetor conectado ao PC. Cada sensor possui as dimensões de 17 x 23 mm (Fig. 3.5). O controlador principal serve como meio intermédio, interface entre os dados dos sensores e a aplicação final no PC, sendo responsável pela programação e configuração dos sensores e envio dos dados por Bluetooth para o recetor no PC. Como pode ser observado no esquemático do circuito (Fig. 3.7, cima), ele é composto por:

- **ATMega 328p:** Micro-controlador ATMega 328p da Atmel [U6] (Atmel, 2015) com um oscilador de 8 MHz [U8]
- **FT232RL:** Conversor USB-UART [U7] que converte os dados da porta USB Micro [U10] para o micro-controlador e vice-versa.
- **MCP73831:** Carregador de baterias LiPo [U11], configurado através da resistência de 1 kOhm [R16] uma corrente máxima de 400 mA.
- **BD523G:** Circuito supervisor de tensão [U12]. Acende um LED [L8] avisando o utilizador de falta de bateria. Está configurado para acender o LED para tensões menores ou iguais a 3,6 V.
- **TPS13733:** Regulador de voltagem [U14], regula de 5 V do USB ou da bateria para a voltagem geral do circuito de 3,3 V.

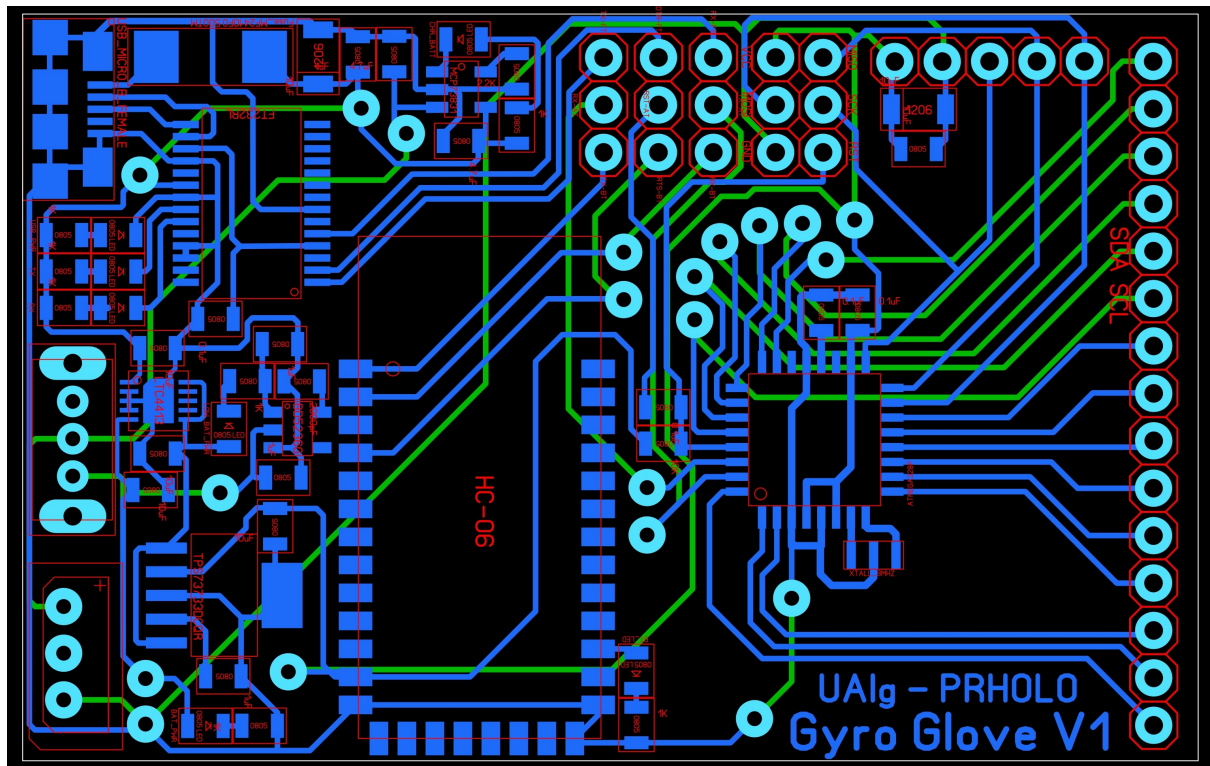
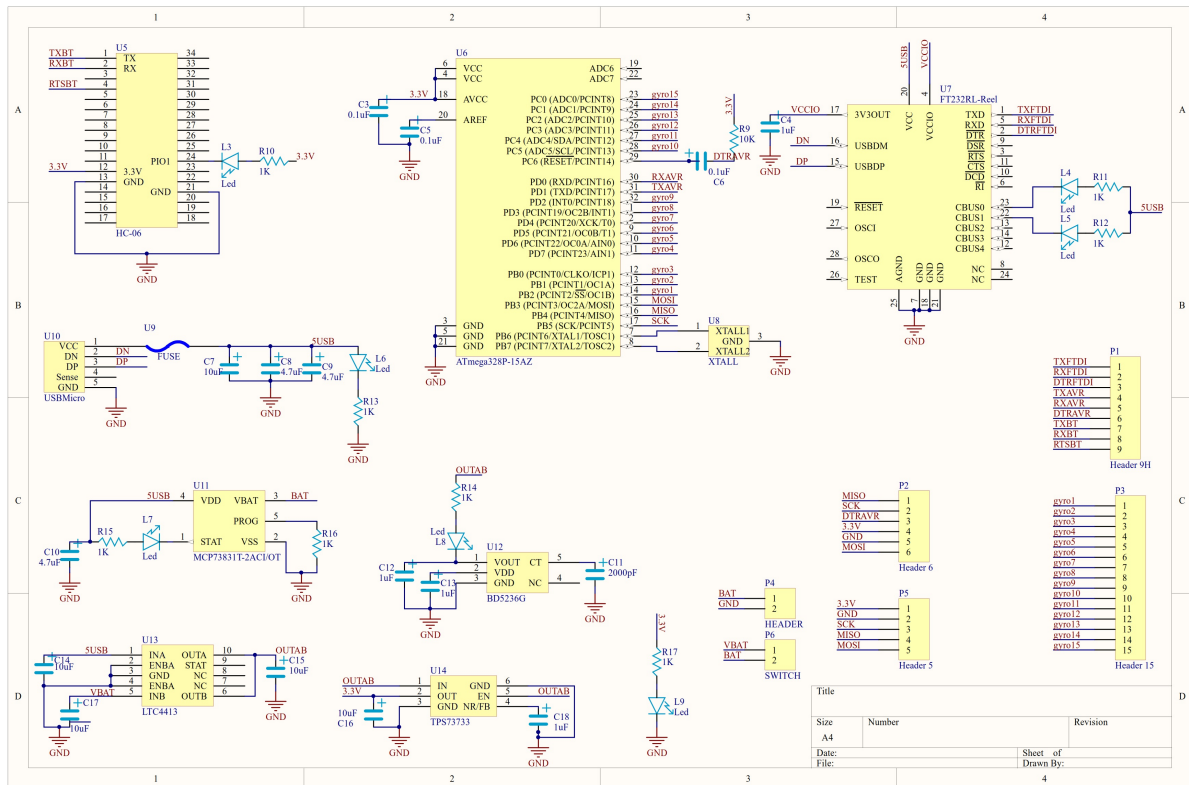


Figura 3.7: Esquemático do circuito do módulo controlador e o seu respetivo desenho da PCB.

- **LTC4413:** Distribuidor de carga [U13]. Serve para selecionar automaticamente a fonte do circuito. Quando apenas a bateria está ligada, apenas esta fonte é utilizada no circuito. Quando o USB está ligado, essa será a fonte principal, independentemente de quantas fontes estão ligadas (USB e/ou bateria).
- **HC-06:** Módulo Bluetooth [U5].

Na Fig. 3.7, baixo, pode-se observar o desenho em PCB do respetivo circuito.

O recetor (Fig. 3.8, cima) consiste num módulo Bluetooth semelhante ao utilizado no controlador, mas com a capacidade de poder ser usado como *Master Device*, i.e. tem a iniciativa de se ligar a outros módulos idênticos. Semelhante ao esquemático do circuito do controlador, este possui:

- **FT232RL:** Conversor USB-UART [U3] que converte os dados da porta USB Micro [U1] para o Bluetooth [U4]
- **HC-06:** Módulo Bluetooth [U4].

Na Fig. 3.8, em baixo, pode-se observar o desenho em PCB do respetivo circuito. Tanto a PCB do módulo controlador como do recetor foram integralmente produzidas por métodos não industriais, métodos esses que não fazem parte do objetivo da dissertação, pelo que o processo não será explicado. Na Fig. 3.9 pode-se observar o resultado final das três PCBs que o sistema é constituído, em cima. Na figura de baixo, apresenta-se o protótipo, com todo o sistema montado na luva.

Interface entre luva e aplicação

É de maior interesse que a transmissão dos dados para o PC se faça com a maior rapidez possível para evitar atrasos entre a captação das rotações dos dedos através dos sensores e a sua receção no PC (*lag*). Por esse motivo, o micro-controlador e os módulos Bluetooth foram programados com uma velocidade de transmissão de 115200bps (Fig. 3.10). Esta é a velocidade máxima possível. A limitação deve-se ao facto do oscilador (8 MHz) do micro-controlador não permitir maiores velocidades. Pela mesma

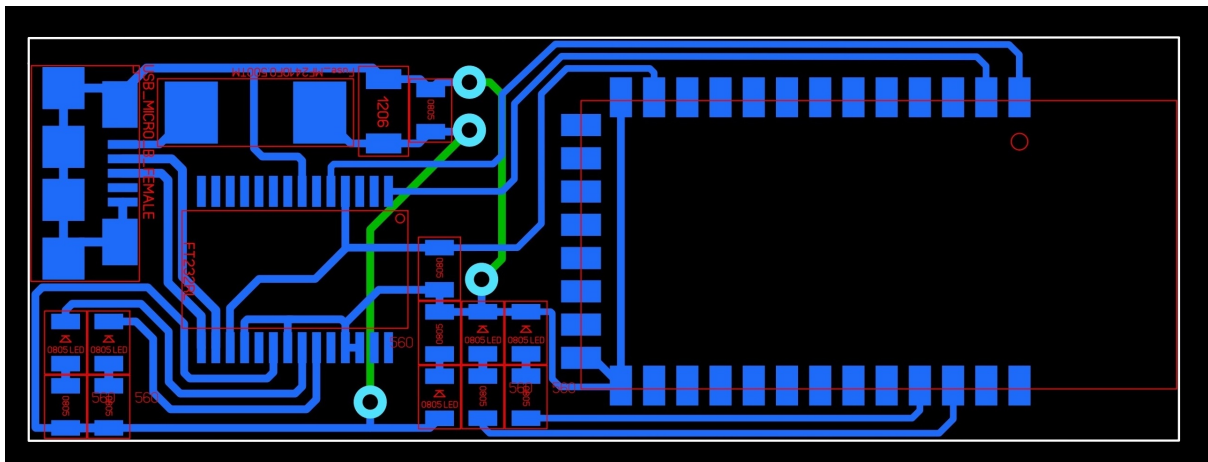
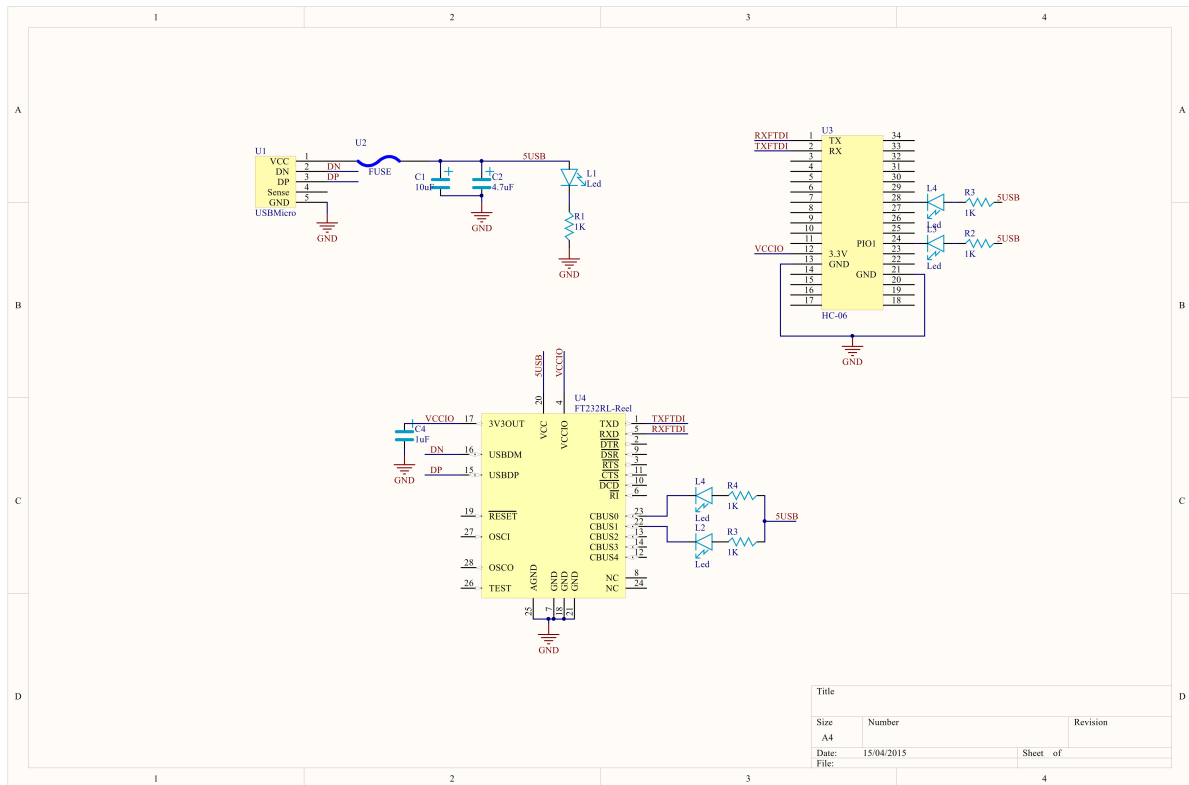


Figura 3.8: Esquemático do circuito do módulo recetor e o seu respetivo desenho da PCB.

razão, a interface entre o micro-controlador e os sensores faz-se com uma velocidade máxima da comunicação SPI de 2 MHz, sendo que a frequência máxima da interface SPI do ATmega328p é 1/4 da frequência oscilador. Estas velocidades de transmissão permitem atualizar todos os dados no PC a uma frequência de 33,3 Hz, a máxima conseguida resultante de vários testes de velocidade.

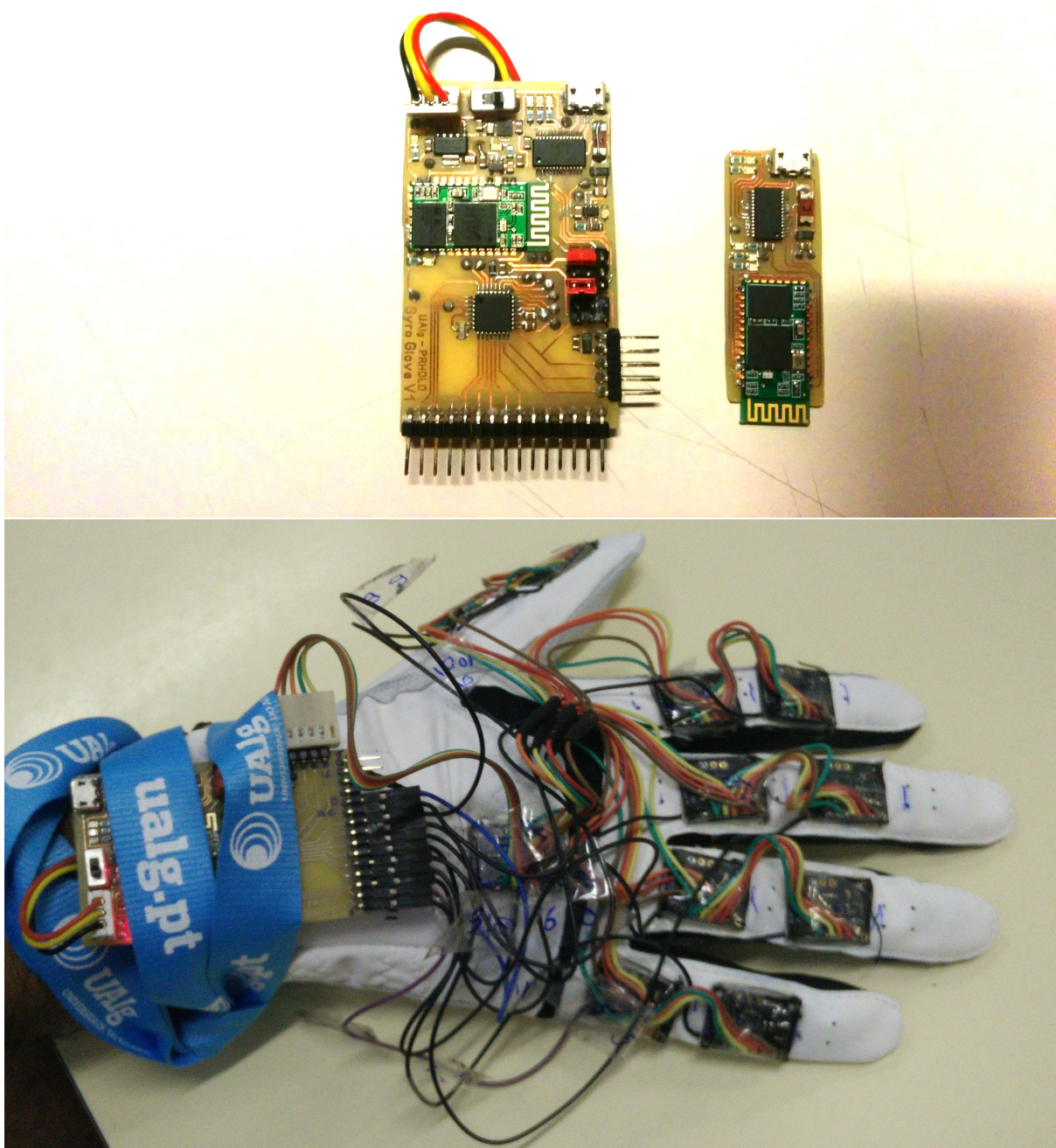


Figura 3.9: As PCBs dos protótipos dos módulos finais do sistema, em cima. O protótipo da luva em baixo.

