

INTERAÇÃO PESSOA-MÁQUINA

INTERFACE MULTIMODAL



RELATÓRIO TÉCNICO

Projeto FEUP 2018/2019:

- Coordenadores gerais: Manuel Firmino e Sara Ferreira
- Coordenadores do curso MIEIC (Informática): José Magalhães Cruz

Equipa 1MIEIC06_03:

- Supervisor: Jorge G. Barbosa
- Monitora: Julieta Frade

Autores:

- Diogo Fernandes
(up201806250@fe.up.pt)
- Diogo Peixoto
(up201806834@fe.up.pt)
- Eduardo Brito
(up201806271@fe.up.pt)
- Hugo Guimarães
(up201806490@fe.up.pt)
- Paulo Ribeiro
(up201806505@fe.up.pt)
- Pedro Ferreira
(up201806506@fe.up.pt)

Resumo

No presente relatório foi tornado explícito o conceito de *Interação Pessoa-Máquina*, sobre a forma de *Interfaces Multimodais*. Esta relação foi analisada desde os seus primórdios até ao presente, sendo realçados os aspetos mais importantes da sua evolução, convergindo na fusão de tecnologias e no aparecimento de sistemas verdadeiramente multimodais.

Foram transmitidos os principais desafios e cuidados necessários à construção de uma *Interface Multimodal*, sendo clarificadas as suas inúmeras vantagens e desvantagens. Destacou-se, também, a importância desta temática no futuro, revelando-se as principais consequências e contrapartidas sociais da integração de determinadas *Interfaces*.

Por fim, idealizou-se um produto que tenha por base o tema deste relatório, pondo, assim, em prática os conceitos que foram analisados e compreendidos neste trabalho.

Palavras-chave

Interface Multimodal; Interação Pessoa-Máquina; Sentidos; Multimodalidade; Sistemas digitais; Tecnologia Multimodal; Comunicação; Fusão sensorial.

Agradecimentos

A realização deste relatório apenas foi possível com o contributo de várias entidades, que nos auxiliaram e se disponibilizaram para melhorar a qualidade do nosso trabalho, não só a nível da apresentação, como também a nível do conteúdo.

Gostaríamos de deixar o nosso apreço às muito úteis informações técnicas que nos foram apresentadas nas palestras da primeira semana do ano letivo, deixando um especial obrigado ao Professor Doutor Manuel Firmino, pelas importantes informações sobre a temática “Técnicas de apresentação”, às doutoras Cristina Lopes e Teresa Ramos e ao Professor Doutor Carlos Oliveira, pelas informações transmitidas sobre os temas “Relatórios em engenharia” e “Comunicação visual”, respetivamente.

Por último, agradecemos as importantes ideias dadas pelo nosso supervisor Jorge Barbosa, que se mostrou sempre disponível, e à nossa monitora, Julieta Frade, que esteve sempre presente, para ajudar e melhorar a qualidade do nosso trabalho, bem como pelos conhecimentos transmitidos durante a unidade curricular do Projeto FEUP.

Índice

Lista de figuras	5
1. Introdução	6
2. Definição e Conceitos	8
3. Linha Temporal	10
3.1. Primórdios	10
3.2. Evolução Histórica	12
3.3. Fusão Tecnológica	16
3.4. Presente	18
4. Desafios e Cuidados	21
4.1. Pontos Fundamentais – Construção de uma <i>Interface Multimodal</i>	22
4.2. Vantagens e Desvantagens	27
5. Futuro	30
5.1. Idealização de um Produto	32
6. Conclusão	34
7. Referências bibliográficas	35

Lista de figuras

- [Fig. 1] Nuvem de Conceitos sobre *Interface Multimodal*
- [Fig. 2] Arma usada na antiguidade
- [Fig. 3] Protótipos de um sistema *Multimodal* – cinema imersivo
- [Fig. 4] “ENIAC”
- [Fig. 5] “HMD” de Comeau e Bryan
- [Fig. 6] Desenho do protótipo do “Sensorama”
- [Fig. 7] “Put-That-There”
- [Fig. 8] Vários smartphones
- [Fig. 9] “Siri”
- [Fig. 10] “Eversight’s Raptor smartglasses”
- [Fig. 11] Exemplo de utilização do “AutoDraw”
- [Fig. 12] “Objectifier”
- [Fig. 13] Perspetiva VR
- [Fig. 14] “Orion SDK”
- [Fig. 15] Dispositivo VR “Leap Motion”
- [Fig. 16] Sistema “EyeTracking”
- [Fig. 17] “Tap Strap”
- [Fig. 18] Funcionalidade de Teclado Virtual
- [Fig. 19] Micro projetor num smartwatch

1. Introdução

A *Interface Multimodal*, na sua vertente relacional *Pessoa-Máquina*, é o mote principal deste relatório. Conjugando a informação mais relevante sobre as principais técnicas, aplicações e algoritmos, esta investigação visa, primeiramente, estabelecer as bases conceituais da matéria abordada e, conseqüentemente, apresentar e avaliar o acervo de desafios que estão presentes, ao longo do processo de criação e dispersão desta tecnologia.

Como tópico inicial, é definida *Interface Multimodal* como uma interligação entre sistemas distintos, que atuam em áreas específicas do panorama sensorial humano e através dos quais se efetuam trocas de informação. Esta interpenetração, baseada na congregação de pontos específicos de várias tecnologias, permite uma melhor relação entre o mundo digital e o utilizador, isto porque apela aos seus diferentes sentidos e a uma comunicação mais livre, mais natural e, sobretudo, mais abrangente. Servindo, principalmente, para conectar o interveniente ao sistema de entrada e saída das diferentes modalidades comunicativas, a *Interface Multimodal* tem vindo a ganhar considerável interesse, facto que resulta, também, da integração da informática nas ações e comportamentos naturais do dia-a-dia. É de prever que esta tendência continue, à medida que aparecem novos desafios, problemas e necessidades, como, por exemplo: gerir a grande sobrecarga de dados e de informação atual, melhorar o desempenho de sistemas baseados em reconhecimento pessoal, obter uma maior sensação de imersão em ambientes de realidade virtual e muitas outras metas desafiantes que a revolução tecnológica faz emergir, todos os dias.