As aplicações no PC foram feitas através do Unity 3D (Unity, 2014), um programa de criação de jogos. Este editor de jogos processa rotações com quatérnios, pelo que é o ideal para aplicar as rotações 3D diretamente extraídas dos sensores. A comunicação entre a luva e o PC é feita através de um pacote de transmissão de caracteres com o formato da Tab. 3.1. Por cada sensor é transmitido um pacote com os seus respetivos

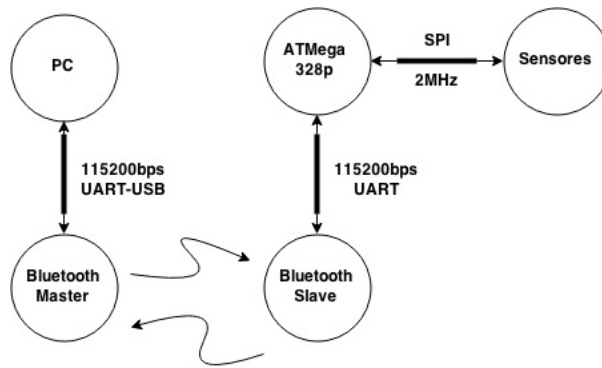


Figura 3.10: Esquema generalizado das comunicações entre os módulos principais.

dados. O início do pacote é feito pelo símbolo \$, de seguida, X, Y, Z e W são conjuntos de 16 bits com os parâmetros do quaternio (para X, Y, Z e W , a nomenclatura X_l, Y_l, Z_l, W_l e X_m, Y_m, Z_m, W_m representam respetivamente o conjunto de 8 bits menos e mais significativos) e i , o índice do sensor respetivo (ver Fig. 3.6, direita). Os caracteres CR e LF , *Carriage Return* e *Line Feed* respetivamente, indicam o final de pacote.

1 Byte	1 Byte	1 Byte	1 Byte	1 Byte	1 Byte	1 Byte	1 Byte	1 Byte	1 Byte	1 Byte	1 Byte
\$	X_m	X_l	Y_m	Y_l	Z_m	Z_l	W_m	W_l	i	CR	LF
	X		Y		Z		W				

Tabela 3.1: Pacote de transmissão dos dados entre a luva e o PC

O sistema DMP consegue calcular a rotação exata de cada sensor mas necessita de se auto configurar inicialmente, sendo este um processo no qual não existe controlo. Esta configuração tem como propósito reduzir e possivelmente eliminar o problema de *drift* provocado pelo giroscópio, como foi explicado no início deste capítulo. Logo após o sistema iniciar, todos os sensores entram em auto-configuração que pode durar entre 10 a 20 segundos, dependendo de quanto estável o sensor está. Durante este tempo, os dados extraídos não são utilizáveis uma vez que sofrem de extremo *drift*, e mesmo depois de estabilizarem, estão desfasados da orientação correta. Para corrigir o problema de desfasamento é necessário saber à priori o estado, rotação, de todos os

sensores, i.e. posicionar todos os sensores numa orientação conhecida.

$$K = Q^{-1} \times Q \quad (3.2)$$

$$K_i[t] = U_{it1}^{-1} \times S_i[t] \quad (3.3)$$

$$K_i[t] = (U_{it1}^{-1} \times D_i)^{-1} \times S_i[t] \quad (3.4)$$

Aplicando a equação 3.2, sendo Q um quaternio com uma orientação desconhecida, o resultado K é um quaternio com rotação nula, estabilizado. Considerando que na aplicação no PC, todos os objetos que replicam a orientação da mão do utilizador estão inicialmente na posição de rotação nula, basta colocar todos os sensores da mão da mesma forma ao iniciar o sistema. Sendo i o índice do sensor e $t1$ o tempo no qual a configuração se finalizou, todas as rotações dos sensores atuais (na finalização da configuração) são gravadas no quaternio U_{it1} . Aplicando a equação 3.3, sendo $S_i[t]$ os dados retirados dos sensores imediatamente e continuamente após a configuração inicial ao longo do tempo (após $t1$), $K_i[t]$ contém as rotações finais e iguais às rotações da mão do utilizador. Se a mão não for colocada com todos os sensores na posição de orientação nula, basta aplicar a equação 3.4, sendo que D_i representa o quaternio com a rotação de desvio (diferente de rotação nula) de cada sensor, ou seja, a rotação que cada sensor possui antes de se iniciar o sistema. A rotação D_i tem que ser conhecida antes da aplicação se iniciar.

É aconselhável e preferível que utilizador mantenha a mão o mais direita possível sobre uma mesa ou plano horizontal durante a configuração. Desta forma, aplica-se maioritariamente a equação 3.3 evitando desvios de rotações, exceto para os sensores posicionados no dedo polegar. Como se pode verificar facilmente pela Fig. 3.9, baixo, quando a mão está posicionada horizontalmente sobre uma mesa, todos os sensores estão direitos (sem rotação), exceto os dois do dedo polegar. Neste caso particular, é preciso estimar qual a rotação destes dois sensores e aplicar a equação 3.4. No caso

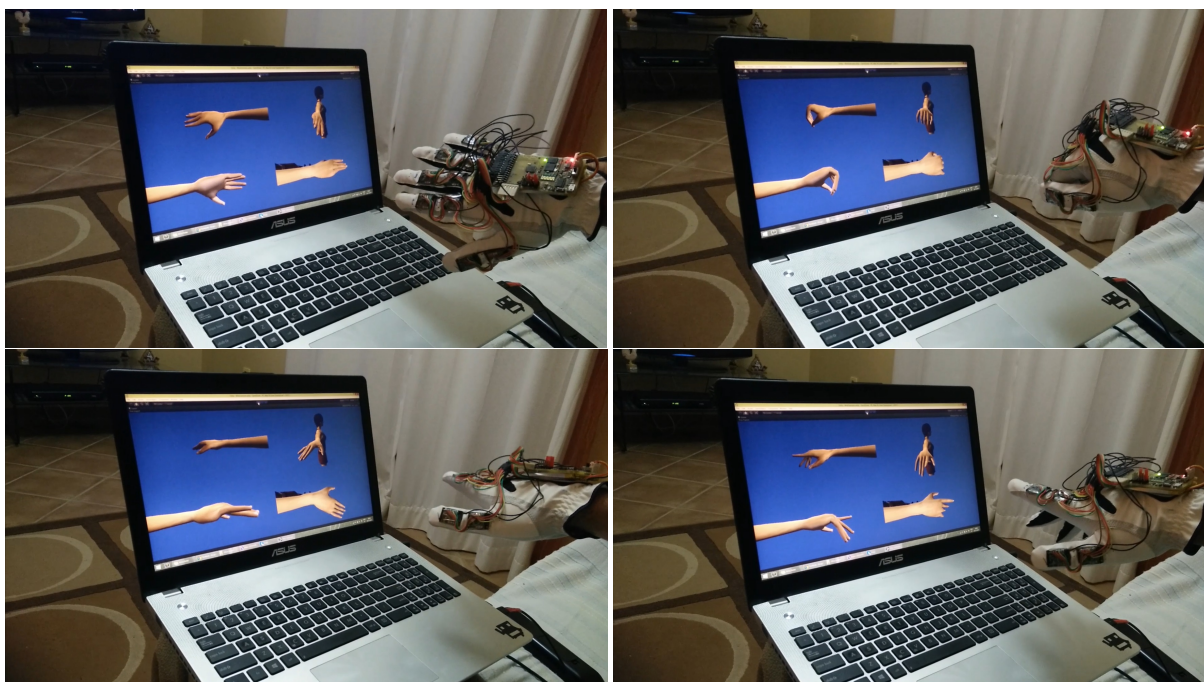


Figura 3.11: Quatro exemplos da replicação da mão do utilizador num modelo 3D. Da esquerda para a direita e de cima para baixo. Mão inclinada para a esquerda, mão fechada, mão estendida e direita e mão semi-aberta com indicador apontado.

atual, foi aplicada uma rotação inicial de -45° ao eixo x e z , obtido de forma empírica. Para verificação do funcionamento da luva, foi feita uma aplicação simples com a ferramenta Unity 3D. Esta aplicação apenas replica a mão do utilizador na mão de um modelo 3D de uma personagem virtual. Podem-se observar na Fig. 3.11, exemplos em quatro situações diferentes da posição da mão do utilizador.

3.2.3 Aplicação de aprendizagem do alfabeto da língua gestual portuguesa (LGP)

Como continuação da aplicação anterior, foi desenvolvido um pequeno jogo com o objetivo de ajudar e estimular a aprendizagem de uma língua gestual com a GyroGlove. É de conhecimento geral, que uma linguagem gestual pode ser algo difícil de aprender pela primeira vez para algumas pessoas, pelo que um jogo destinado à sua aprendizagem é uma mais valia. Como a luva apenas permite detetar as rotações de todos os dedos e mão, não é possível aprender a língua gestual de palavras, expressões ou fra-

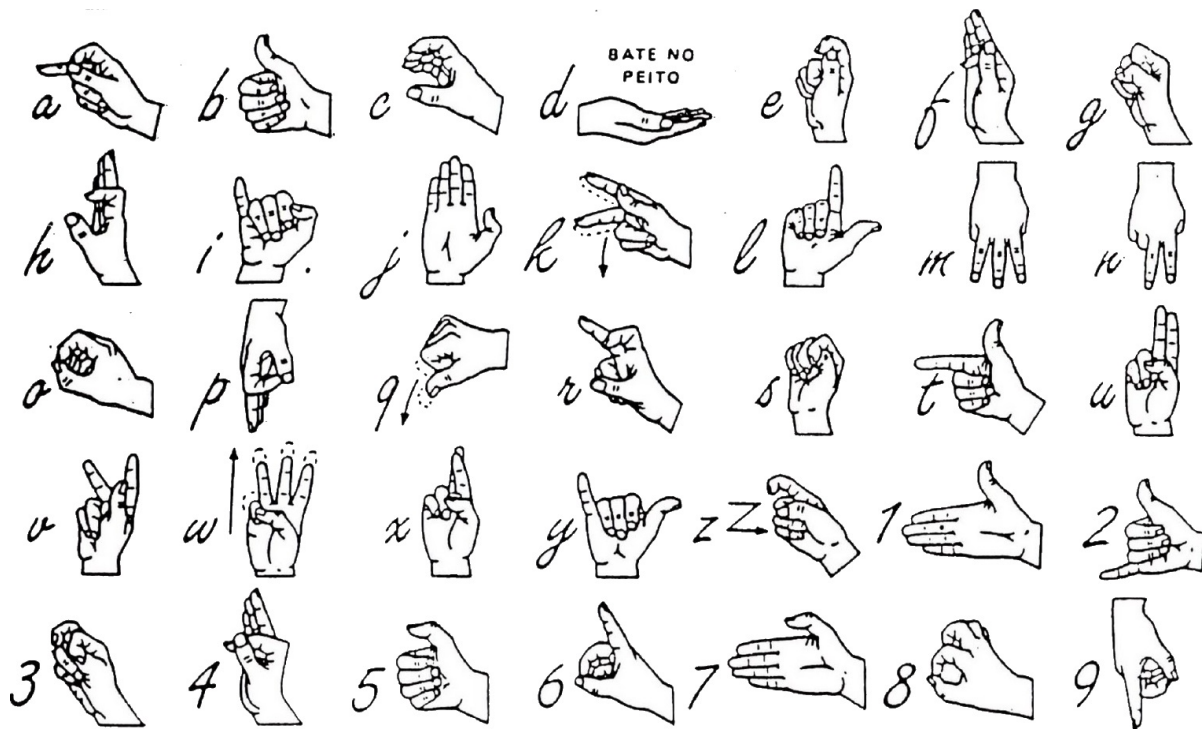


Figura 3.12: Alfabeto da Língua Gestual Portuguesa (LGP) (Wiki, 2014)

ses, mas sim apenas a língua gestual do alfabeto, onde cada letra possui uma relação direta com a posição da mão e dos dedos. Uma vez que o sistema da luva foi desenvolvido em Portugal, utilizou-se o alfabeto da Língua Gestual Portuguesa (Wiki, 2014) como a primeira língua gestual que a luva consegue reconhecer. Pode-se observar o LGP na Fig. 3.12 a língua gestual Portuguesa correspondente ao alfabeto e números. Nesta aplicação apenas será efetuado o reconhecimento do alfabeto.

Integração do LGP no sistema do GyroGlove

Para reconhecer cada posição da mão é necessário comparar todas as rotações dos sensores da mão com todas as rotações associadas a uma determinada letra do alfabeto pré-gravadas na aplicação. Em vez de se compararem os quaternios de cada sensor, compara-se a relação dos quaternios entre cada dois sensores adjacentes, pré-definidos i.e., o ângulo absoluto entre cada dois sensores. Compara-se também a relação do quaternio do sensor da mão (com número 11 na luva) com o eixo vertical, permitindo por essa via determinar a orientação global da mão (virada para cima, baixo etc)

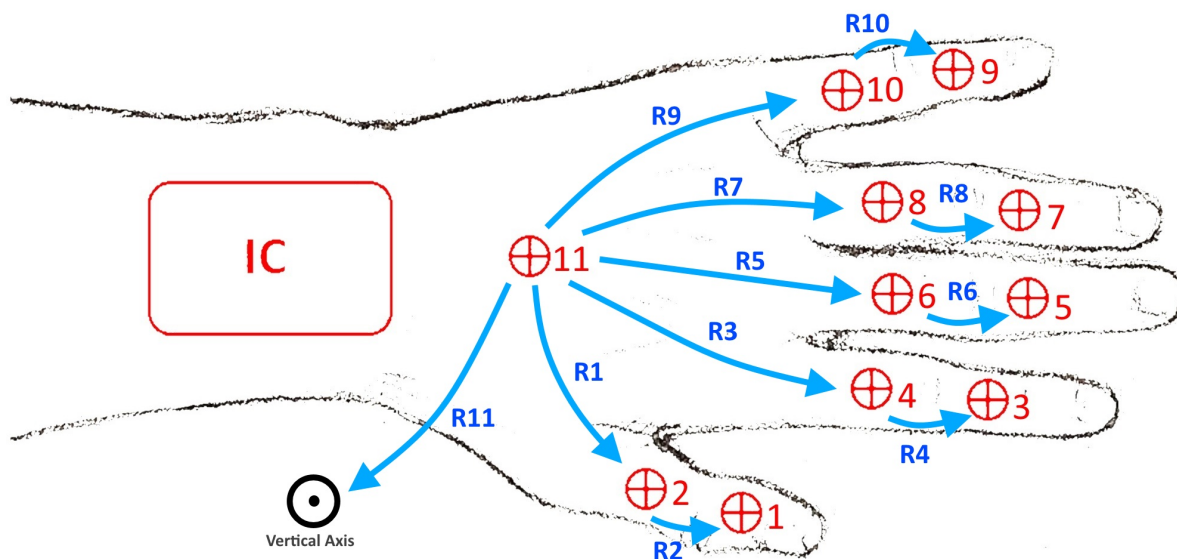


Figura 3.13: Relações entre sensores adjacentes e entre o eixo vertical

	ID: 1 Polegar Distal	ID: 2 Polegar Médio	ID: 3 Indicador Médio	ID: 4 Indicador Proximal	ID: 5 Médio Médio	ID: 6 Médio Proximal	ID: 7 Anelar Médio	ID: 8 Anelar Proximal	ID: 9 Mínimo Médio	ID: 10 Mínimo proximal	ID: 11 Hand
Relação com Sensor ID n ^o i (R _j):	2 (R ₂)	1 (R ₂) 11 (R ₁)	4 (R ₄)	3 (R ₄) 11 (R ₃)	6 (R ₆)	5 (R ₆) 11 (R ₅)	8 (R ₈)	7 (R ₈) 11 (R ₇)	10 (R ₁₀)	9 (R ₁₀) 11 (R ₉)	2 (R ₁) 4 (R ₃) 6 (R ₅) 8 (R ₇) 10 (R ₉) Vetor Vertical (R ₁₁)

Tabela 3.2: Relações entre sensores / partes da mão para comparação e detecção de gestos.

No total existem 11 relações distintas (R_j), tal como representado na Fig. 3.13 e Tab. 3.2. Cada seta azul na imagem tenta apresentar essas relações. Para determinar cada uma, é necessário calcular o ângulo absoluto entre cada conjunto de dois sensores. Como exemplo, a primeira relação (R_1) faz-se entre o sensor 11 na posição global da mão e o sensor 2 posicionado no polegar proximal.

Como um quaternio (os dados retirados diretamente dos sensores) define uma rotação e não um vetor no espaço, é necessário aplicar a rotação de cada quaternio a um vetor conhecido V_{front} , dado por exemplo em Eq.3.5. Aplicando a Eq.3.6, resulta \vec{V}_i , um

vetor com a orientação de K_i (quaternios obtidos dos sensores).

$$V_{front}^{\vec{}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (3.5)$$

$$\vec{V}_i = K_i \times V_{front}^{\vec{}} \quad (3.6)$$

Depois de obtido o vetor \vec{V}_i que define a orientação no espaço de cada um dos 11 sensores, é necessário calcular o ângulo absoluto, e mais curto, entre cada dois vetores de orientação. Este ângulo traduz a relação entre cada par de sensores. Aplicando a Eq.3.8, para cada conjunto de dois sensores (ilustrado em Tab. 3.2 e Fig. 3.13), obtém-se A_j , o ângulo em graus ($^\circ$) formado entre cada conjunto de dois sensores n e m (a variar entre 1 e 11) e também entre o sensor principal da mão (sensor 11) e o vetor vertical \vec{V}_{up} representado na Eq.3.7. Os ângulos absolutos A_j podem variar entre 0° a 180° .

$$\vec{V}_{up} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad (3.7)$$

$$A_j = \left| \arccos \frac{\vec{V}_n \cdot \vec{V}_m}{\|\vec{V}_n\| \|\vec{V}_m\|} \times 180 / \pi \right|, \quad j = \{R1, \dots, R11\} \quad (3.8)$$

Para a detecção de gestos, através das posições da mão, é necessário que previamente se gravem estas posições, calculando os ângulos absolutos entre cada sensor (A_j) através da Eq.3.8, entre os pares de sensores da luva estabelecidos através da Tab. 3.2. Para isso, para cada uma das letras $l = a, \dots, z$ do alfabeto, procedeu-se ao cálculo dos ângulos máximos ($Smax_{j,l} = \max A_j[w]$ com $t - \Delta_{sample} \leq w \leq 0$) e mínimos ($Smin_{j,l} = \min A_j[w]$ com $t - \Delta_{sample} \leq w \leq 0$) de cada ângulo associado a cada letra. A mão foi posicionada na forma da respetiva letra durante o tempo Δ_{sample} .

Sempre que estão disponíveis novos dados da mão, K_i , são calculados todos os ângulos atuais A_j e verificado se estes dados estão dentro do intervalo entre em $Smin_{j,l}$ e $Smax_{j,l}$, para cada letra e durante um espaço de tempo $\Delta t_s = 1s$, definido empiricamente. Analisando todos os dados no intervalo de tempo, se a condição 3.9 se verificar em todo o espaço de tempo, será considerado uma deteção positiva do gesto, indicando uma posição da mão correta. Na aplicação prática, será considerado que o utilizador colocou a mão corretamente de acordo com uma letra do alfabeto. Na expressão 3.9, Δm é um valor de margem com o objetivo de permitir ajustar a sensibilidade da deteção. Nos testes utilizou-se um valor de Δm igual a 20, podendo o mesmo ser ajustado.

$$Smin_{j,l} - \Delta m \leq A_j \leq Smax_{j,l} + \Delta m \quad (3.9)$$

Aplicação

Na Fig. 3.14 pode-se observar a aplicação desenvolvida para a estimulação da aprendizagem do alfabeto da linguagem gestual. A interface foi desenvolvida através da ferramenta Unity 3D e é constituída por 4 partes: (i) a letra e a imagem da posição da mão respetiva no lado direito, (ii) um *input* visual com a posição da mão replicada num modelo 3D humano do lado esquerdo em cima, (iii) um sistema de pontuação do lado esquerdo no meio e (iv) uma barra de nível a indicar o quanto próximo o utilizador tem a posição da mão próxima do objetivo no lado esquerdo, em baixo. No início da aplicação, o utilizador é convidado a colocar a mão sobre a mesa com a palma virada para baixo e esticada de forma a que o programa configure os sensores. Depois de alguns segundos, a configuração finaliza-se e aparece no ecrã, o design em forma de jogo da Fig. 3.14.

O programa apresenta as letras do alfabeto no ecrã individualmente, começando pela primeira letra, e uma imagem explicativa da posição da mão e dedos que o utilizador tem que replicar (i). Quanto mais próximo da posição certa da mão e dedos, mais preenchida a barra de nível estará (iv), refletindo visualmente o quanto próximo da

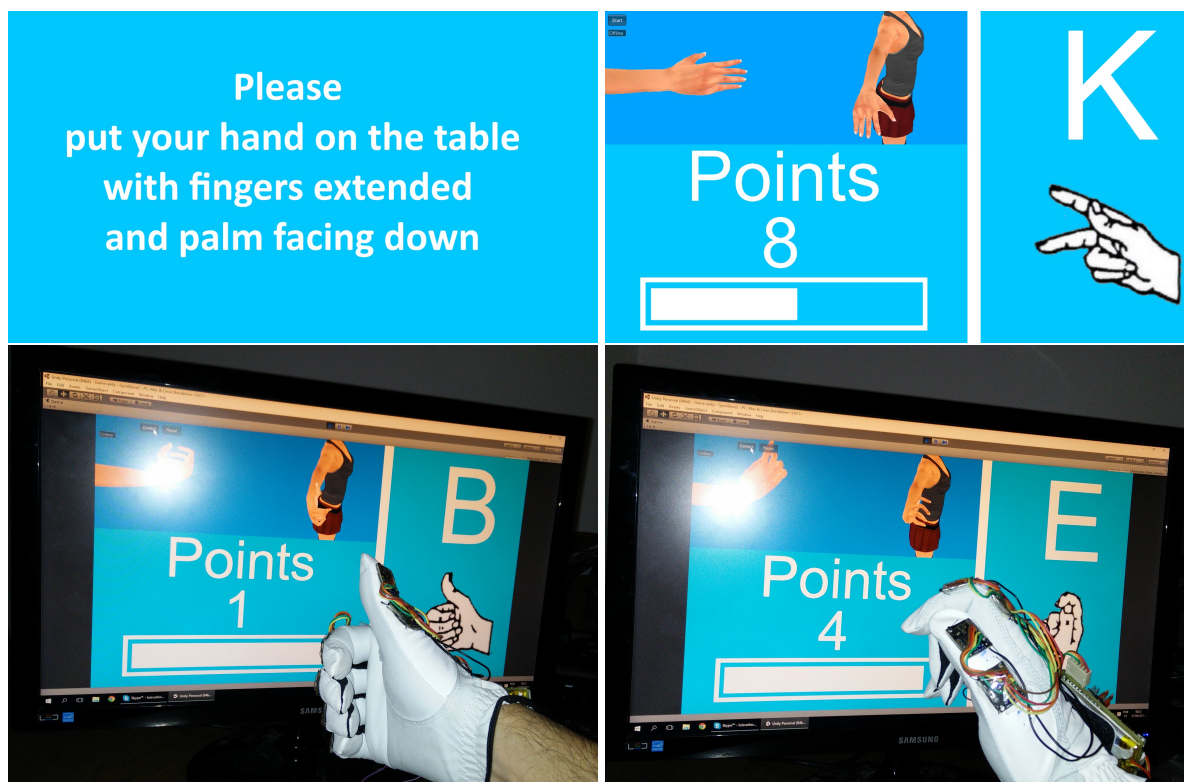


Figura 3.14: Aplicação de estimulação da aprendizagem do alfabeto da linguagem gestual portuguesa. Instruções da aplicação na imagem em cima, à esquerda. O design da aplicação na imagem de cima, à direita, e em baixo, duas imagens durante a utilização.

posição certa o utilizador tem a mão. Cada vez que acertar na posição, os pontos são incrementados de 1 valor (iii) e o programa passa à letra seguinte. O utilizador é assim encorajado a tentar imitar as letras do alfabeto da língua gestual portuguesa do e desse modo aprender mais rapidamente.

3.3 Testes e resultados

Nesta secção são apresentados e discutidos os resultados obtidos através de um conjunto de testes efetuados aos sistemas desenvolvidos neste capítulo. Nas próximas duas secções serão analisados os testes efetuados ao sistema de reconhecimento de gestos usando o sensor Leap Motion e à luva eletrónica, respetivamente.

3.3.1 Reconhecimento de gestos com Leap Motion

Para poder qualificar este sistema, colocaram-se 7 pessoas a experimentar as 3 aplicações desenvolvidas na secção 3.1.1 durante alguns minutos e registaram-se vários parâmetros tais como: i) o tempo total de utilização (com definição mm:ss), ii) total de gestos efetuados, iii) total de gestos detetados corretamente pelo sistema, iv) total de gestos não detetados e v) total de falsos positivos. Os resultados dos parâmetros anteriores estão apresentados nas colunas da Tabela 3.3.

De salientar que como todas as aplicações desenvolvidas usam o mesmo sistema de reconhecimento de gestos, não é feita a distinção dos registos entre elas mas sim do conjunto das 3 aplicações.

		Testes				
		i	ii	iii	iv	v
Identificação da pessoa de teste	1	06:23	95	85(89%)	10 (11%)	5
	2	08:00	102	86 (84%)	16 (16%)	7
	3	05:10	79	67 (85%)	12 (15%)	6
	4	10:50	117	101 (86%)	16 (14%)	10
	5	04:32	60	56 (93%)	4 (7%)	6
	6	06:35	83	72 (87%)	11 (13%)	5
	7	04:52	55	49 (89%)	6 (11%)	2
Total		46:22	591	516 (87%)	75 (13%)	41
Média		6:37	84	74 (87%)	11 (13%)	6

Tabela 3.3: Tabela com dados quantitativos sobre os resultados de várias pessoas que experimentaram as aplicações.

As 7 pessoas testaram as 3 aplicações num total de 46 minutos e 22 segundos, realizando 591 gestos (uma média de 84 cada pessoa). O sistema reconheceu os gestos com um sucesso de 87% (516) sendo que 13% deles (75) não foram detetados. O motivo para a não deteção de 13% de todos os gestos efetuados, deve-se principalmente a 3 motivos: a) o gesto não atingiu a velocidade mínima, b) o gesto não percorreu a distância mínima, ou então o utilizador não possuía a mão na área de visão (campo de ação) do Leap Motion.

Verificaram-se 41 falsos positivos que se deveram a situações indesejáveis. Por exem-

plo, uma situação onde o utilizador ao retirar a mão da área de visão do Leap Motion (quando já não queria interagir com a aplicação), cumpriu os requisitos de velocidade mínima e distância mínima e de o vetor de movimento, que correspondiam a um gesto.

Este sistema de reconhecimento de gestos apresenta uma taxa de sucesso alta, podendo ser ainda melhorada com algumas alterações. Verificou-se que, apenas uma vez no início da aplicação, seria aconselhável configurar ou adaptar os requisitos das deteções de gestos para casa pessoa. Tal deve-se ao facto de diferentes pessoas moverem a mão a velocidades moderadamente constantes, mas diferentes entre si. O sistema apresentado neste capítulo está apenas configurado para uma velocidade, fazendo com que os utilizadores tenham que adaptar-se ao sistema, ao invés do sistema adaptar-se ao utilizador.

3.3.2 Reconhecimento de gestos com o GyroGlove

Para este sistema, pediu-se a 3 pessoas para testarem a aplicação para estimulação da aprendizagem do alfabeto da linguagem gestual (secção 3.2.3) durante alguns minutos e quantificou-se, para cada letra do alfabeto, i) a taxa sucesso ou percentagem de letras reconhecidas com sucesso (%), ii) a taxa de insucesso (%) e iii) o tempo médio que a pessoa tentou replicar a letra até que o sistema a reconheceu corretamente (segundos). As 3 pessoas que testaram o sistema estão identificadas por T1, T2 e T3 sendo que cada uma testou todas as letras do alfabeto 5, 8 e 10 vezes respetivamente. Todos estes dados podem ser observados na tabela 3.4. Entre todas as pessoas que testaram o sistema e para todas as letras, foi registado uma taxa de sucesso média de 76%. Embora este valor não pareça elevado, em 35% das letras foi registado uma taxa de sucesso superior a 90%, sendo que apenas 19% têm uma taxa de sucesso inferior a 70%. Os testes indicaram que algumas letras são particularmente difíceis de serem detetadas pelo sistema. Esta é uma imprecisão do sistema que decorre da estimação que se fez à rotação inicial dos dois sensores colocados no dedo polegar (referido no

		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Média
T1	i	100	100	100	100	80	80	100	80	80	100	60	80	100	100	80	60	40	20	100	80	60	60	60	20	80	60	76
	ii	0	0	0	0	20	20	0	20	20	0	40	20	0	0	20	40	60	80	0	20	40	40	40	80	20	40	24
	iii	3	2	3	3	4	5	3	4	3	2	5	3	2	3	3	6	6	7	2	3	5	6	5	8	3	5	3,8
T2	i	100	100	88	75	88	75	100	88	100	100	38	88	100	100	100	38	38	13	100	88	38	38	88	13	100	75	77
	ii	0	0	12	25	12	25	0	12	0	0	62	12	0	0	0	62	62	87	0	12	62	62	12	87	0	25	23
	iii	2	2	2	3	3	4	2	3	3	2	5	4	3	2	4	6	6	8	2	3	4	4	5	6	3	4	3,4
T3	i	90	100	90	90	90	80	90	70	70	80	70	60	90	100	90	60	50	30	90	90	40	60	70	10	80	70	75
	ii	10	0	10	10	10	20	10	30	30	20	30	40	10	0	10	40	50	30	10	10	60	40	30	90	20	30	25
	iii	2	2	3	3	2	4	3	3	3	3	4	4	2	2	3	5	6	6	2	2	4	5	7	6	7	4	3,6
Média	i	97	100	93	88	86	78	97	79	83	93	56	76	97	100	90	53	46	21	97	86	48	46	73	56	87	68	
	ii	3	0	7	12	14	22	3	21	17	7	44	24	3	0	10	47	54	79	3	14	52	54	27	44	13	32	
	iii	2	2	3	3	3	4	3	3	3	2	5	4	2	2	3	6		7	2	3	4	5	6	7	4	4	

Tabela 3.4: Tabela com os resultados quantitativos dos testes realizados com o sistema de reconhecimento de gestos com a luva GyroGlove.

final da secção 3.2.2). A maioria da taxa de insucessos deveu-se a este facto.

4

Interface humano-computador aplicada a uma instalação holográfica 360°

Neste capítulo será apresentado o desenvolvimento de uma instalação holográfica 360° projetada e desenvolvida no âmbito do projeto "PRHOLO: O Relações Públicas Holográfico Realista" tendo como promotor a empresa SPIC - Creative Solutions. Esta instalação foi desenhada com o principal objetivo de mostrar em todo o seu redor, um holograma que aparente flutuar no ar. Como objetivos, foi definido que o holograma possuirá perspectivas nos 360° em torno da estrutura, permitindo assim, que os utilizadores possam observar o holograma de qualquer ângulo. Deverá também ser transparente para oferecer o maior envolvimento possível e ser de fácil transporte e

montagem. Para além disso, também se tomou como objetivo, a extração da informação sobre os utilizadores que se movimentam à volta da instalação, para efeitos de estatística. A instalação deverá também permitir a interação em forma de gestos, quer com a imagem holográfica, como também com uma interface em forma de menus contendo informações.

No decorrer deste capítulo começamos por apresentar, na secção 4.1, a tecnologia de holografia em que esta instalação se baseou para gerar os hologramas. Na secção 4.2 será detalhado a construção da instalação e elementos de hardware. O software que corre na máquina local, que gera o holograma e a interface de menus interativos e a aplicação Android capaz de interagir com o holograma está apresentado na secção 4.3. Por fim, na secção 4.4 serão apresentados os resultados.

4.1 Técnica *Pepper's Ghost*

O princípio sob o qual assenta o holograma baseia-se na técnica *Pepper's Ghost*, uma técnica que ficou famosa quando John Henry Pepper, um mágico, que a usou em 1862 em palco, num truque de magia. Esta técnica é na verdade um truque de ilusão usando uma imagem do objeto que se deseja transformar em holograma e uma película fina transparente, ou vidro. As grandes vantagens do uso desta técnica são o facto de poder ser económica de implementar e de não possuir partes móveis. Como se pode observar na Fig. 4.1, colocando a imagem do holograma no plano horizontal, faz com que a mesma seja refletida pela película inclinada a 45° . Assim, o utilizador que olhar diretamente para a película verá a imagem flutuando no ar atrás da película. A película apenas reflete a luz da imagem. A distância a que esta se encontra da imagem, determina o quanto afastado (atrás da película) o holograma se encontra, sendo este o principal efeito de ilusão, que o distingue do uso de um espelho por exemplo. Os 45° de inclinação são essenciais para que o holograma se situe direito, i.e., se a inclinação for menor, o holograma ficará inclinado para a frente (para o utilizador), caso

contrário, ficará inclinado para trás, causando distorção no holograma. Embora a técnica apenas apresente uma imagem em 2 dimensões, quando em forma de holograma existe a ilusão de profundidade e de existência de mais uma dimensão.

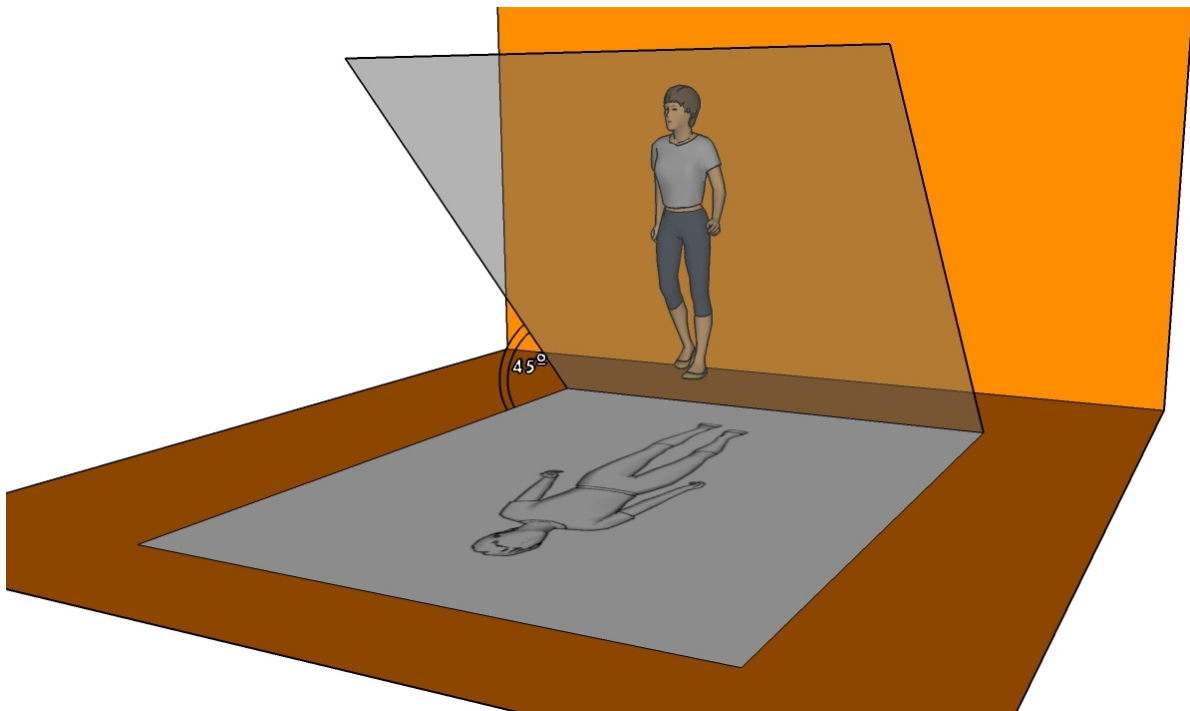


Figura 4.1: Técnica *Pepper's Ghost*. A imagem do objeto está colocada num plano horizontal enquanto uma película refletora transparente fica inclinada 45° sobre ela.