Este relatório, desenvolvido no âmbito do Projeto FEUP 2018, tem, assim, como principal objetivo alargar a rede de noções e o conhecimento científico acerca deste assunto, sendo destinado, sobretudo, à comunidade de engenharia. [1]

2. Definição e Conceitos

Para observarem e compreenderem o mundo que os rodeia, os humanos usam os vários sentidos disponíveis, servindo cada um destes como meio comunicacional com qualquer outra entidade, ou sistema, seja o mundo exterior, seja outro ser humano, ou, simplesmente, um recetor com o qual se quer estabelecer comunicação. Esta troca de informação, em que se usa vários polos distintos da rede sensorial humana, ganha a designação de *Multimodal*, pela referida coexistência de várias modalidades comunicativas.

À medida que as máquinas foram sendo progressivamente introduzidas, aquando da grande revolução industrial e tecnológica, o melhoramento da eficácia comunicativa e relacional entre os utilizadores e estes sistemas tornou-se um dos principais focos de desenvolvimento, surgindo, então, o conceito de *Interface* - genericamente definido como um elemento que estabelece uma interdependência entre duas entidades. Quando este elemento é usado para auxiliar na *Interação Pessoa-Máquina* (Computador), ganha o nome de *Interface do Utilizador*. Uma evolução para este paradigma, criando uma *Interação* mais natural com quem o usa, é o estabelecimento de uma *Interação Multimodal*.

Essa *Interação* contempla uma maior gama de dispositivos de entrada e saída, sendo comumente associada à utilização de computadores, e, assim, ganha o nome de *Interface Multimodal*.



Figura 1- Nuvem de Conceitos sobre "Interface Multimodal"

Explicitamente, uma *Interface Multimodal* processa dois ou mais modos de entrada do utilizador - como fala, caneta, toque, gestos, olhar, movimentos da cabeça e do corpo, entre outros... - de maneira coordenada com a saída do sistema. São uma nova classe de *Interfaces* que visam reconhecer formas naturais de linguagem, ou comportamento humano, e que incorporam essas várias tecnologias baseadas no reconhecimento sensorial.

Este tipo de *Interface* tem por base o uso de diversos sentidos, conjugando, como já foi referido, variações comunicacionais. O objetivo é criar um facilitismo na troca de informação entre os utilizadores e a tecnologia usada, permitindo, portanto, uma maior aproximação dos mesmos. [2]

3. Linha Temporal

Após estarem estabelecidos os conceitos fundamentais de *Interface Multimodal*, segue-se o relato da sua origem e evolução. É de extrema importância compreender, não só a necessidade de evolução por parte dos humanos, como também a maneira como a *Interface* foi evoluindo, desde os seus primórdios até aos dias de hoje.

3.1 Primórdios

O progresso tecnológico não contempla apenas o momento presente da raça humana, mas sim toda a grande evolução a que a humanidade se submeteu, desde o seu estado mais primitivo de desenvolvimento. O salto histórico dá-se a partir do surgimento dos primeiros hominídeos capazes de desenvolver “processos de reconhecimento de padrões, classificação em categorias e interpretação ordenada do mundo.” Processos esses que são, simultaneamente, a base de todos os tipos de comunicação e, também, um mecanismo de transmissão e compreensão da informação. Assim que isto se tornou bem definido, a espécie passou a beneficiar das experiências baseadas na *Interação Multimodal* com o mundo em redor. O desenvolvimento humano nos distintos campos foi fundamental para a natureza aditiva do conhecimento e para a formação e evolução física, biológica e tecnológica. [3]

À medida que se avança no tempo, surgem as primeiras aplicações e as evidências da necessidade de transformar o meio e, sobretudo, de fomentar um processo criacionista, em prol do facilitismo interacional. É disso exemplo o aparecimento dos sistemas e das tecnologias, ainda modais ou bimodais, mais primitivas: as primeiras línguas e dialetos; a criação da roda e dos primeiros mecanismos com engrenagens,

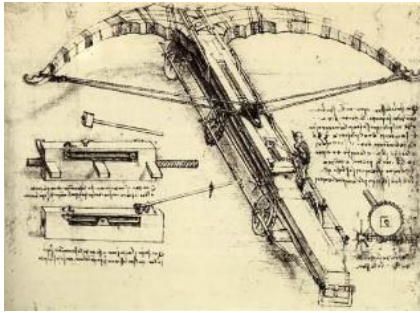


Figura 2 – Arma usada na antiguidade [4]

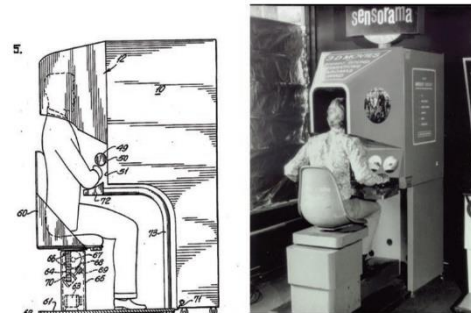


Figura 3 – Primeiros protótipos de um sistema multimodal – cinema imersivo [5]

instrumentos auxiliares usados para inúmeros fins tecnológicos; as armas e máquinas de guerra da antiguidade; o fabrico de espelhos, de faróis e o surgimento das primeiras lentes oculares, que levaram à criação de óculos e telescópios; as máquinas a vapor e, posteriormente, os motores de combustão usados em fábricas, máquinas industriais e automóveis; o desenvolvimento do telefone, como meio de comunicação à distância; os computadores; a televisão; o cinema imersivo; as consolas de jogos, e muitas outras invenções que se congregaram e convergiram na criação de máquinas e tecnologias verdadeiramente interativas e totalmente *Multimodais*, capazes de fazer valer a variedade sensorial, auxiliando e facilitando a relação com o seu criador – o ser humano – e deste com o meio. Hoje em dia, atribui-se a noção de *Interface* e *Interação Multimodal* aos dispositivos tecnológicos, sobretudo, computadores, máquinas computacionais e seus derivados.