A fonte da imagem pode ser uma superfície de retro-projeção (com um projetor a projetar a imagem por trás), uma superfície branca de projeção frontal ou simplesmente um monitor. Quanto mais brilhante for a imagem e/ou usando uma superfície refletora com maior qualidade, melhores serão os resultados que se conseguem obter. A película refletora pode ser uma folha simples de acrílico transparente, poliéster, ou folha de *mylar*, sendo esta última a que possui melhores resultados por ser bastante transparente, fina e resistente, mas também a mais dispendiosa.

Na estrutura do PRHolo foi usada uma superfície acrílica de retro-projeção para produzir a fonte da imagem e em cima, uma folha de acrílico com 2 milímetros de espessura como película refletora e um projetor como fonte das imagens.

4.2 Instalação

Embora a versão original da técnica *Pepper's Ghost* acrescenta a uma imagem de 2 dimensões, a ilusão de possuir 3 dimensões, ela continua a apresentar uma perspectiva da imagem e apenas pode ser observada pelo utilizador de uma zona restrita, na parte frontal. Para criar um holograma 360° é necessário replicar várias técnicas *Pepper's Ghost* dispondo-as numa forma circular (Fig. 4.2, imagens azuis/roxas), com a reflexão do holograma direcionado para o exterior. Como referido na secção anterior e ilustrado na Fig. 4.1, a fonte da imagem foi colocada na parte inferior. Em alternativa poder-se-ia rodar a instalação de 180° , ficando a fonte da imagem na superior. Tal é possível, mas caso se use um projetor, como é o nosso objetivo, requeria uma estrutura na parte superior bastante alta. Por esta razão, definiu-se que a fonte da imagem ficará na parte inferior da estrutura, paralelamente ao chão.

Por cada técnica *Pepper's Ghost*, é acrescentada à instalação mais uma perspectiva. Quanto mais perspectivas a estrutura possuir, maior será a continuidade do holograma entre perspectivas. No entanto, a área útil do holograma reduz-se-á se se quiser manter o mesmo tamanho da estrutura. Para além disso, se a estrutura possuir mais de 4 perspectivas (menos de 90° entre cada perspectiva) dispostas de uma forma circular uniforme, começa a sofrer de interferências entre perspectivas, um problema que ocorre quando diferentes perspectivas do holograma começam a sobrepor-se na mesma área do holograma. Este problema aumenta com o aumento das perspectivas, mas consegue-se controlar, e até eliminar através de algumas técnicas explicadas em detalhe nas próximas secções deste capítulo.

Antes de definir o número de perspectivas e tamanho da estrutura, vários testes de hologramas que variaram de 1 até 10 perspectivas foram efetuados de forma a se conseguir qualificar a continuidade do holograma em todo o seu redor. Nas imagens da Fig. 4.2 (duas figuras de cima) pode-se observar dois modelos 3D com a configuração de 4 e de 8 faces, à esquerda e direita respetivamente. Para a estrutura 360° foi deci-

dido usar 8 perspectivas no holograma pelo facto de ser um valor, segundo os testes efetuados, que permite uma continuidade superior entre as perspectivas, uma evolução da comum pirâmide com 4 faces, oferecendo assim uma maior imersão no uso da tecnologia.

Um holograma com mais de 2 perspectivas que possua dimensões em altura semelhantes a um humano ($\pm 1,8$ metros) requer uma instalação com mais do dobro desse valor em largura. Como um dos objetivos é a fácil montagem e o transporte, a representação holográfica de uma figura humana torna-se assim proibitiva. Desse modo, foi definida uma estrutura (área útil do holograma) com uma largura máxima (diâmetro de face a face) com $d_1 = 100$ centímetros.

A combinação de várias técnicas de *Pepper's Ghost* para formar um único holograma 360° requer que as faces sejam dispostas numa configuração circular, produzindo assim uma estrutura com forma octogonal (para 8 faces), como se pode observar na Fig. 4.2 (cima à direita). Com as faces dispostas circularmente, inclinadas de 45° e tendo em conta a dimensão d_1 correspondente à largura máxima, a altura máxima possível seria $h_o = \frac{d_1}{2} = 50$ centímetros com as faces em forma de triângulo. As faces não podem possuir exatamente esta forma geométrica, uma vez que se assim for, a estrutura não se conseguirá suportar por si só. Para se construir a estrutura desta forma, teria que existir algum suporte na parte exterior, diminuindo assim a visibilidade do holograma e conseqüentemente a sua qualidade. Pelos motivos mencionados, torna-se necessário criar uma base que não influencie diretamente a qualidade do holograma. A solução passou por cortar a parte inferior do triângulo de cada face, criando assim uma base onde toda a estrutura se pode apoiar sem suportes extras (Fig. 4.2, meio à esquerda). O corte desta parte da base fez diminuir a altura do holograma, e por isso, foi definido assim uma nova altura com $h_o = 30$ centímetros, mantendo o mesmo diâmetro superior d_1 . Embora a altura da estrutura holográfica seja um pouco inferior, em termos de altura ao definido anteriormente, a área que foi cortada perfaz uma pequena área, deixando assim ainda uma área bastante grande

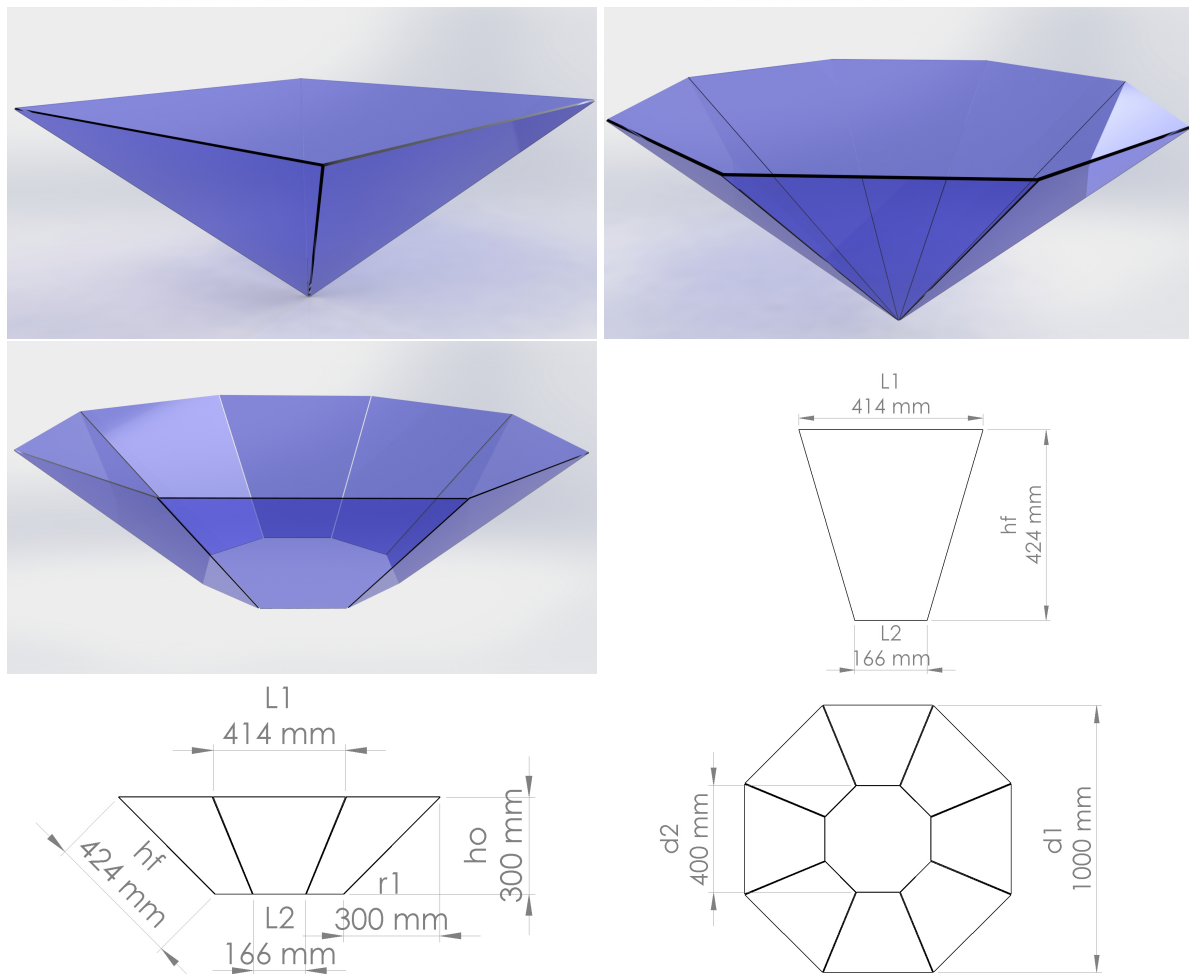


Figura 4.2: Modelos 3D de exemplos e soluções da estrutura holográfica 360° replicando a técnica *Pepper's Ghost* de uma forma circular. Nas duas imagens de cima pode-se observar a configuração de 4 e 8 faces, à esquerda e direita respetivamente. Estes são dois exemplos onde as faces estão desenhadas em forma de triângulo. Na imagem do meio à esquerda, a solução final para a estrutura holográfica aplicando a técnica *Pepper's Ghost* com 8 faces. Nas restantes figuras, as dimensões da da estrutura. A cor azul/roxo apenas tem como objetivo a melhor visualização das placas acrílicas. Na verdade, elas são transparentes.

para refletir a luz, a imagem, com a técnica *Pepper's Ghost*.

Como as faces estão inclinadas a 45° , o diâmetro inferior da estrutura octogonal possui uma dimensão de $d_2 = 40$ cm calculado através de Eq. 4.1. Cada face possui uma aresta na parte superior com um comprimento $l_1 = 41,4$ cm (Eq. 4.2), a aresta inferior $l_2 = 16,6$ cm (Eq. 4.3) e uma altura (não a altura da estrutura, mas sim da face) $h_f = 42,4$ cm (Eq. 4.4), sendo $r_1 = 30$ cm, o comprimento da projeção de cada face

no plano horizontal (Eq.4.5). Todas estas dimensões estão ilustradas na Fig. 4.2 (meio à direita e nas duas em baixo). Todos estes cálculos foram feitos para uma estrutura holográfica de 8 faces/perspetivas, podendo ser recalculadas para uma estrutura de maior ou menores dimensões.

Estas dimensões são suficientemente grandes para permitir a visualização de um holograma com um tamanho razoável que pode ser visto facilmente em toda a volta da instalação, até uma distância de 5 metros.

$$d_2 = d_1 - 2 \times \frac{h_0}{\tan 45^\circ} \quad (4.1)$$

$$l_1 = \frac{d_1}{1 + \sqrt{2}} \quad (4.2)$$

$$l_2 = \frac{d_2}{1 + \sqrt{2}} \quad (4.3)$$

$$h_f = \sqrt{h_0^2 + r_1^2} \quad (4.4)$$

$$r_1 = \frac{d_1 - d_2}{2} \quad (4.5)$$

Como se pode observar nas figuras em 4.2, o modelo 3D apenas está construído com placas acrílicas. Para a estrutura suportar o seu próprio peso, foi necessário construir algumas peças para formar um esqueleto (Fig. 4.3), reforçando toda a estrutura. Nos vértices entre cada face, na zona superior e inferior, foi colocado uma junção (peças azuis) e entre elas, nas arestas, um perfil em alumínio em forma de U (peça vermelha). Entre cada placa, na zona lateral, também foram colocados dois espaçadores (peça amarela) e um perfil de alumínio em forma de L (peça verde). As peças azuis e amarela foram desenhadas e impressas em 3D. Nas quatro figuras de baixo pode-se observar o esqueleto da estrutura holográfica. As peças apenas são coloridas nos modelos 3D, na verdade elas foram pintadas de preto.

Como referido na secção 4.1, a fonte da imagem é um ecrã de acrílico de retro-

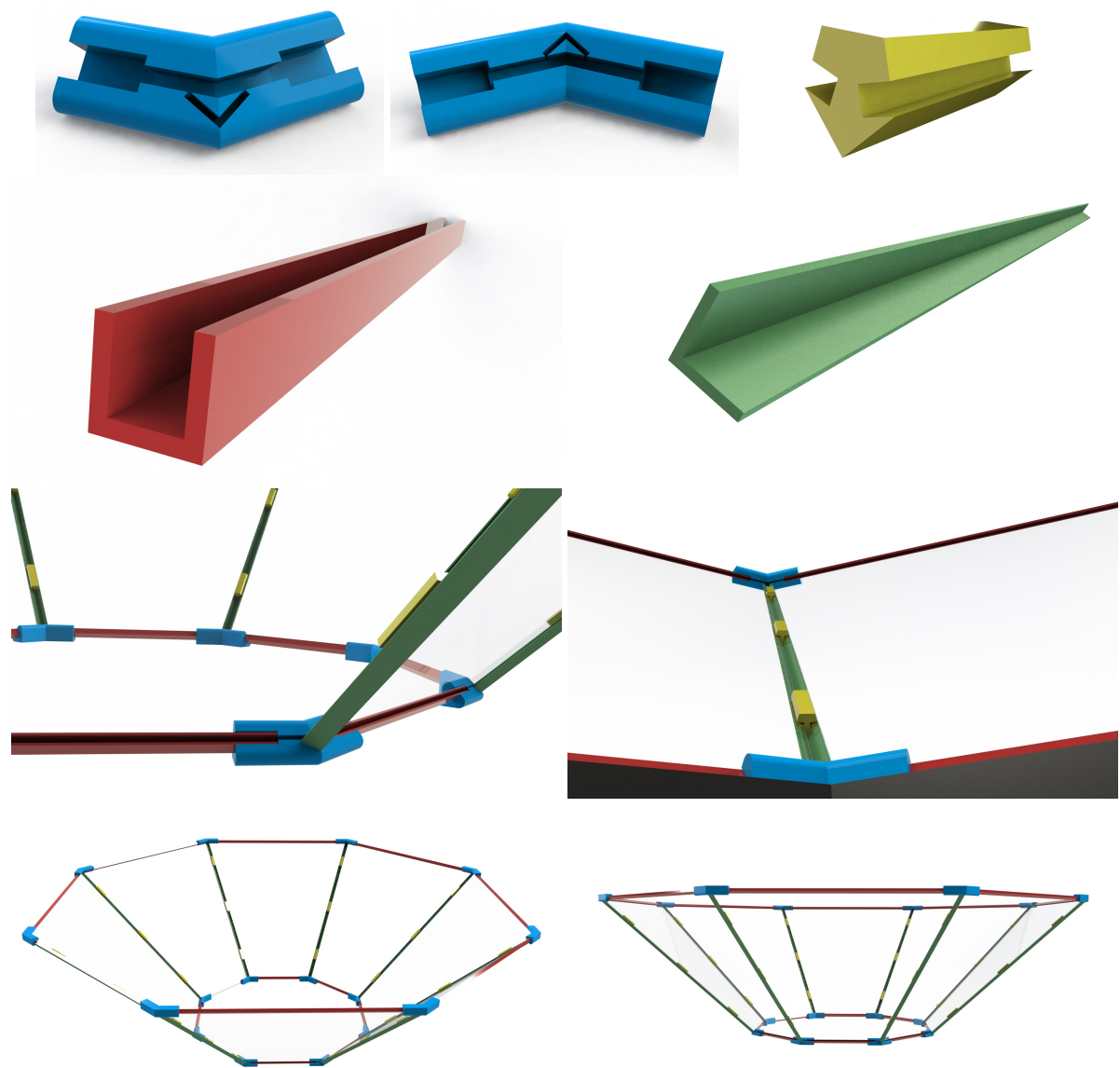


Figura 4.3: Modelos 3D de peças de alumínio para fortalecer a estrutura holográfica. De cima para baixo e esquerda para a direita, junção de cima e baixo (partes azuis), espaçador entre faces (parte amarela), perfil de alumínio em forma de U (vermelho), perfil de alumínio em forma de L (verde) e estrutura montada e fixa. As cores apenas contribuem para a melhor visualização dos contornos, na verdade as peças são pretas para refletir a luz o menos possível.

projeção de forma octogonal e é colocado abaixo da estrutura holográfica transparente. Cada lado da estrutura reflete a luz vinda de cada uma das 8 partes desta superfície com a imagem a ser criada por um único projetor colocado na parte de baixo e virado para cima. Desta forma, a estrutura dará a percepção que essas imagens estão a flutuar

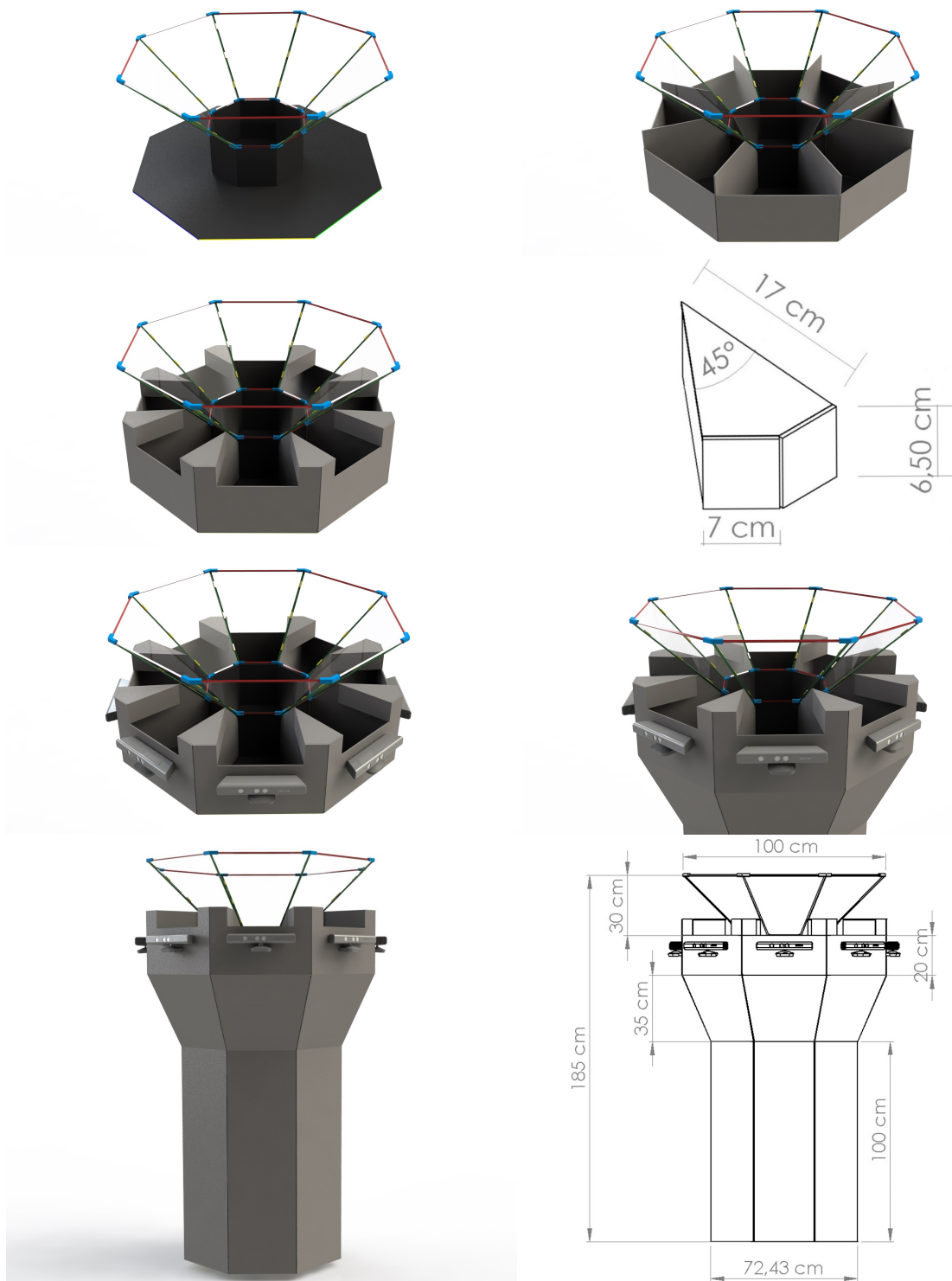


Figura 4.4: Sequência da construção da estrutura base. De cima para baixo e esquerda para a direita, fixação do ecrã de projeção pela retaguarda, divisórias para bloquear luz vinda de outras perspetivas (interferência entre perspetivas), estrutura em forma de "fatia de bolo" com o mesmo propósito anterior e as suas dimensões à direita, posicionamento das câmaras Kinect e estrutura base. As restantes figuras são a estrutura vista de várias perspetivas e as suas dimensões.

dentro da estrutura, criando o holograma segundo a técnica *Pepper's Ghost*. Um computador, com um software específico cria essas imagens/vídeos a partir de um objeto 3D ou então vários, um para cada perspectiva do objeto holográfico que se quer criar (ver secção 4.3.3). O holograma está atrás da película de acrílico à mesma distância que o ecrã está colocado dele. Por isso, para dar a ilusão que o holograma está colocado dentro e no meio da estrutura, o ecrã de retro-projeção terá que descer a mesma distância da qual se quer recuar o holograma. A base da estrutura tem 40 cm de diâmetro (d_2), por isso o ecrã de retro-projeção é colocado a uma distância $d_3 = 20$ cm ($d_3 = \frac{d_2}{2}$) da base da estrutura (Fig. 4.4 em cima à esquerda).

Na base da estrutura do holograma também foi colocado um ecrã com uma orientação vertical (representado nos modelos 3D na Fig.4.6 e Fig.4.13). Para além do holograma, este ecrã irá mostrar menus e informação auxiliar que o utilizador poderá controlar.

Como referido anteriormente nesta secção, a criação de um holograma disposto numa forma circular uniforme com mais de 4 perspectivas (menos de 90° entre cada perspectiva) provoca interferência entre perspectivas diminuindo a qualidade do holograma. Para resolver este problema foi necessário colocar placas opacas com o objetivo de bloquear a luz entre perspectivas de uma forma especial. Esta operação foi feita em duas etapas:

- (a) Entre cada perspectiva e entre a placa de retro-projeção e a base da estrutura holográfica da parte superior, aplicaram-se placas finas, colocadas na vertical com uma altura igual a d_3 (Fig. 4.4 em cima à direita). Este processo resolve 90% do problema das interferências.
- (b) Por cima das placas descritas em (a), foi colocada uma estrutura em forma de "fatia de bolo" como se pode observar na Fig. 4.4 (2ª linha, do lado esquerdo) e as suas dimensões Fig. 4.4 (2ª linha, do lado direito).

Para garantir a eficácia contra as interferências desta estrutura, foram realizados vários testes de reflexões de luz no modelo 3D da estrutura holográfica. Como imagem

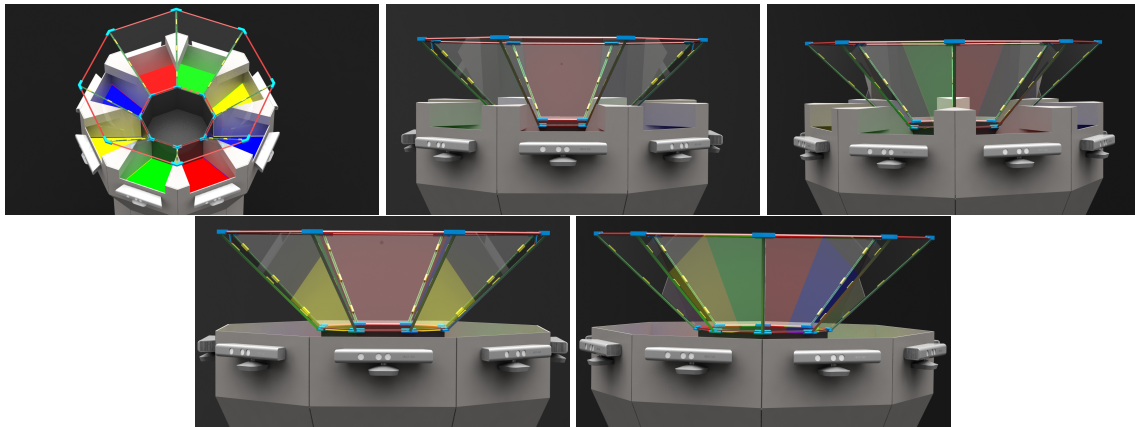


Figura 4.5: Testes de reflexão de luz nos modelos 3D da instalação holográfica. De cima para baixo e da esquerda para a direita, as luzes de testes, observação diretamente a uma face e na esquina com correção de interferências e a mesma observação à face e esquina sem correção de interferências.

de teste, cada uma das 8 partes da placa de retro-projeção foi iluminada com uma cor (Fig. 4.5, cima à esquerda). Cada lado teve assim uma cor igual ao seu lado oposto, num total de 4 diferentes cores. Na Fig. 4.5, meio e direita em cima, pode-se observar as reflexões da luz olhando diretamente para uma face e entre duas perspectivas respetivamente. Nestes dois caso, verifica-se que as luzes das outras perspectivas não interferem nas faces que se está a observar. É preciso reparar que no caso da imagem da direita, a luz de ambas as cores verde e vermelha são refletidas, mas cada uma na sua face e apenas em metade, o que é o ideal. Se a imagem fosse por exemplo um carro com uma perspectiva lateral, conseguia-se observar metade do carro refletido numa face e a outra metade na outra face, completando o total do carro. Nas duas imagens de baixo pode-se observar as reflexões no caso onde não foram aplicadas estruturas para corrigir as interferências. No caso da imagem da esquerda, apenas se deveria observar a perspectiva frontal, mas sem as correções, as perspectivas em amarelo também são refletidas. Quando se observa de uma esquina (entre duas faces), para além das metades de cada uma das perspectivas serem refletidas, também são refletidas todas as perspectivas vizinhas, erradamente.

As estruturas de correção de interferências retificam efetivamente o problema e permitem que os utilizadores possam andar à volta da estrutura observando uma conti-

nuidade suave em todas as 8 perspetivas mesmo quando olham diretamente para as junções entre cada face acrílica (resultados práticos na secção 4.4).

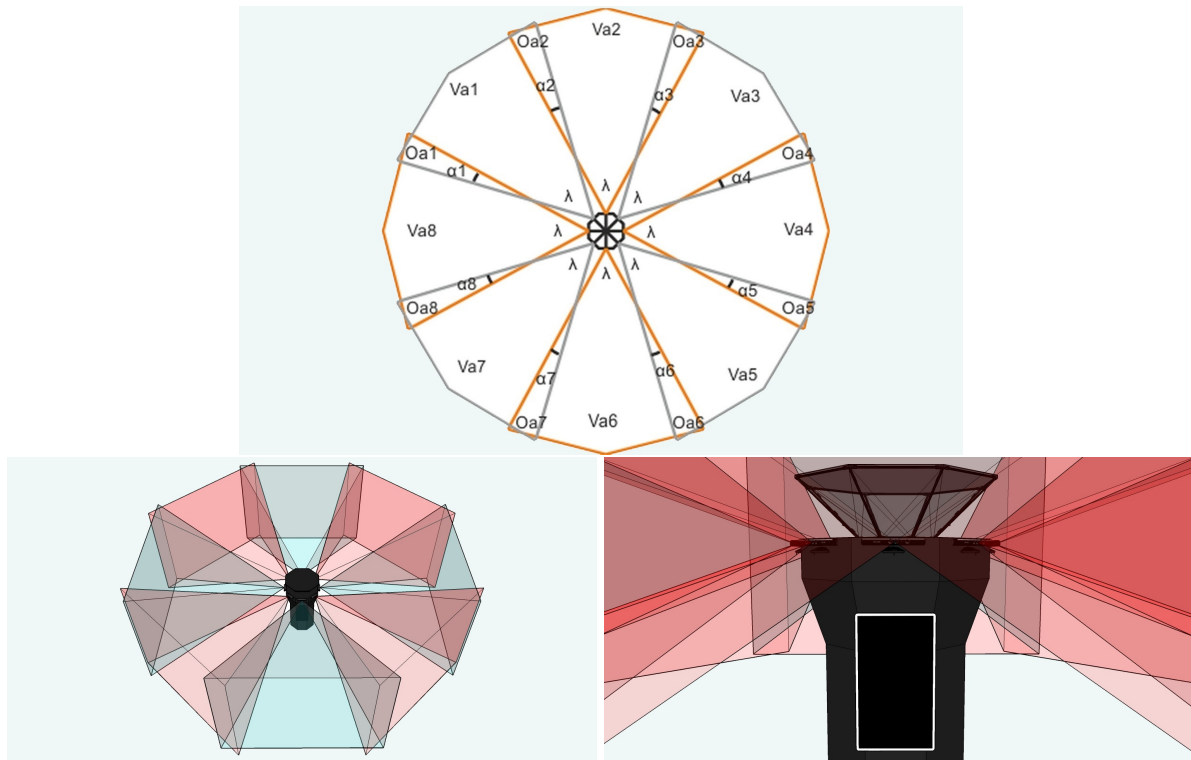


Figura 4.6: Ângulos do posicionamento entre cada Kinect à esquerda e visualização dos campos de visão das 8 Kinect nas duas imagens à direita.

Como mencionado anteriormente, foram usadas 8 Kinect na instalação, onde cada uma tem um campo de visão de 57° na horizontal e 47° na vertical, fazendo com que todas elas interceitem o campo de visão das Kinect vizinhas por $\alpha = 12^\circ$. Este ângulo de interceção pode ser calculado multiplicando o número de Kinect pelo seu campo de visão horizontal ($\lambda = 57^\circ$), subtraindo 360° e dividindo por 8, $\alpha = (8 \times \lambda - 360^\circ)$ (Fig. 4.6, cima). As outras figuras em baixo na Fig. 4.6 representam a instalação com os 8 campos de visão das câmaras Kinect representando as zonas de interceção. Também é mostrado na imagem da direita a zona onde o ecrã lateral está colocado.

4.2.1 Especificações de hardware

Como referido atrás, a instalação contém um projetor no seu interior, que atua como a fonte das imagens holográficas. O projetor necessita de projetar a área completa da película de retro-projeção (forma octogonal com 1 metro de diâmetro) que está colocada a 1,35 metros do solo por isso ele necessita de ser de ultra-curta distância. Optamos pelo projetor Optoma W307UST por conseguir projetar a área pretendida e a tão curta distância.

Como a instalação possui 8 câmaras Kinect, necessitamos de um computador com alguma potência, e por esse motivo escolhemos um processador Intel i7 de 4^o geração com 16GB de RAM e uma placa dedicada PCI Express HUB USB para conseguir ligar todas as câmaras ao computador.

4.3 Interface Humano-Computador

Para controlar e gerir a instalação holográfica foram criados quatro programas, cada um com uma função distinta, mas relacionada, na instalação:

- (i) Menu interativo com controlo pelo utilizador (secção 4.3.4). Este programa corre no ecrã lateral e o utilizador pode navegar por uma série de menus, imagens, textos e vídeos.
- (ii) O controlo dos menus utilizando gestos captados através das câmaras Kinect (secção 4.3.5).
- (iii) Gerador de holograma (secção 4.3.3). Este programa produz os hologramas a partir de um objeto 3D ou de um conjunto de vídeos com várias perspetivas. O projetor dentro da estrutura projeta as imagens no ecrã de retro-projeção, criando o holograma através da técnica *Pepper's Ghost*.
- (iii) Aplicação Android (secção 4.3.9) opcional para controlo do menu interativo (i).

4.3.1 Base de dados

Todos os três programas possuem informações armazenadas numa base de dados comum (Fig. 4.7). Esta base de dados está localizada num servidor da aplicação e tem como objetivo, o armazenamento das configurações dos programas, inserção de novos menus, conteúdos, imagens holográficas, armazenamento de estatísticas dos utilizadores, entre muitas outras informações. As tabelas da base de dados serão explicadas em detalhe nas secções seguintes.

4.3.2 Sistema de Kinects e utilizadores

Embora toda a instalação tenha sido desenhada, quer em termos de hardware como de software, de modo a acomodar o uso das câmaras Kinect, o programa foi desenvolvido de forma a funcionar com ou sem as câmaras. O campo *KinectNumber* da tabela *configurations* define o número de câmaras usadas na instalação entre 0 a 8 câmaras, sendo o valor 0 escolhido quando não se usam câmaras.