3.2 Evolução Histórica

Na construção e desenvolvimento das primeiras tecnologias e máquinas, o manuseamento e a utilização dos aparelhos eram tarefas complexas e demoradas, sendo que apenas os seus criadores sabiam como operá-las. Contudo, ao longo dos anos, o avanço da tecnologia e a sua crescente importância e influência nas vidas das pessoas obrigou ao desenvolvimento de tecnologias que permitissem maior conforto e facilidade de acesso. Este conceito aplicou-se não só nas áreas comerciais, como também nas áreas de lazer. Por exemplo, o primeiro computador do mundo, denominado “ENIAC”, é completamente diferente das concepções presentes.

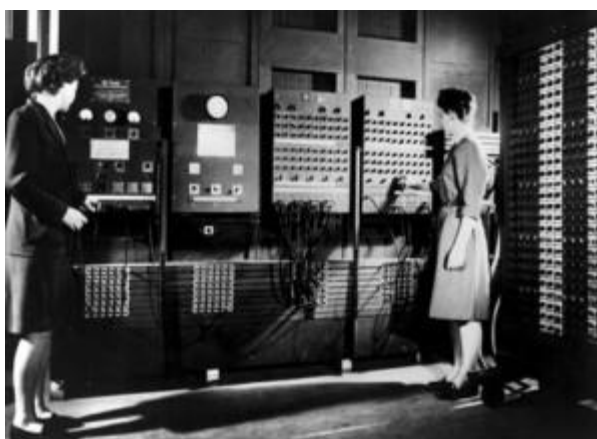


Figura 4 – “ENIAC” [6]

Para além de ter um conceito bastante díspar da atualidade, controlar este computador era um processo demoroso e complexo. No entanto, com o passar do tempo, os sistemas foram atualizados e aperfeiçoados, até chegar às tecnologias de hoje, que permitem um maior facilitismo e tornam a sua utilização mais intuitiva.

O primeiro desenvolvimento na matéria *Interface Pessoa-Máquina* foi da autoria de Comeau e Bryan, em parceria com a empresa Philco. Essa nova invenção ganhou o nome de “Head-Mounted Display” (HMD), tendo sido um protótipo de um capacete com monitores e sensores de deteção de

movimento ligados a um par de câmaras remotas, criado em 1958. Os sensores comandavam as câmaras, seguindo os movimentos da cabeça, dando ao utilizador a sensação de presença. [6]



Figura 5 – “HMD” de Comeau e Bryan [6]

Em 1962, Morton Heilig, desenvolveu um simulador denominado “Sensorama”, onde o utilizador experienciava uma simulação de viagem num veículo de duas rodas. Para tal, eram utilizadas as imagens 3D, vibrações e simuladores de vento e aromas. Esta máquina é vista como o primeiro dispositivo de realidade virtual, pois foi uma das primeiras invenções que conseguiu unir os diversos sentidos humanos de maneira eficiente. [7]

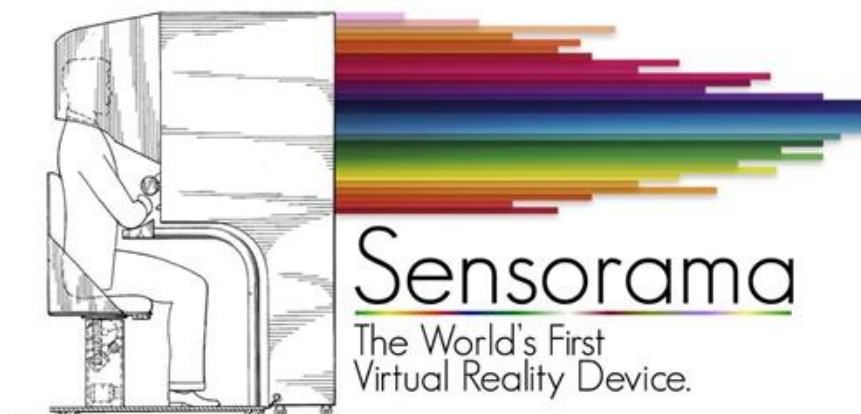


Figura 6 – Desenho do protótipo do “Sensorama” [7]

Mais tarde, no ano de 1968, Ivan Sutherland criou o primeiro sistema “Head-Mounted Three Dimensional Display”, ou seja, um capacete 3D,

baseando-se no protótipo criado em 1958, que ficou conhecido como “The Sword of Damocles”. Este capacete tinha como função conciliar os sentidos das pessoas, dando a sensação de que o que se via era real. Pelos seus esforços e criações, Sutherland ficou conhecido como o precursor da realidade virtual. Este tipo de *Interface* facilitou e melhorou trabalhos importantes relacionados com a Força Aérea Americana e a NASA.

A nível da computação, é de notar que houve uma evolução ao longo dos tempos. Começou por se utilizar algo mais simples, com a introdução de dados a partir do teclado e do rato. Na elaboração destas *Interfaces Multimodais* iniciais, tinha-se como objetivo fornecer um processamento da linguagem natural mais rico de modo a que os utilizadores facilmente manipulassem a máquina e extraíssem a informação necessária.

Por volta dos finais da década de 80, começou a introduzir-se o *spoken input*, de modo a substituir a introdução de texto a partir do teclado. É exemplo disto a interface “Put-That-There” (1980) de Richard A. Bolt. Ao manusear esta interface, o utilizador fornece as ordens através da voz e com o dedo, ou seja, a partir de gestos, aponta para o ecrã, para indicar a localização onde a ordem deveria ser executada. [8]

Após esse avanço nos sistemas modais, deu-se início à descoberta de interfaces bimodais com duas formas de introdução de dados paralelas que fornecem uma informação semântica mais rica.



Figura 7 – “Put-That-There” [8]

Servem como exemplos as interfaces de *speech* e *pen input* ou de *speech* e *lip movements input*, sendo que ambas excluem o uso do teclado e do rato. A partir daí, as interfaces foram sendo desenvolvidas cada vez mais até se tentar condensar a maior variedade de *inputs* possível, desde reconhecimentos de voz a interpretação de expressões faciais.