Caso se utilize uma ou mais câmaras, todas elas funcionam em conjunto para reconhecer os utilizadores à volta da instalação, ou seja, um utilizador pode percorrer toda a área da instalação sendo reconhecido por todas as câmaras como o mesmo utilizador. Só existe no máximo um utilizador num determinado tempo que pode interagir com a instalação, obviamente isto não é uma limitação mas sim uma funcionalidade pelo motivo de não fazer sentido estarem dois utilizadores em simultâneo a interagir com o holograma. Por essa razão, o primeiro utilizador que entra na instalação tem a prioridade de interação, e só a perde quando sair da área da qual as Kinects conseguem detetar os utilizadores

4.3.3 Sistema gerador de hologramas

O programa que gera os conteúdos holográficos nas 8 faces foi também produzido em Unity 3D e corre num programa separado do que gere os menus e câmaras Kinect. Os

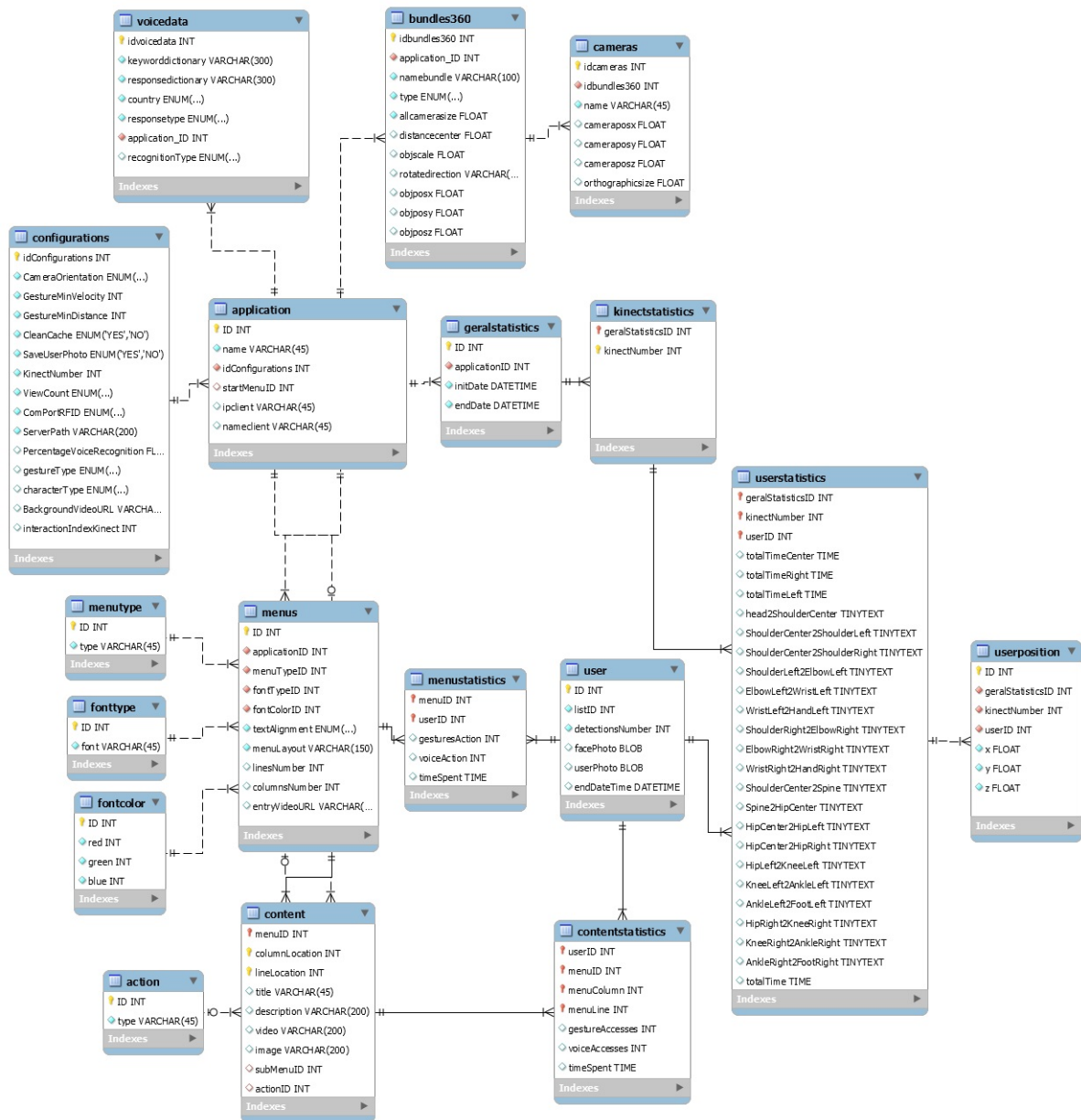


Figura 4.7: Base de dados no servidor da estrutura holográfica.

dois programas comunicam por TCP, com o programa do menu principal a funcionar como servidor TCP, e este programa o cliente. O objetivo deste programa é produzir as imagens necessárias para criar o holograma a partir de:

- Vídeos
- Objetos 3D

No caso do holograma ser composto por vídeos, este é criado a partir de um vídeo que inclua as perspetivas nas suas posições corretas (imediatamente abaixo de cada película refletora da instalação holográfica). Na Fig. 4.8 (1ª linha e coluna), pode-se observar um exemplo de uma imagem do vídeo da cabeça de uma mulher. Para ativar o holograma a partir de vídeos, no campo *characterType* da tabela *configurations* da base de dados terá que estar definido numa quarta opção como já discutido na secção Personagens digitais. Neste caso, cada vez que o utilizador escolher um menu, é enviado numa ligação TCP/IP o comando para este programa, com a localização do vídeo que se quer correr no holograma associado a esse menu.

No caso dos objetos 3D, o holograma apresenta objetos 3D que podem ser de dois tipos, personagens 3D ou objetos/modelos estáticos. Para criar as 8 perspetivas a partir de um objeto 3D, colocaram-se 8 câmaras virtuais (no editor Unity 3D) à volta do objeto e direcionado para ele, como se pode observar na Fig. 4.8 (1ª linha, 2ª e 3ª coluna). Para criar a imagem holográfica com as 8 perspetivas, dispuseram-se as imagens de cada câmara à volta de um ponto central, criando a imagem final com todas perspetivas do objeto colocadas de forma circular (Fig. 4.8, 1ª linha, 4ª coluna). Da mesma forma, também se pode criar a mesma imagem mas com 4 perspetivas em vez de 8, caso se use uma instalação com 4 perspetivas (Fig. 4.8, 2ª linha e 1ª coluna). Nas restantes imagens estão representados outros objetos ou personagens 3D. A grande vantagem da utilização dos objetos 3D ao invés dos vídeos deve-se ao facto deles oferecerem mais interação com o utilizador e poderem ser configurados facilmente para qualquer tamanho de instalação holográfica. Isto deve-se em primeiro lugar, ao facto dos objetos 3D possam ter várias animações associadas (no caso de uma personagem animada), podendo interagir de forma mais natural com o utilizador que um vídeo que foi produzido apenas para um propósito. Cada vez que um novo utilizador entra na instalação, é enviado um comando por TCP, e ativada uma animação onde o boneco 3D levanta a mão acenando para o utilizador. Para além disso, também é enviado por TCP, a direção na qual o utilizador foi detetado fazendo a personagem 3D

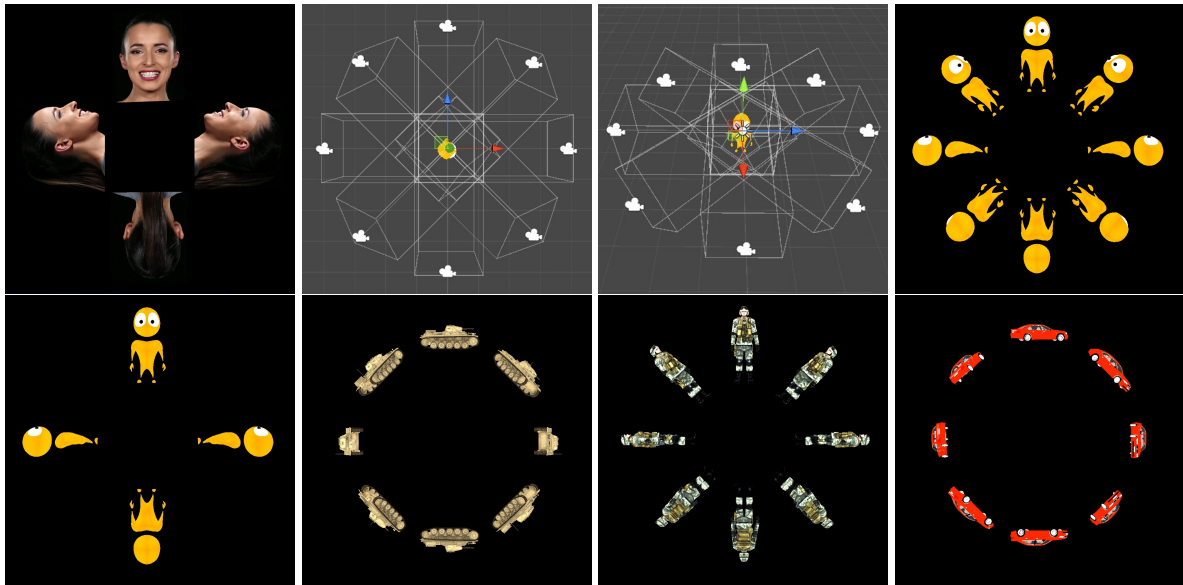


Figura 4.8: Vários exemplos das imagens do holograma de 4 e 8 perspectivas. Da esquerda para a direita e de cima para baixo, video da cabeça de uma mulher com 4 perspectivas, editor no Unity com as câmaras a apontar para um avatar 3D, avatar amarelo com 8 perspectivas, o mesmo avatar com 4 perspectivas, carro com 8 perspectivas, um soldado e um carro.

virar-se para o utilizador. A outra grande vantagem do uso dos objetos 3D reside no facto de facilmente poder-se corrigir ou configurar a posição, independente de cada uma das 8 perspectivas, colocando todas as perspectivas no sítio correto no holograma. Estes parâmetros podem ser modificados na tabela *bundles360* e na *cameras* da base de dados, com os campos:

- *distancecenter* - Define o afastamento interior (diâmetro interior) entre as várias perspectivas. No holograma, este parâmetro provoca a subida e descida do objeto.
- *objscale* - Define o tamanho do objeto para todas as perspectivas.
- *objpos x, y, z* - Define a posição global do objeto de todas as perspectivas.
- *camerapos x, y, z* - Semelhante à opção anterior mas independente para cada perspectiva.
- *orthograficsize* - Esta opção permite alterar a dimensão do objeto independente para cada perspectiva.

4.3.4 Sistema de menus interativos

Como explicado anteriormente, na zona lateral da estrutura está colocado um ecrã. Este programa apresenta nesse ecrã, um sistema de menus (i) onde se podem observar conteúdos em forma de texto, imagens e vídeos. Para além das informações visuais, este programa também é responsável pelas câmaras Kinect à volta da instalação, ou seja, a aplicação recolhe e processa todos os dados acerca dos utilizadores à volta da instalação, guardando-os na base de dados em forma de estatísticas para além de também usar esses dados para detetar gestos (usados para navegação) e reconhecimento de voz. Adicionalmente, esta aplicação também é responsável pelo controlo de um sistema RFID que permite ao utilizador fazer *check in* e despoletar uma ação na instalação.

Menus

A base de dados pode conter mais que uma aplicação e diferente entre elas, inseridas na tabela *application*. Cada aplicação introduzida na base de dados refere-se a uma instalação holográfica diferente. Cada aplicação pode conter vários menus (tabela *menus*), podendo ser escolhidos entre 7 de diferentes tipos (Fig. 4.9) gravados na tabela *menuType* na base de dados. Os 7 diferentes tipos de menus são os seguintes:

- **Simples** (Fig. 4.9, cima, primeira da esquerda): Este menu apenas apresenta um título. É usado principalmente como o menu principal, o menu de entrada na aplicação, mostrando opções para entrar nos outros menus que contêm mais informações. As várias opções movem-se verticalmente consoante as escolhas de navegação do utilizador.
- **Titulo_Texto** (Fig. 4.9, cima, segunda da esquerda): Os conteúdos são títulos e descrições. A opção que estiver selecionada (no meio do ecrã) mostra a descrição

enquanto as outras são escondidas e só mostram os títulos.

- **Titulo_Texto_Imagem** (Fig. 4.9, cima, terça da esquerda): Cada opção é composta por um título, uma descrição e uma imagem ou vídeo. A opção escolhida situa-se no meio do ecrã onde a imagem ou vídeo é apresentado.
- **Diagonal** (Fig. 4.9, cima, quarta da esquerda): Este menu possui a mesma funcionalidade que o menu *Titulo_Texto_Imagem*, mas em vez de apresentar as opções na vertical, apresenta-os na diagonal.
- **Imagem** (Fig. 4.9, baixo à esquerda): Este menu apenas possui um conteúdo na forma de uma imagem ou de um vídeo que ocupa totalmente a metade de baixo do ecrã.
- **Mural** (Fig. 4.9, baixo ao meio): Apresenta vários conteúdos dispostos na vertical com um título e uma imagem ou um vídeo.
- **Selfie** (Fig. 4.9, baixo à direita): Tem a mesma configuração do menu *Mural* mas com a característica de ter a opção de gravar uma fotografia do utilizador (com ação tomada pelo próprio através de um gesto com o braço).

Configurações

A fonte de letra, o alinhamento e a cor do texto dos títulos e descrições dos conteúdos de cada menu podem respetivamente ser configurados pelos campos *fontTypeID*, *textAlignment* e *fontColorID* na tabela *menus*. Também na tabela *configurations* podem ser definidos valores que especificam qual a câmara Kinect (das 8 à volta da instalação) possui a funcionalidade de interação com o utilizador, o tipo de personagem digital (secção Personagens digitais), o tipo de deteção de gestos com a câmara Kinect e os parâmetros da distância e velocidade mínima que o utilizador terá que movimentar a mão para navegar no menu (secção Reconhecimento de gestos), a percentagem de certeza de reconhecimento de voz secção Reconhecimento de voz, o número da porta

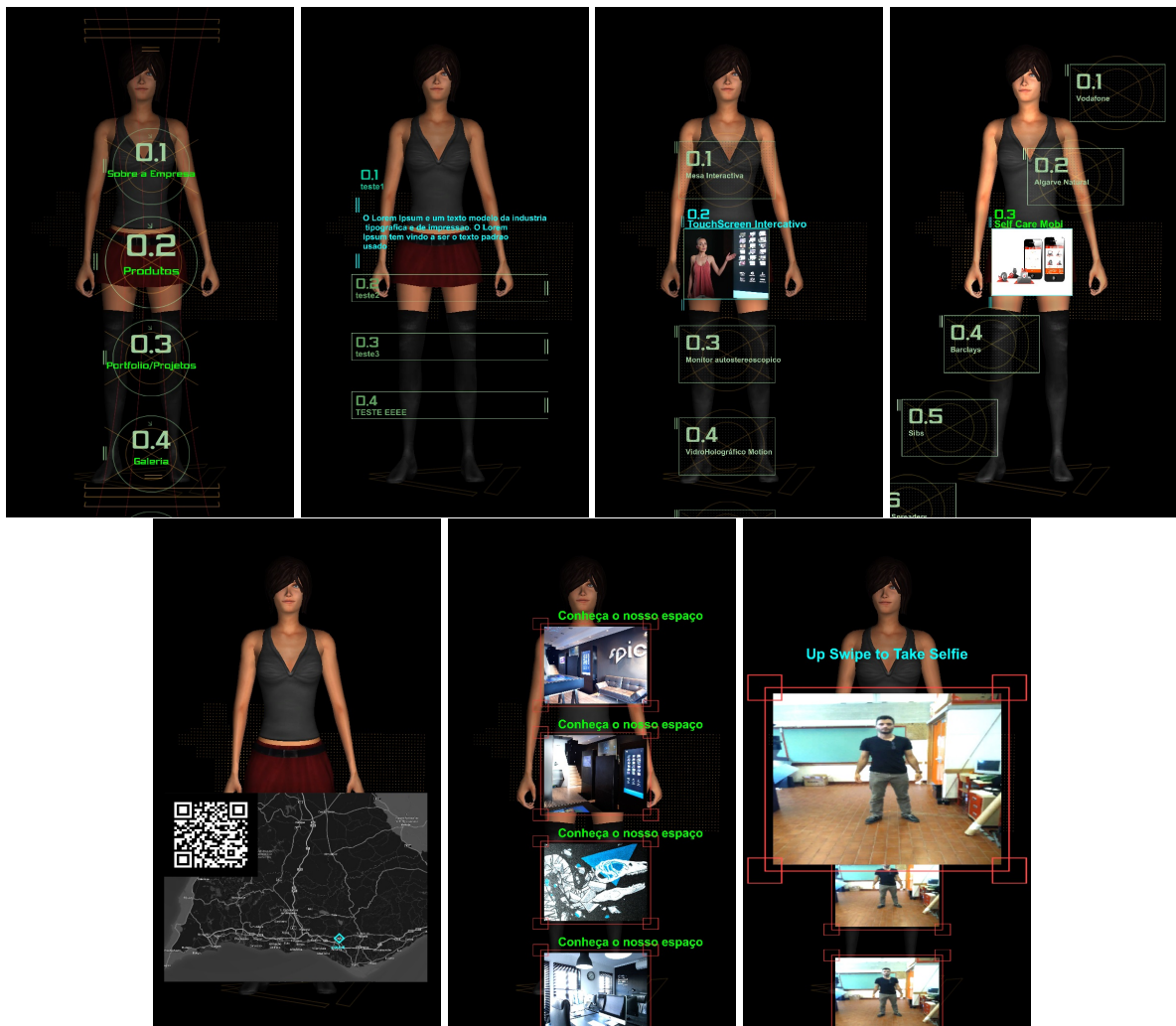


Figura 4.9: Os 7 diferentes tipos de menus. Da esquerda para a direita e de cima para baixo, menu *Simples*, *Titulo_Texto*, *Titulo_Texto_Imagem*, *Diagonal*, *Imagem Mural* e *Selfie*.

série (*COM Port*) caso se use um sistema de detecção com cartões RFID (secção Sistema RFID), o número de perspectivas usado no holograma e o número de câmaras Kinect.

Personagens digitais

O campo *characterType* da tabela *configurations* define três tipos de personagens existentes juntamente com os menus. Pode ser definido uma personagem 3D com uma animação simples (Fig. 4.10, esquerda), vídeos que apresentam/introduzem os menus ou um vídeo de fundo (Fig. 4.10, meio). Estes vídeos são definidos no campo *entryVideoURL* da tabela *menus*. Cada vez que o utilizador escolhe um menu, o vídeo respetivo

aparece, e imediatamente após o vídeo acabar, o vídeo desaparece e o menu volta a aparecer. Por fim, no campo *BackgroundVideoURL* da tabela *configurations* define a localização de um vídeo, de uma pessoa ou não, que se deseja que apareça por trás dos menus (Fig. 4.10, direita).



Figura 4.10: Os 3 diferentes tipos de personagens nos menus. Da esquerda para a direita, personagem 3D, vídeos de apresentação para cada menu e vídeo de fundo (atrás das opções dos menus).

Existe uma quarta opção neste campo da base de dados, caso se queira que no holograma, na parte superior da instalação, sejam usados vídeos em vez de objetos 3D. Neste caso, o campo *entryVideoURL* da tabela *menus* define a localização dos vídeos que aparecem no holograma. Cada vez que o utilizador escolhe um menu, esse vídeo aparecerá como holograma. Mais informações na secção 4.3.3.

4.3.5 Reconhecimento de gestos

Quando se usa uma câmara Kinect, o utilizador tem a possibilidade de interagir com os menus através de um sistema de reconhecimento de gestos. O SDK da câmara Kinect deteta todo o esqueleto da pessoa à sua frente, por isso, para este caso, o programa

retira a posição espacial principalmente da mão direita do utilizador entre outros descritos de seguida. Para tornar a interação o mais natural possível, decidiu-se reconhecer os gestos *swipes* (movimento rápido com a mão numa determinada direção), dependendo da distância e velocidade com que é realizado. Se o utilizador realizar um *swipe* com a mão direita, cumprindo uma velocidade mínima (v_s) e distância mínima (d_s) estabelecida, então é considerado um gesto reconhecido com sucesso. Para detetar estes movimentos, é criada uma janela de tempo Δt_s calculado através de $\Delta t_s = d_s/v_s$ com o tempo necessário para analisar no histórico da posição da mão do utilizador ($p_{\{x/y\}}$). Analisando essa posição entre t até $t - \Delta t_s$, se a posição da mão ter percorrido uma distância total (d_t) superior a d_s e uma velocidade média (v_m) superior a v_s então será considerado que houve um *swipe*. Para calcular a distância total, é somado dentro da janela de tempo Δt_s , todas as diferenças temporais de $p_{\{x/y\}}$ através de $d_t = \sum_{j=t-\Delta t}^{t-1} (p_{j+1_{\{x,y\}}} - p_{j_{\{x,y\}}})$. De forma semelhante, para a velocidade média, são calculadas todas as velocidades entre as várias posições da mão $p_{\{x/y\}}$ na janela de tempo Δt_s e calculado a sua média.

Os valores de d_t e v_t podem possuir valores positivos e negativos, sendo que a direção do *swipe* é o que o define. Se $d_{t_x} > d_s$ e $v_{t_x} > v_s$ então é considerado um *swipe* da direita para a esquerda. Se $d_{t_y} < -d_s$ e $v_{t_y} < -v_s$ então é considerado um *swipe* de baixo para cima. Se $d_{t_y} > d_s$ e $v_{t_y} > v_s$ então é considerado um *swipe* de cima para baixo.

Estes dois critérios (d_s e v_s) são usados para definir a sensibilidade e podem ser alterados nos campos *GestureMinDistance* e *GestureMinVelocity*, respetivamente, na tabela *configurations* da base de dados. Nos testes efetuados utilizaram-se os valores de 100 *cm/s* e 30 *cm* para a velocidade e distância mínima respetivamente. Para seleccionar os menus ou voltar ao menu anterior, o utilizador terá que movimentar a mão direita continuamente da direita para a esquerda (*swipe* para a esquerda).

Para navegar pelos vários conteúdos em cada menu o utilizador terá duas opções:

- **Swipe cima/baixo** - Esta opção é semelhante à anterior (seleccionar o menu), mas

realiza-se com o movimento de *swipe* para cima e para baixo com os mesmos parâmetros de velocidade e distância mínima que os anteriores.

- **Braço direito inclinado para cima/baixo** - Este movimento é diferente do *swipe* e faz-se ao manter o braço inclinado para cima ou baixo. Para detetar a inclinação do braço, para além da posição espacial da mão, também é retirado a do cotovelo, pescoço e início da coluna, podendo assim calcular facilmente a sua inclinação como descrito de seguida.

Considerando o vetor definido pelo corpo do utilizador $\vec{V}_{bd} = (x_{bd}, y_{bd}, z_{bd})$, calculado ao subtrair as posições do esqueleto do pescoço com posição do início da coluna, e usando outro vetor que define o braço do utilizador $\vec{V}_{ar} = (x_{ar}, y_{ar}, z_{ar})$ calculado ao subtrair a posição da mão com o cotovelo. O ângulo do braço é calculado através de $\theta = \text{acos}((\vec{V}_{bd} \cdot \vec{V}_{ar}) / (||\vec{V}_{bd}|| ||\vec{V}_{ar}||))$. Para valores de θ entre $100^\circ \leq \theta \leq 160^\circ$ é considerado que o utilizador efetuou o gesto para cima, e inversamente, para θ entre $20^\circ \leq \theta \leq 80^\circ$ é considerado que o utilizador efetuou o gesto para baixo. Todos os outros valores indicam que o utilizador não está a efetuar gestos.

Se o utilizador mantiver o braço inclinado, de 1 em 1 segundo, as opções nos menus movem-se de acordo com a direção da inclinação. Se o utilizador colocar a mão horizontalmente com o chão, ou completamente estendida para baixo (em repouso), é considerado que não existe movimento.

4.3.6 Reconhecimento de voz

A tabela *voicedata* configura o sistema de reconhecimento de voz com resposta por voz. Nesta tabela existe um dicionário de palavras/códigos para deteção (*keyword-dictionary*) e outro dicionário de resposta (*responsedictionary*). A deteção das palavra chave pode ser uma frase para deteção pelo sistema de reconhecimento de voz da Kinect, ou então o código atribuído ao cartão RFID (secção Sistema RFID). A resposta

pode ser de dois tipos diferentes (definidos pelo campo *responseType*), a reprodução de um ficheiro de som (.wav) ou a conversão de texto para voz (*Text to Speech*). No caso da reprodução do ficheiro de som, o campo *responsedictionary* será composto pelo endereço onde o ficheiro se encontra na máquina, e caso se utilize a conversão de texto para voz, este campo possuirá a frase a ser convertida. Quando usado o reconhecimento de voz pela Kinect, cada vez que a câmara retorna uma palavra/frase reconhecida, é devolvido também um grau de certeza entre 0-100% de que aquela palavra/frase foi dita. Este grau de certeza é definido no campo *PercentageVoiceRecognition* da tabela *configurations*.

Para mais detalhes consultar Alves et al. (2015).

4.3.7 Sistema RFID

Para o sistema RFID, usou-se um leitor e cartões RFID de 125 kHz que se conecta ao computador por porta COM. Para ler os cartões, eles têm que se aproximar no máximo 10 cm do leitor, e quando detetados, o leitor envia o código do cartão por USB. O campo *PortComRFID* da tabela *configurations* define a porta COM do leitor, conectado ao computador. O sistema de reconhecimento com os cartões RFID faz-se, como na secção anterior, também pela tabela *voicedata*. O campo *recognitionType* define o tipo de deteção da palavra/código, e neste caso como é usado um cartão RFID, o dicionário de palavras para deteção será o código atribuído ao cartão. O dicionário de resposta funciona da mesma forma que o explicado na secção anterior e a resposta poderá ser também a reprodução de um ficheiro de som ou a conversão de texto para voz.

4.3.8 Estatísticas

Os dados retirados das câmaras Kinect, para além de serem usadas para detetar os gestos dos utilizadores e reconhecimento de voz, também são usadas para preencher toda uma série de dados estatísticos na instalação holográfica.

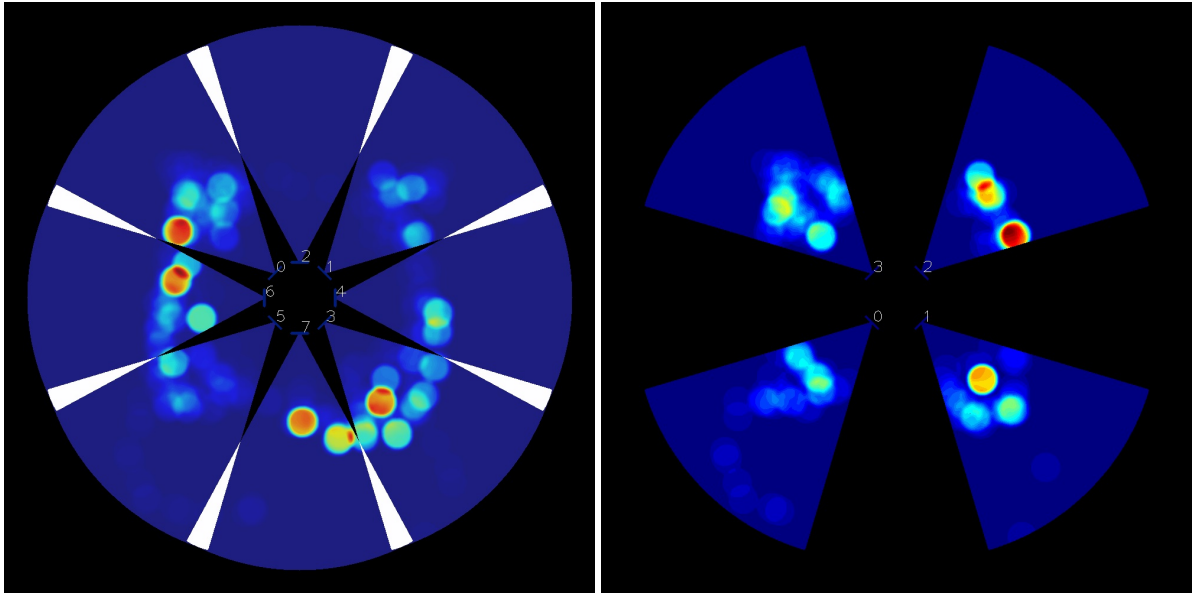


Figura 4.11: Mapas de calor de vista aérea numa instalação com 8 e 4 Kinect na imagem à esquerda e direita respetivamente.

Sempre que um novo utilizador entra no campo de visão de uma câmara, à volta da instalação, é adicionado um novo utilizador na tabela *user* na base de dados. Se esse utilizador for o que está a interagir com a aplicação, também será gravada a foto da face e corpo inteiro, na mesma tabela. Para além disso, também será atualizado o número de vezes que o utilizador entrou na instalação (caso tenha sido reconhecido como um utilizador anteriormente presente na instalação).

Quando o utilizador sai da área da instalação, são atualizados uma série de dados estatísticos. Primeiro, é introduzido na tabela *menustatistics* o número de vezes que o utilizador esteve em cada menu e quando tempo permaneceu em cada um deles. Da mesma forma, também são introduzidos na tabela *contentstatistic* o número de vezes que o utilizador visualizou cada conteúdo e quanto tempo permaneceu a visualizar. Na tabela *userstatistics* também são introduzidos o tempo que o utilizador olhou para a sua esquerda, centro e direita, o tempo total na instalação, como também dados relativos ao comprimento entre cada duas junções do corpo. Estes últimos dados têm como objetivo, ajudar a identificar o utilizador, caso este volte a entrar na instalação e só funcionam se o utilizador for o que esteve a interagir. Por fim, também é gra-

vada a posição do utilizador à volta da instalação com o objetivo de produzir mapas de calor, ou seja, os locais onde os utilizadores estiveram mais e menos tempo. Na Fig. 4.11 podem-se observar dois mapas de calor (vista aérea) usando 8 e 4 câmaras Kinect, respetivamente, do lado esquerdo e direito. Os locais mais frequentados são representados a cores mais quentes e vice-versa.

Para mais detalhes consultar Alves et al. (2015).

4.3.9 Aplicação Android

A aplicação Android foi desenvolvida como um complemento para o controlo da instalação holográfica. Quando a aplicação holográfica é iniciada, é gravado na base de dados, no campo *ipclient* da tabela *applications*, o IP da máquina respetiva. Assim, quando a aplicação Android se inicia, são retirados vários dados através de um *web-service* no servidor em forma de JSON, permitindo à aplicação ler o IP da máquina e como também todos os dados acerca dos menus e conteúdos presentes na instalação holográfica. Sabendo o IP da máquina do holograma, é também estabelecida uma comunicação TCP por um porto previamente definido entre a instalação holográfica e a aplicação Android.

Se tudo correr bem, e os dados forem retirados com sucesso, a aplicação inicia-se com um menu lateral (Fig. 4.12, esquerda).

O menu de interação na aplicação Android (Fig. 4.12, 2ª coluna) lista todos os menus existentes no holograma, e ao selecionar um, é enviado um comando para o holograma com o nome do menu que o utilizador escolheu, ativando esse menu no holograma. Alternativamente, também existe outro menu de interação em forma de comando (Fig. 4.12, 3ª coluna). Como se pode observar na figura, existem 5 botões diferentes permitindo o controlo total do holograma. Os botões de cima e baixo servem para navegar entre os vários conteúdos no holograma, o botão central para selecionar o menu, o da esquerda para voltar ao menu anterior e o botão da direita para começar a experiência holográfica do início.



Figura 4.12: A aplicação Android. Da esquerda para a direita, o menu lateral da aplicação, seleção dos menus da instalação holográfica, seleção dos menus em forma de comando, leitor de códigos QR e submissão de opiniões.

Existe também um leitor de códigos QR (Fig. 4.12, 4ª coluna) capaz de ler um código disponível na instalação holográfica e abrir o seu conteúdo num navegador de Internet. Por fim, também existe um outro menu na aplicação que permite ao utilizador submeter a sua opinião (Fig. 4.12, 5ª coluna).

4.4 Resultados

O protótipo da instalação holográfica esteve ativo durante vários dias na Universidade do Algarve e também na empresa SPIC - Creative Solutions. Todos os algoritmos e hardware holográfico foram testados independentemente e em combinação. Foram realizados vários testes da instalação durante várias horas e vários dias consecutivos envolvendo:

Reconhecimento de gestos - A deteção de gestos com as câmaras Kinect tiveram uma taxa de sucesso superior a 90%. As falhas detetadas devem-se ao facto do utilizador não cumprir os requisitos da distância e velocidade mínima necessárias para haver deteção. Das duas opções possíveis para realizar os gestos verticais, os utilizadores preferiram a técnica de *swipes* do que manter a mão inclinada para cima ou para baixo, para navegar nos menus verticalmente.

Reconhecimento de voz - Os testes de reconhecimento de voz foram feitos em 3 diferentes cenários. O utilizador localizado 2 metros em frente à instalação pronunciou 100 palavras diferentes em 3 diferentes níveis de som ambiente. Testou-se com os níveis de áudio de ruído exterior da sala (42 dB), depois para um nível de ruído de 57,5 dB e finalmente para 67,5 dB. Os últimos dois ruídos foram captados num bar com muitas pessoas a falar. No primeiro teste, foram detetadas 94% das palavras com sucesso, no segundo teste foram detetadas 85% palavras com sucesso e por fim, no último teste, foram detetados 53% das palavras. Estes resultados são esperados uma vez que as palavras ditas pelo utilizador são camufladas pelo ruído de muitas pessoas a falar.

Reconhecimento RFID - O reconhecimento de cartões RFID não apresentou nenhuma falhas uma vez que na instalação apenas é feita a interface entre a aplicação e um leitor RFID. Qualquer leitor RFID possui deteções com taxas de deteções próximos de 100%.

Qualidade do holograma - Durante os dias de testes, vários estudantes e professores foram convidados a utilizar a instalação. Todos os sistemas funcionaram como pretendido e sem problemas durante todos os dias. A não existência de interferências entre perspetivas permitiu a visualização do holograma em todo o redor da instalação de uma forma bastante suave, mesmo quando existem um número limitado de perspetivas.

A altura o holograma foi definida de acordo com o tamanho médio das pessoas, pelo que crianças com menos de 1,5 metros de altura, quando a uma distância inferior a 1 metro da instalação têm alguma dificuldade em visualizar o holograma completo. De forma semelhante, pessoas maiores de 2 metros de altura e quando próximos da instalação não conseguirem observar todo o holograma.

Na Fig. 4.13 pode-se observar várias imagens da instalação holográfica. Na primeira linha, a instalação completa com o ecrã lateral e nas restantes, vários ob-

jetos 3D vistos de várias perspetivas, inclusive nas zonas que poderiam ser mais problemáticas (junção entre duas perspetivas).



Figura 4.13: Vários resultados da instalação. Instalação completa na primeira linha e detalhes nas transições entre as prespetivas nas imagens das duas linhas centrais. As últimas 3 imagens da última linha representam outra instalação, baseada na descrita neste capítulo, onde as principais diferenças são o uso de 4 perspetivas/faces invés de 8 e apenas uma Kinect , colocada na perspetiva do ecrã lateral.

5

Conclusões e trabalho futuro

Nesta dissertação efetuou-se o estudo e desenvolvimento de vários sistemas de reconhecimento de gestos e aplicações interativas usando sensores 3D. Globalmente, foram apresentadas três soluções de aplicações de interfaces humano-computador, usando tecnologias diferentes.

Começou-se por apresentar uma solução de um sistema de reconhecimento de gestos através do sensor Leap Motion e três desenvolvimentos de aplicações/interfaces humano-computador que recorrendo ao mesmo, suportam aplicações sem fios e não táteis. Alternativamente ao Leap Motion, foi também desenvolvido um sistema de uma luva (GyroGlove) capaz de adquirir todos os ângulos tridimensionais dos dedos/mão (deteção da mão e dedos), e apresentada uma aplicação no âmbito educa-

cional usando esta nova tecnologia que permite uma mais rápida aprendizagem do alfabeto da linguagem gestual Portuguesa. Por fim, foi também desenvolvida uma instalação holográfica no âmbito do projeto "PRHOLO: O Relações Públicas Holográfico Realista" que recorrendo a outro sensor 3D (Kinect), permite às pessoas circundantes interagir com uma interface humano-computador com a capacidade de reconhecimento de gestos e voz.

Nas próximas secções serão apresentadas as conclusões e trabalho futuro para cada um destes sistemas de interfaces humano-computador com reconhecimento de gestos.

5.1 Reconhecimento de gestos com Leap Motion

O sistema descrito no capítulo 3, secção 3.1, apresentou alguns conceitos de uma interface natural de utilizador pela ilustração e desenvolvimento de três aplicações interativas de interfaces humano-computador. Estas interfaces são utilizadas em cenários em que o uso de outro tipo de soluções de interação mostrar-se pouco cómoda ou agradável (rato, teclado, ecrã tátil etc.). Embora os sistemas propostos ainda estejam em fase de prova de conceito, os testes realizados com vários utilizadores indicam que o sistema possui boas hipóteses de ser implementado nas situações reais propostas. A interface para o sistema PoolLiveAid é usada atualmente.

No entanto, alguns problemas ainda persistem quando se faz uso de um sensor 3D. Como foi descrito na introdução da dissertação, todos os sensores 3D apresentados funcionam com luz infravermelha, o que induz a algumas possíveis falhas se o mesmo for usado no exterior, ao sol, ou na proximidade de outra fonte de luz infravermelha (tal como alguns tipos de lâmpadas usadas dentro de edifícios). Nestes casos, a luz infravermelha provoca interferências no sensor acabando por prejudicar o seu funcionamento. O Leap Motion é o único sensor, dos apresentados, que contém menos problemas relacionados com as condições de luz, principalmente dentro de edifícios.