É também importante referir que a evolução destas interfaces teve como objetivo facilitar o uso móvel de diversos produtos. Por exemplo, a introdução de sistemas que envolvem *speech*, *pen* ou *touch input* são boas modalidades para tarefas móveis, pois, ao serem combinadas, permitem que o utilizador mude facilmente de uma para a outra, consoante a situação.

Conclui-se, assim, que as interfaces multimodais foram evoluindo ao longo do seu percurso consoante a necessidade de melhorar a acessibilidade, relativamente à ligação entre o Homem e a Máquina. [9]

3.3 Fusão Tecnológica

A fusão tecnológica corresponde à congregação das mais variadas *Interfaces*, de modo a criar um dispositivo que permita, verdadeiramente, uma *Interação Multimodal* entre o utilizador e a máquina. Esta convergência é vulgar perante aparelhos que cumpram semelhantes funções e invulgar em tecnologias avançadas de alto teor de complexidade. A temática exposta está intimamente ligada à contínua evolução tecnológica das *Interfaces Multimodais*, dado que esta evolução tem preconizado a criação de dispositivos que consigam fazer por si só aquilo que antigamente seria feito por dispositivos distintos.

Um bom exemplo para a fusão tecnológica é o telemóvel. Inicialmente, este dispositivo foi criado com o simples intuito de realizar chamadas telefónicas. Porém, nos dias de hoje, estes dispositivos eletrónicos estão equipados com sistemas operativos (como por exemplo, iOS ou Android) que contêm *Interfaces Naturais* e *Multimodais* que permitem o substancial aumento da quantidade de funcionalidades do aparelho. Assim sendo, o telemóvel é visto como um verdadeiro “supercomputador” que foi evoluindo. É equipado com um *touch-screen* e um sistema de comunicação (microfone e colunas), podendo ainda existir um sistema de reconhecimento de impressões digitais, um software de interação conversacional, inteligência artificial, e outras funcionalidades e funções audiovisuais que estimulam os vários sentidos do seu utilizador. [10][11]

Nos dias de hoje, também os veículos modernos possuem sistemas de *Interfaces* audiovisuais que permitem ao condutor realizar chamadas telefónicas, escolher estações de rádio e orientar-se na sua viagem (GPS) sem desviar a sua atenção da estrada com elevada frequência, ou seja, foi integrado no veículo um sistema eletrónico *Multimodal*, controlado

tanto por voz, como por simples toques num *display*, que contribui simultaneamente para um maior conforto do condutor e para a segurança rodoviária. [12]



Figura 8 - Vários smartphones [13]



Figura 9 – “Siri” [14]

Deste modo, pode-se concluir que a fusão tecnológica é uma temática de elevada importância, visto que a melhoria da qualidade de vida da sociedade é alcançada não só pela descoberta de novas tecnologias, mas também pela sua correta integração nas existentes, originando produtos polivalentes de melhor qualidade.

3.4 Presente

Recentemente, foi surgindo uma nova geração de aplicações com um grau de interatividade superior às do passado, que eram baseadas apenas em dispositivos de indicação e seleção de informação, como o rato. Este sistema mais simples predominou até inícios do século XXI e começa agora a ser substituído. Além disso, nos modelos antigos, praticamente toda a *Interação* se concentrava num único sentido, o da visão, sobrecarregando-o, tendo surgido a necessidade da criação de novas séries que permitissem o uso de diversos sistemas de entrada de dados, ao mesmo tempo. O surgimento de novos modelos também nos desprende do conceito de computador de mesa, sendo agora possível realizar as mesmas tarefas de forma simples e prática com o uso de um smartphone, como foi anteriormente observado. Todos esses desafios, que vão surgindo para melhorar o desempenho e facilitar as tarefas, exigem mais conhecimento por parte dos criadores de sistemas interativos, havendo uma contínua evolução no sentido de satisfazer ao máximo todas as necessidades impostas na atualidade.

Um exemplo disso é o sistema “EverySight's Raptor *smart-glasses*”, semelhante a um painel de um automóvel, sendo especialmente desenvolvido para ciclistas. Todo o *display* apresenta uma cor verde e permite obter dados como a velocidade e a distância percorrida. Também apresenta um sensor de frequência cardíaca e um sistema que permite criar os próprios percursos que se pretende seguir.



Figura 10 - Utilização em primeira pessoa dos “EverySight's Raptor smartglasses” [16]

Para além disso, este dispositivo apresenta uma configuração onde, se a frequência cardíaca subir a um nível máximo, as cores dos seus dados mudarão de verde para laranja, ou para vermelho, dependendo da diferença dos dados em relação ao estado normal. O mesmo acontece se o utilizador optar por colocar uma meta para o seu exercício, sendo neste caso a cor correspondente à performance do utilizador. Como controladores, contém uma câmara e comandos de voz que se ativam ao ser dita a frase "Go Every sight", podendo ser pedida a realização de um vídeo, ou fotografia. Permite também a sincronização de filmagens com o smartphone por meio de um aplicativo complementar que possibilita a sincronização de músicas com o aparelho.

Uma das grandes inovações destes óculos é a integração de um sistema de localização pessoal. Sempre que o utilizador se desvia da rota, a seta do seu movimento fica vermelha, indicador de que a direção está errada. O controlo da cabeça é também notório e importante. Quando a cabeça é virada para um lado pode ser visto, no mapa, um pequeno cone a indicar a direção do local observado. Este dispositivo permite ainda a visualização de *feeds* e notícias, com um simples movimento leve da cabeça para cima, ou para baixo. [16]

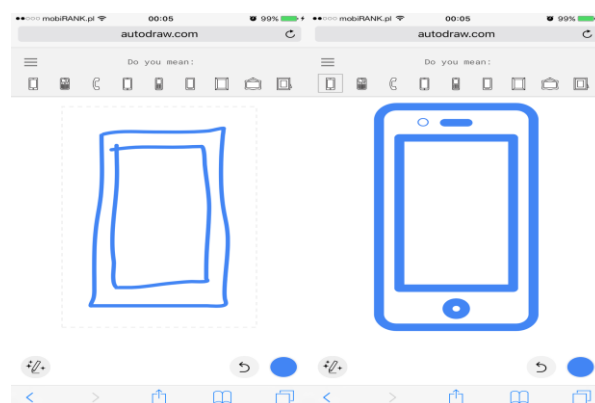


Figura 11 - Exemplo de utilização do "AutoDraw" [17]