Este sensor consegue-se ajustar ligeiramente e automaticamente à intensidade de interferência externa e auto compensar essa interferência desde que não seja extrema.

Outro problema deve-se ao facto do sensor poder ter falsas deteções de movimentos. Para minimizar este problema, é possível alterar as propriedades de configuração da API do Leap Motion para cada tipo de aplicação ou pessoa que usa a interface (descrito na secção 3.1.1)). Para cada pessoa existem estes valores ideais que diminuem as falsas deteções de movimentos, pelo que poderá ser necessário uma configuração inicial da interface com o objetivo de determinar os melhores valores.

Para o trabalho futuro deste sistema, e tentando resolver alguns problemas descritos anteriormente, estão a ser estudadas novas formas de gestos que poderão conter uma menor taxa de falsos positivos e que ao mesmo tempo aumentem a facilidade de interação. Também para cada tipo de interface, justifica-se a criação de uma pequena configuração inicial para determinar a velocidade média e distância percorrida para cada gesto e programar a aplicação com esses valores, com o objetivo de diminuir as falsas deteções, fazendo com que o programa se adapte ao utilizador e não o contrário. Outra possibilidade, com a finalidade da diminuição de problemas relacionadas com a interferência da luz infravermelha, poderá passar pela criação de um pequeno habitáculo à volta do sensor, com uma entrada apenas para a mão/braço, com o objetivo de bloquear a luz infravermelha vinda do exterior caso o sistema esteja colocado em zonas não muito apropriadas para este tipo de sensores 3D.

5.2 Sistema GyroGlove

No capítulo 3, secção 3.2 foi proposto um protótipo de uma luva eletrónica capaz de detetar a posição e rotação 3D da mão e dos dedos e uma aplicação em forma de jogo, desenvolvida com o objetivo de estimular a aprendizagem da língua gestual Portuguesa do alfabeto. Esta luva introduziu uma nova abordagem aos sistemas tradicionais que se baseiam na aquisição de imagens, e.g., Motion (2014), Kinect (2014),

ASUSTeK (2015) e Structure (2014). Ao contrário destes sistemas, o GyroGlove não possui os mesmos problemas relacionados com o espaço restrito ou problemas com as condições luminosas. Para além disso, é um sistema sem fios e não necessita de ter as mãos próximas de um dispositivo especial.

Comparativamente, os sistemas descritos em Patel et al. (2011) e Kessler et al. (1995) mostram-se mais limitados ao apresentado nesta dissertação devido ao facto de não possuírem sensores 3D, ao invés disso, possuem resistências flexíveis ao longo de cada dedo, não permitindo saber a inclinação/rotação separadamente de cada zona anatómica dos dedos nem no espaço tridimensional. Assim, a proposta dos autores, apenas permite obter os valores aproximados das inclinações de cada dedo em relação à mão. O sistema GyroGlove, permite obter os mesmos dados com mais precisão e mais detalhe de toda a mão.

Esta nova tecnologia permite encontrar soluções para um variado conjunto de aplicações com utilidades educativas. Em complemento com a aplicação de aprendizagem da linguagem gestual, o utilizador poderá também treinar ao tentar traduzir linguagem gestual para texto. Também da mesma forma mas ao contrário, uma pessoa que tem como único meio de comunicação a linguagem gestual, poderá usar esta tecnologia para conseguir traduzir a sua linguagem gestual para texto, permitindo assim a comunicação com outras pessoas que não percebam linguagem gestual.

Como trabalho futuro, pretende-se em primeiro lugar, desenvolver uma versão do sistema GyroGlove mais prática e simples de usar. A versão atual (protótipo), necessita de um pequeno esforço para ser colocada, devido ao facto da luva necessitar de estar bem apertada (os sensores não podem estar soltos). Pretende-se assim, construir um modelo do sistema em estilo de esqueleto que assenta por cima da mão, de fácil colocação e que não limite o utilizador de fazer os movimentos da forma mais natural possível com a mão. Para além disso, também se pretende desenvolver um novo conjunto de aplicações interativas onde este dispositivo é uma mais valia.

5.3 Sistema holográfico

No sistema descrito no capítulo 4 foi apresentado uma instalação holográfica que pode ser visualizada em todo o seu redor e também um sistema que permite a interação por gestos, voz, cartões RFID e uma aplicação Android entre o utilizador e o holograma.

A instalação possui 8 câmaras Kinect que captam toda uma série de dados acerca dos utilizadores que podem interagir com a instalação, ou que apenas andam à sua volta, podendo assim criar vários dados estatísticos acerca das vezes que esteve em cada menu e conteúdo como também o tempo passado. O sistema também consegue criar mapas de calor (*heat maps*) que indicam os locais onde os utilizadores estiveram mais tempo na instalação e reconhecer utilizadores que já lá estiveram. Todos estes dados são introduzidos na base de dados, permitindo o fácil acesso por parte de uma empresa promotora do sistema holográfico.

A instalação pode ser configurada e ajustada ao conteúdo que se deseja mostrar. Todas as imagens, vídeos e textos podem ser ajustados ao tipo de menu, as letras podem mudar de cor, tipo de letra e alinhamento. Os vídeos e/ou objetos 3D que se deseja mostrar no holograma também podem ser configurados de acordo com o que se pretende mostrar.

Esta instalação foi o protótipo para uma versão comercial atualmente disponível através da empresa SPIC - Creative Solutions.

Para o trabalho futuro, está a ser estudado a criação de uma *framework* que permita a criação de novos tipos de menus, além dos 7 apresentados, para que facilmente sejam configurados outros menus com diferentes apresentações de imagens, textos, vídeos e animações. Para além disso, também está a ser estudada uma instalação holográfica que apresente uma área holográfica maior que a atual. O maior problema da escalabilidade do holograma é o facto de que a estrutura externa que o suporta tem que acompanhar esse crescimento, produzindo uma instalação maior e por vezes com impossibilidades práticas. Uma das soluções passa por virar a estrutura 180° vertical-

mente, ficando o projetor e a película de retro-projeção na parte superior, enquanto a zona transparente do holograma na parte de baixo da instalação. Assim, toda a parte da estrutura que oferece suporte ao holograma ficaria junto ao teto do edifício, libertando área junto ao holograma.

5.4 Lista de publicações

No que diz respeito à disseminação do trabalho desenvolvido, foram publicados os seguintes artigos:

- Rodrigues, J.M.F., Alves, R., Sousa, L., Negrier, A., Cardoso, P.J.S, Monteiro, J., Felizberto, P., Figueiredo, M.J.G., Mendes da Silva, B., Gomes, M., Bica, P. (2016) *PRHOLO: 360° Interactive Public Relations*, accepted for Handbook of Research on Human-Computer Interfaces, Developments, and Applications, IGI Global
- Sousa, L., Alves, R., Rodrigues, J.M.F. (2015) Augmented reality system to assist inexperienced pool players, accepted for The Visual Computer
- Sousa, L., Monteiro, J., Cardoso, P.J.S. e Rodrigues, J.M.F. (2015). Interface natural de utilizador baseado em reconhecimento de gestos usando o sensor Leap Motion. *Dos Algarves: A Multidisciplinary e-Journal*, 26(1), 107-121. DOI: 10.18089.DA-MeJ/2015.26.1.6
- Cardoso, P.J.S., Rodrigues, J.M.F., Carlos, L., Mazayev, A., Ey, E., Corrêa, T., Saleiro, M. (2015) *A freehand system for the management of orders picking and loading of vehicles*, In M. Antona and C. Stephanidis (Eds.): *Universal Access in Human-Computer Interaction 2015, Part IV*, LNCS 9178, pp. 422–431. DOI: 10.1007/978-3-319-20687-5_41
- Alves, R., Negrier, A., Sousa, L., Rodrigues, J. M. F., Felisberto, P., Gomes, M., & Bica, P. (2015). *Interactive 180° Rear Projection Public Relations*. *Procedia Computer Science*, 51, 592–601. <http://doi.org/10.1016/j.procs.2015.05.327>

- Alves, R., Sousa, L., Negrier, A., Rodrigues, J.M.F., Cardoso, P.J.S., Monteiro, J., Gomes, M., Bica, P. (2015) *PRHOLO: Interactive Holographic Public Relations*, In Proc. 3rd Int. Conf. on Advances in Computing, Communication and Information Technology, Birmingham, UK, May 26-27, pp. 124-128. ISBN: 978-1-63248-061-3. doi: 10.15224/978-1-63248-061-3-74
- Alves, R., Sousa, L., Negrier, A., Rodrigues, J.M.F., Monteiro, J., Cardoso, P., Felisberto, P., Bica, P. (2015) *Interactive 360° holographic installation*, submitted to The Visual Computer
- Sousa, L., Rodrigues, J.M.F., Monteiro, J., Cardoso, P.J.S., Semião, J., Alves, R. (2014) *A 3D Gesture Recognition Interface for Energy Monitoring and Control Applications*, In Proc. 13th Int. Conf. on Applications of Computer Engineering (ACE '14), Lisbon, Portugal, Oct. 30-Nov. 1, pp. 62-71. ISBN: 978-960-474-393-3
- Alves, R., Madeira, M., Ferrer, J., Costa, S., Lopes, D., Mendes da Silva, B., Sousa, L., Martins, J., Rodrigues, J.M.F. (2014) *Fátima revisited: An interactive installation*, In Proc. Int. Multidisciplinary Scientific Conf. on Social Sciences and Arts, Varna, Bulgaria, 2-9 Set., pp. 141-148
- Figueiredo, M.J.G., Sousa, L., Cardoso, P.J.S., Rodrigues, J.M.F., Gonçalves, C., Alves, R. (2014) *Learning Technical Drawing with Augmented Reality and Holograms*, In Proc. 13th Int. Conf. on Education and Educational Technology (EDU '14), Lisbon, Portugal, Oct. 30-Nov. 1, pp. 11-20. ISBN: 978-960-474-395-7
- Alves, R., Sousa, L., Rodrigues, J.M.F. (2013) *PoolLiveAid: Augmented reality pool table to assist inexperienced players*, In Proc. 21st Int. Conf. on Computer Graphics, Visualization and Computer Vision, Plzen, Czech Republic, 24-27 June, pp. 184-193. ISBN: 978-80-86943-75-6

Bibliografia

- Active8-3D (2015). 3D holographic projection displays. Retirado de <http://www.activ8-3d.co.uk>. Acedido em 30 Junho de 2015.
- Alves, R., Madeira, M., Ferrer, J., Costa, S., Lopes, D., da Silva, B. M., Sousa, L., Martins, J., and Rodrigues, J. (2014). Fátima revisited: an interactive installation. In *Int. Multidisciplinary Scientific Conf. on Social Sciences and Arts*, pages 141–148. SGEM.
- Alves, R., Negrier, A., Sousa, L., Rodrigues, J. M., Felisberto, P., Gomes, M., and Bica, P. (2015). Interactive 180° rear projection public relations. *Procedia Computer Science*, 51:592–601.
- Alves, R., Sousa, L., and Rodrigues, J. (2013). Poolliveaid: Augmented reality pool table to assist inexperienced players. In *21st Int. Conference on Computer Graphics, Visualization and Computer Vision*, pages 24–27.
- Antonio, S., Herrera, R., and Enriquez, E. (2013). Projection’s panel of models for touch screen. *Int. J. of Innovative Research in Computer and Communication Engineering*, 1(9):2057–2064.
- ASUSTeK (2015). Xtionprolive. Retirado de http://www.asus.com/Multimedia/Xtion_PRO_LIVE/. Acedido em 30 Julho de 2015.
- Atmel (2015). Atmel. Retirado de <http://www.atmel.com/devices/atmega328p.aspx>. Acedido em 10 Junho de 2015.
- Cardoso, P., Rodrigues, J., Carlos, L., Mazayev, A., Ey, E., Corrêa, T., and Saleiro, M. (2015). A freehand system for the management of orders picking and loading of vehicles. In *M. Antona and C. Stephanidis (Eds.): Universal Access in Human-Computer Interaction 2015, Part IV, LNCS 9178, pp. 422-431. DOI: 10.1007/978-3-319-20687-5_41*.
- Chen, F., Fu, C., and Huang, C. (2003). Hand gesture recognition using a real-time tracking method and hidden markov models. *Image and vision computing*, 21(8):745–758.
- Chung, I., Huang, C., Yeh, S., Chiang, W., and Tseng, M. (2014). Developing kinect games integrated with virtual reality on activities of daily living for children with developmental delay. In *Advanced Technologies, Embedded and Multimedia for Human-centric Computing*, pages 1091–1097. Springer.
- D’Strict (2014). 3D sensing holographic installation. <http://goo.gl/CyqKXe>. Retrived: Sep. 1, 2015.

- Dumas, B., Lalanne, D., and Oviatt, S. (2009). Multimodal interfaces: A survey of principles, models and frameworks. In *Human Machine Interaction*, pages 3–26. Springer.
- Figueiredo, M. J., Cardoso, P. J., Goncalves, C. D., and Rodrigues, J. M. (2014). Augmented reality and holograms for the visualization of mechanical engineering parts. In *Int. Conf. on Information Visualisation*, pages 368–373. IEEE.
- Flyway (2015a). 3D holographic projection the future of advertising. Retirado de <http://flyawaysimulation.com/news/3630/3d-holographic-projection-future-of-advertising>. Acedido em 10 Junho de 2015.
- Flyway (2015b). Advanced virtual assistant. Retirado de <http://airportone.com/airportvirtualassistancesystem.htm>. Acedido em 10 Junho de 2015.
- Gordon, N., Ristic, B., and Arulampalam, S. (2004). Beyond the kalman filter: Particle filters for tracking applications. *Artech House, London*.
- InvenSense (2015). Invensense. Retirado de <http://www.invensense.com/>. Acedido em 10 Junho de 2015.
- Kessler, G. D., Hodges, L. F., and Walker, N. (1995). Evaluation of the cyberglove as a whole-hand input device. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 2(4):263–283.
- Kinect (2014). Kinect for Windows. Retirado de <http://goo.gl/FGZT8X>. Acedido em 10 Novembro de 2014.
- Kubrick, S. (1968). 2001: A space odyssey. USA, UK: Metro-Goldwyn-Mayer (MGM).
- Li, Y. (2012). Hand gesture recognition using kinect. In *2012 IEEE 3rd International Conference on Software Engineering and Service Science (ICSESS)*, pages 196–199. IEEE.
- Mihaylova, E., Cody, D., Naydenova, I., Martin, S., and Toal, V. (2013). Holography-basic principles and contemporary applications. *InTech, Chapters*, 1:3–35.
- Monteiro, J., Cardoso, P. J., Serra, R., and Fernandes, L. (2014a). Evaluation of the human factor in the scheduling of smart appliances in smart grids. In *Universal Access in Human-Computer Interaction. Aging and Assistive Environments*, pages 537–548. Springer.
- Monteiro, J., Eduardo, J., Cardoso, P. J., and Semião, J. (2014b). A distributed load scheduling mechanism for micro grids. In *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 278–283. IEEE.
- Motion, L. (2014). Leap Motion. Retirado de <https://www.leapmotion.com/>. Acedido em 10 de Novembro de 2014.
- Patel, B., Shah, V., and Kshirsagar, R. (2011). Microcontroller based gesture recognition system for the handicap people. *Engineering Research and Studies, Surat, India*.

- Potter, L. E., Araullo, J., and Carter, L. (2013). The leap motion controller: a view on sign language. In *of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration*, pages 175–178. ACM.
- Ren, Z., Meng, J., Yuan, J., and Zhang, Z. (2011). Robust hand gesture recognition with Kinect sensor. In *of the 19th ACM international conference on Multimedia*, pages 759–760. ACM.
- Ren, Z., Yuan, J., Meng, J., and Zhang, Z. (2013). Robust part-based hand gesture recognition using Kinect sensor. *IEEE Transactions on Multimedia*, 15(5):1110–1120.
- Rennie, J. (2015). The tupac hologram, virtual ebert, and digital immortality. Retirado de <http://goo.gl/fQYUt2>. Acedido em 10 Junho de 2015.
- Rybka, J. and Janicki, A. (2013). Comparison of speaker dependent and speaker independent emotion recognition. *Int. Journal of Applied Mathematics and Computer Science*, pages 797–449.
- Sousa, L., Rodrigues, J., Monteiro, J., P. J., Semião, J., and Alves, R. (2014). A 3D gesture recognition interface for energy monitoring and control applications. In *13th Int. Conf. on Applications of Computer Engineering*, volume 1, pages 62–71. ACE '14.
- Spielberg, S. (2002). *Minority report*. USA: 20th Century Fox; DreamWorks Pictures.
- Structure (2014). Structure sensor. Retirado de <http://structure.io/>. Acedido em 10 Novembro de 2014.
- Sutton, J. (2013). Air painting with corel painter freestyle and the leap motion controller: a revolutionary new way to paint! In *ACM SIGGRAPH 2013 Studio Talks*, page 21. ACM.
- SweetHome (2015). Sweet home 3D. Retirado de www.sweethome3d.com. Acedido em 10 Janeiro de 2015.
- Unity (2014). Unity 3D. Retirado de <https://unity3d.com/pt>. Acedido em 10 Novembro de 2014.
- Vizoo (2015). Cheoptics. Retirado de <http://www.vizoo.com/flash>. Acedido em 30 Junho de 2015.
- Wachowski, A. & Wachowski, L. (1999). *The matrix*. USA, Australia: Warner Bros.
- Wells, S. (2002). *The time machine*. USA: Warner Bros.
- Wiki (2014). Dactiologia portuguesa. Retirado de https://pt.wikipedia.org/wiki/L%C3%ADngua_gestual_portuguesa. Acedido em 10 Novembro de 2014.
- Yoon, H.-S., Soh, J., Bae, Y. J., and Yang, H. S. (2001). Hand gesture recognition using combined features of location, angle and velocity. *Pattern recognition*, 34(7):1491–1501.