Outro exemplo é o "AutoDraw", que usa *machine learning* para completar desenhos (rabiscos, realizados ao acaso) que os utilizadores fazem no site. Apesar de parecer algo sem utilidade, este mecanismo é capaz de

interpretar textos escritos à mão, voz, gestos, humor, traços de identidade, e interagir com o utilizador. [17]

Um outro projeto, que é da autoria da Google, concede-nos o poder de treinar objetos do nosso dia a dia, de modo a que estes respondam aos comportamentos dos utilizadores, isto é, fornece a experiência de treinar um objeto com inteligência artificial. Interagir com o “Objectifier” assemelha-se ao treino de um animal de estimação, uma vez que apenas se lhe ensina o que realmente nos importa, ou seja, o modo como desejamos que ele realize determinada função.



Figura 12 – “Objectifier” [18]

É através da *Visão Computacional* (processo de moldagem e réplica da visão humana usando software e hardware) e *Redes Neurais* (modelos computacionais inspirados no sistema nervoso central de um animal, em particular o cérebro, que são capazes de realizar a aprendizagem de uma máquina, bem como o reconhecimento de padrões), que certos comportamentos são associados a uma ordem dada ao “Objectifier” pelo utilizador. Por exemplo, poderíamos querer ligar o rádio com o nosso passo de dança favorito. Para isso, precisaríamos de sincronizar o rádio com o “Objectifier”, e treiná-lo de modo a que, sempre que fizéssemos o passo de dança, ele o ligasse.

É com estes exemplos e deste modo que os utilizadores podem experimentar novas maneiras interativas de controlar objetos, estabelecendo uma relação criativa e *multimodal* com a tecnologia. [18]

4. Desafios e Cuidados

A frase de Larry Tesler, cientista da Apple, define bem a importância da experiência de utilização que o sistema proporciona ao seu utilizador - “A *Interface* é a única maneira que o utilizador tem de avaliar o sistema” [19]. A este não lhe interessa a linguagem de programação na qual o sistema foi desenvolvido, ou a metodologia empregue na sua conceção e desenvolvimento. O conceito vai além da estética da tela (disposição de menus, cores, etc.), uma vez que se estende a questões como o tempo de resposta, o grau de dificuldade de utilização, a rapidez no desempenho de tarefas, o nível de erros cometidos por utilizadores durante o uso, a facilidade na aprendizagem, o cansaço resultante do uso prolongado, a acomodação de utilizadores com necessidades especiais, entre outros fatores.

Assim, no planeamento de *Interfaces Multimodais*, há que ter em mente uma série de cuidados, de modo a que o produto final seja aplicado na situação mais adequada. Por essa razão, o processo de criação traz imensos desafios aos criadores, pois têm de efetuar um estudo sobre o ambiente, o tipo de utilizadores e a conveniência, entre outros parâmetros, de modo a produzir uma solução funcional. Para tal, é essencial entender não apenas como o computador opera, mas também como o ser humano “funciona”. Desse modo, cientistas da computação juntam-se a psicólogos, dando início a atividades de pesquisa na área de *Interação Pessoa-Máquina*.

Seguindo, então, esse raciocínio, deve-se ter em conta que é necessário escolher as modalidades que, consoante a ocasião, nos permitem recolher a maior quantidade de todo o tipo de informação existente, através da combinação de modos de input, ou de alternância entre

os mesmos, daí que as *Interfaces Multimodais* são a melhor solução, por permitirem uma maior flexibilidade no uso desses modos. [20]

4.1 Pontos Fundamentais – Construção de uma Interface

Para que a *Interface Multimodal* satisfaça todas as necessidades impostas pelo Homem e seja considerada uma boa *Interface*, tem que seguir um conjunto de parâmetros e regras [21]:

- **Conveniência segundo a ocasião**

É de extrema importância que se reconheça a situação onde a *Interface Multimodal* vai ser aplicada. Um estudo prévio da mesma permite aos criadores inventar uma solução eficaz a partir da escolha acertada das modalidades a ser aplicadas. Por exemplo, num caso onde exista muito movimento (como quando andamos de carro), é mais viável que as interações com a tecnologia sejam feitas por voz. Já noutro caso, onde o ruído é a grande perturbação, é mais eficaz aplicar uma *Interface* com modalidade visual.

- **Confiança no sistema**

O utilizador precisa de confiar no sistema. Um sistema no qual ele não acredita está destinado ao fracasso, sendo a confiança algo frágil, pois basta uma falha e esta fica comprometida por muito tempo. Os comandos devem funcionar como especificados e conforme descritos nos manuais do utilizador. Os dados e valores exibidos em telas e relatórios devem corresponder à realidade, àquilo que está armazenado em bases de dados.

- **Disponibilidade do sistema**

Outro fator importante é a disponibilidade. Por um lado, o sistema precisa de estar sempre disponível. Nos dias de hoje, não se espera nada menos do que sistemas que operem vinte e quatro horas por dia, sete dias por semana. Os sistemas de computação não devem parar, nem para fazer “backup”. Devem ser concebidas estratégias de atualizações para deixar o sistema o mínimo de tempo indisponível. Posto isto, as *Interfaces Multimodais* permitem que não haja apenas um único caminho de execução/utilização, mas sim várias alternativas, de maneira a que, mesmo que um dos modos falhe, a operação possa ser feita por outra via. Por outro lado, havendo uma grande complexidade nesses sistemas, uma só modalidade não é suficiente para que todos tenham acesso a certa tecnologia. Por essa razão, a *Interface Multimodal* resume-se também em permitir que todos os utilizadores consigam interagir eficazmente em todos os casos, sendo, por isso, importante ter em conta as idades, os níveis de habilidades, a sua língua-mãe, os estilos cognitivos e as incapacidades/deficiências temporárias ou permanentes. Por exemplo, quando se projeta a *Interface* tendo em conta as necessidades de pessoas especiais, oferece-se a oportunidade de se tornarem produtivas e úteis à sociedade. *Interfaces* acessíveis não envolvem apenas pessoas com alguma deficiência sensorial, como também diminuem as limitações de *Interação* trazidas pela idade, isto é, dos utilizadores idosos.

- **Segurança do sistema**

O conceito de segurança está ligado ao de disponibilidade. Um sistema com pouca segurança tem muito mais probabilidade de se tornar indisponível. No entanto, a segurança não está

somente ligada a essa fração do conceito. Muitas vezes a informação está correta, mas não deve ser acedida por determinado utilizador – um funcionário, por exemplo, não deve ter acesso aos salários dos restantes funcionários da empresa. Outro aspeto a considerar é o de que a segurança deve espelhar-se igualmente no bem-estar do utilizador, sendo necessário criar uma *Interface* que não provoque uso excessivo, ou dano a qualquer nível na pessoa, principalmente quando se tratar de um uso de longa duração.

- **Padronização da Interface**

A padronização da *Interface*, dos objetos (normalmente visuais) com os quais os utilizadores interagem é um fator decisivo para o sucesso de um sistema interativo. Pequenas diferenças entre aplicações irritam o utilizador, aumentam o tempo de aprendizagem e a incidência de erros. Assim, a padronização refere-se a características comuns entre as *Interfaces*, tanto entre programas de uma mesma aplicação, como entre diferentes aplicações. Por outro lado, este aspeto também se refere à facilidade de utilização das *Interfaces*, normalmente associadas à proximidade do seu uso com a realidade. Por exemplo, quando tocamos numa fotografia de um “tablet” com dois dedos e afastamos um do outro, criamos mais espaço entre os dois dedos, consequentemente aumentando o tamanho da fotografia. O gesto, pela semelhança que teria com aquele que faríamos no mundo real, permite uma experiência sensorial mais rica e intuitiva para o utilizador da aplicação. Essa experiência aumenta o grau de satisfação deste, diminui o esforço de aprendizagem e aumenta o nível de retenção, isto é, lembrar-se-á com mais facilidade de como utilizar.

- **Consistência da Interface**

O conceito de consistência está muito próximo do de padronização: refere-se a sequências de ações comuns, termos, unidades e cores utilizadas numa aplicação. A consistência é, por si só, um fator determinante de sucesso, ou fracasso, de um projeto. *Interfaces* inconsistentes causam os mesmos males discutidos no tópico anterior. O utilizador também espera que a consistência seja mantida entre as diversas versões da aplicação. Mudar o modo de se realizar uma tarefa de uma versão para outra é algo bastante perigoso de se fazer. O conjunto de ações necessárias para o utilizador atingir o mesmo objetivo utilizando a versão nova deve ser menor, as ações devem ser mais simples e intuitivas, caso contrário essa quebra de continuidade pode levar à frustração e perda de produtividade, além de poder trazer a perda de mercado, para a empresa que o construiu. A consistência também pode ser observada entre as ações do mundo real e o digital. Quanto mais consistente, em termos de ação do utilizador e do sistema, com as reações que ele esperaria no mundo real, melhor.

- **Portabilidade da Interface**

Já este tópico diz respeito ao potencial de se converter dados e compartilhar *Interfaces* entre vários sistemas e equipamentos de diferentes tipos. Hoje em dia, essa característica está bem valorizada devido às tecnologias de computação que surgiram ao longo do tempo – a informação está disponível aos utilizadores, onde quer que eles estejam (desde que tenham acesso à internet) e em vários dispositivos diferentes (desktops, tablets, smartphones, etc.).

- **Integração da Interface**

O conceito de integração está ligado à transferência automática de dados entre sistemas distintos. Por exemplo, o sistema de compras pode estar integrado ao de pagamento fazendo com que a pessoa que vai fazer o pagamento já saiba qual é a forma de pagamento negociada, sem ter que ir ao seu colega de trabalho e perguntar. [22]

4.2 Vantagens e Desvantagens

Como já foi referido, as *Interfaces Multimodais* visam tornar a comunicação entre o homem e a máquina mais prática e cómoda. Para tal, é necessária uma maior imersão no universo computacional, para que a comunicação com estas máquinas seja mais simples e fluída.

Para a maioria das pessoas, a utilização dos sentidos para controlar um computador pode parecer insignificante. No entanto, isto traz muitas vantagens, mas também tem algumas desvantagens associadas. Apesar disso, os utilizadores preferem este tipo de *Interface* por oferecer mais alternativas e opções no momento de operar os produtos. [23]

4.2.1 Vantagens

- Ajudar e acomodar as capacidades que o utilizador tem de perceber e comunicar, através do uso de *Interfaces* mais naturais, com novas maneiras de interagir.
- Integração de habilidades computacionais no mundo real, oferecendo formas mais naturais, no âmbito da *Interação Pessoa-Máquina*.
- Maior robustez devido à combinação de diferentes fontes de informação, a partir da complementação de diferentes modalidades.
- Personalização flexível, ao usar diferentes modalidades consoante as preferências do utilizador e o contexto em que se encontra.
- Facilita a *Interação Pessoa-Máquina* para pessoas com incapacidades/deficiências, de modo a que possam contribuir mais na sociedade.

- Se uma das modalidades falhar, há sempre outra via para recolher informação e permitir que a máquina continue em funcionamento. [24]

4.2.2 Desvantagens

- É conhecido que a ambiguidade é uma característica intrínseca aos atos humanos. Essa ambiguidade pode levar a máquina a não interpretar da melhor maneira aquilo que lhe é pedido, e tal situação é mais comum do que parece. Um exemplo claro disso ocorre quando se envia um comando por voz a um sistema. Este, apesar de captar aquilo que se profere, pode apresentar uma interpretação completamente diferente e executar uma tarefa que não era suposto executar.
- Uma *Interface Multimodal* tem que permitir ao utilizador uma *Interação* fluída. Para tal, é necessário que exista tecnologia de ponta, ainda dispendiosa e, por vezes, de difícil acesso, retirando-se disto a conclusão de que a construção e implementação destas *Interfaces*, que permite que os utilizadores comuniquem através dos seus sentidos, tem um elevado custo associado.
- A introdução de sistemas e tecnologias *Multimodais* inovadoras vai fazer com que estes sistemas sejam cada vez mais utilizados pelos humanos, criando-se uma maior proximidade com a máquina. No entanto, isto traz efeitos adversos, como, por exemplo, a menor comunicação e uma menor ligação entre as pessoas, o que se vai traduzir numa perda de contacto com o mundo real.

- É presumível que a introdução deste tipo de *Interfaces* computacionais provoque uma perda de empregos, visto que estas permitem que uma única pessoa consiga executar vários trabalhos simultaneamente, excluindo, assim, a necessidade de mão de obra.
- É uma *Interface* que facilmente cria um hábito no utilizador e uma certa dependência, a necessidade de ter mais do que um meio para resolver as situações importantes do dia-a-dia.

5. Futuro da Interface Multimodal

Como foi visto, hoje em dia, qualquer indivíduo possui no bolso um dispositivo que seria considerado, no início do século, um supercomputador. A evidente evolução tecnológica com a qual nos deparamos nos dias de hoje leva-nos a pensar no futuro das *Interfaces Multimodais* e no seu possível impacto na sociedade. Nas últimas décadas, tem-se estado a utilizar o “rato” como principal método de *Interação* com a tecnologia, e ainda que o *touch-screen* tenha surgido como um novo modelo de ligação *Pessoa-Máquina*, depende também do toque. [25]

“Leap Motion” é uma empresa que visa revolucionar o mundo tecnológico através de um dispositivo de hardware que utiliza sensores de movimentos executados pela mão e dedos do utilizador como modo de entrada de informação, sem qualquer necessidade de contacto físico (como um clique de um rato, ou um *touch-pad*). [26]

Outro aparelho, denominado “Orion SDK”, atua em conjunto com um sistema de realidade virtual, ou seja, uma tecnologia de *Interface* que permite ao utilizador interagir em tempo real com o mundo à sua volta numa amplitude de 360°, através de recursos gráficos a 3 dimensões, formando um ambiente imaginário que se assemelha ao real através da estimulação de uma grande diversidade de sentidos: o visual, o auditivo e o tato. Esta sinestesia cria uma melhor e maior evolução na *Interação Pessoa-Máquina*. [27] [28]

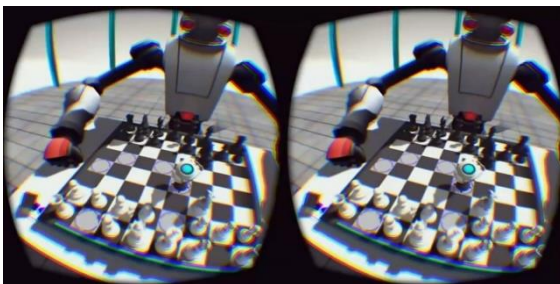


Figura 13 - Perspetiva VR [25]



Figura 14 - Orion SDK [27]

O sistema de “EyeTracking” está a ser explorado por várias empresas, nomeadamente, a “Google”. Este dispositivo deteta o movimento dos olhos do utilizador, tendo como objetivo permitir realizar com os olhos aquilo que se pode atualmente fazer com o rato. Uma das suas aplicabilidades é a sua união com um aparelho de realidade virtual, tendo como objetivo ser um mecanismo que facilite a *Interação* com a máquina. [29]



Figura 15 - Dispositivo VR "Leap Motion" [26]



Figura 16 - Sistema “EyeTracking” [29]

O desenvolvimento destas novas tecnologias é uma prova da crescente influência das *Interfaces Multimodais* no futuro, dado que abrem um novo horizonte de possibilidades de criação de novas tecnologias que possam ser utilizadas, quer no âmbito do lazer, quer a nível profissional, sendo em ambas ocasiões um potencial criador de emprego e um mecanismo que potencializa o aumento da qualidade de vida da nossa sociedade. [30] [31]

5.1 Idealização de um Produto

Para a idealização do nosso produto, baseámo-nos nalguns já existentes e tentámos criar uma versão que reunisse todas as funcionalidades destes, concebendo algo mais sofisticado e preferencialmente multimodal.

A ideia a propor é uma fita, esteticamente construída para encaixar nos dedos do utilizador, tal como faz o produto “Tap Strap”. [32]



Figura 17 - "Tap Strap" [32]



Figura 18 – Funcionalidade de Teclado Virtual [32]

Congregaria as funcionalidades desta invenção com outras pensadas por nós: teria incorporado um sistema de reconhecimento de gestos que permitisse não só o controlo de um teclado virtual, funcionalidade da “Tap Strap”, como também o controlo de sistemas e equipamentos sincronizados com o produto, por exemplo, o sistema de luzes de casa, *smart TVs*, *media players*, eletrodomésticos, o automóvel, entre muitos outros. Poderia conjugar, da mesma maneira, um sistema de reconhecimento por voz, respondendo a comandos e executando as respetivas tarefas, estando tudo sincronizado com o smartphone do utilizador (imaginar-se-ia a existência de um microfone e de um altifalante incorporados que permitissem a realização de chamadas telefónicas, por exemplo).



Figura 19 - Micro projetor num smartwatch [33]

Outra funcionalidade, já pensada para outro tipo de equipamentos, como os *smartwatches*, seria incorporar um micro projetor a laser, que projetasse, através de um dos anéis da fita, em qualquer superfície, a imagem das horas, por exemplo, ou, então, de uma notificação que o utilizador receba no telemóvel. [33] Além disso, juntar-se-ia a funcionalidade de vibração, que funcionaria da mesma maneira que a implementada nos smartphones.

6. Conclusão

Apesar de terem sido apresentadas algumas desvantagens, são bastantes os benefícios que a evolução e o desenvolvimento do ramo da *Interface Multimodal* podem trazer, uma vez que facilitam a interação com os sistemas do quotidiano, tornando esta ligação mais prática e simples. Se neste momento existem já projetos com esta *interface* implementada e bastante desenvolvida, capazes de colocar noutra patamar os avanços tecnológicos presentes, certamente que, no futuro, será possível a criação de um sistema totalmente completo, onde este tipo de *Interface* estará implementado com todo o seu potencial.

7. Referências Bibliográficas

- [1] Santos, Marcelo Alves dos. 2013. “*Interface Multimodal de Interação humano-computador em sistema de recuperação de informação baseado em voz e texto em português*”, Faculdade de Ciência da Informação, Universidade de Brasília. Acedido a 15 de outubro de 2018. <http://repositorio.unb.br/handle/10482/14843>
- [2] Júnior, Valter dos Reis Inácio. 2007. “Um Framework para Desenvolvimento de *Interfaces Multimodais* em Aplicações de Computação Ubíqua”, Escola de Engenharia de São Carlos, Universidade de São Paulo. Acedido a 15 de outubro de 2018. <http://www.teses.usp.br/teses/disponiveis/55/55134/tde-11122007-105843/publico/DissertacaoMScValterFinal20070216.pdf>
- [3] Rebelo, Mauro. “Quando o homem começou a falar?” Science Blogs (blog), 1 de maio de 2007. Acedido a 16 de outubro de 2018. <http://scienceblogs.com.br/vqeb/2007/05/quando-o-homem-comecou-a-falar/>
- [4] Da Vinci, Leonardo. 1500. “Crossbow, Europe, 1200 - wood and iron”, Besta (arma). Acedido a 16 de outubro de 2018. http://www.sandia.gov/tp/SAFE_RAM/CRSBW.HTM
- [5] Heilig, Morton. 1961. “Sensorama”, Illustration of Morton Heilig's Sensorama device, precursor to later virtual reality systems. Acedido a 16 de outubro de 2018. Figure 5 of U.S. Patent #3050870 <http://patft.uspto.gov/>
- [6] Cam Robinson. 2014. “History of Virtual Reality - Reality Check”. Youtube vídeo. Acedido a 14 de outubro de 2018. https://www.youtube.com/watch?time_continue=1&v=43mA_ypfwKg
- [7] Thomas DeMichele. 2016. “The Sensorama Was the First VR Device”. Fact/Myth (blog), 25 de janeiro de 2016. Acedido a 10 de outubro de 2018. <http://factmyth.com/factoids/the-sensorama-was-the-first-vr-device/>

- [8] Salter, Carson. 2012. "Put That There (Original)". *Arena*.
Acedido a 10 de outubro de 2018.
<https://www.arena.com/carson-salter/nicholas-negroponte-and-an-eames-chair>
- [9] Sears, Andrew e Jacko, Julie A. 2007. "*Multimodal Interfaces*". Em "The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications, Second Edition", 414-417. CRC Press. Acedido a 14 de outubro de 2018
https://books.google.pt/books?hl=en&lr=&id=A8TPF_O385AC&oi=fnd&pg=PA413&dq=Interface+Multimodal+evolution&ots=fmVJkOjSZh&sig=X30S6joXYmbWE2ghuvvpX07Lvs&redir_esc=y#v=onepage&q&f=false
- [10] Federico Caviggioli. 2016. "Technology fusion: Identification and analysis of the drivers of technology convergence using patent data". Acedido a 23 de outubro de 2018.
<https://www.sciencedirect.com/science/article/pii/S0166497216300293>
- [11] Ana Duarte. 2018. "A evolução do telemóvel em imagens". Acedido a 23 de outubro de 2018. <https://www.e-konomista.pt/artigo/evolucao-do-telemovel/>
- [12] Pedro Kutney. 2012. "A revolução da *Interface* homem/carro" Acedido a 23 de outubro de 2018. <http://www.automotivebusiness.com.br/noticia/15729/a-revolucao-da-Interface-homemcarro>
- [13] Julian Chokkattu. 2018. "How to choose a smartphone". Acedido a 19 de outubro de 2018. <https://www.digitaltrends.com/mobile/cell-phone-buying-guide/>
- [14] Ric Richardson. 2018. "How to use SIRI to record Audio or video". Acedido a 19 de outubro de 2018.
<https://www.google.com.sg/url?sa=i&source=images&cd=&cad=rja&uact=8&v>

[ed=2ahUKEwjdo-r50ZLeAhWHtI8KHe9yA6AQjRx6BAgBEAU&url=https%3A%2F%2Fhackernoon.com%2Fhow-to-use-siri-to-record-audio-or-video-1b46d8102f8e&psig=AOvVaw2_U_6QHlZLA8_xHip2A3_&ust=1540043146384973](https://www.youtube.com/watch?v=ed=2ahUKEwjdo-r50ZLeAhWHtI8KHe9yA6AQjRx6BAgBEAU&url=https%3A%2F%2Fhackernoon.com%2Fhow-to-use-siri-to-record-audio-or-video-1b46d8102f8e&psig=AOvVaw2_U_6QHlZLA8_xHip2A3_&ust=1540043146384973)

- [15] Apple. 2018. "Siri does more than ever. Even before you ask".
Acedido a 19 de outubro de 2018. <https://www.apple.com/siri/>
- [16] Husain Sumra. 2017. "AR meets cycling: Trying out EverySight's Raptor smartglasses". Acedido a 19 de outubro de 2018.
<https://www.wareable.com/ar/eversight-raptor-ar-smartglasses-cycling-231>
- [17] Rodrigues, Eli. 2017. "Google Autodraw e as Interfaces Multimodais".
Acedido em 19 de outubro de 2018.
<https://pt.linkedin.com/pulse/google-autodraw-e-Interfaces-Multimodais-eli-rodrigues>
- [18] Bjørn Karmann. 2018. "Objectifier Spatial Programming". Acedido a 19 de outubro de 2018.
<https://experiments.withgoogle.com/objectifier-spatial-programming>
- [19] Larry Tesler, 2015. Acedido a 9 de outubro de 2018.
https://en.wikipedia.org/wiki/Larry_Tesler
- [20] Francisco Carlos de Mattos Brito Oliveira e Fernando Antônio de Mattos Brito Oliveira. 2016. "Computação: *Interação* Humano-Computador". Acedido a 9 de outubro de 2018.
www.uece.br/computacaoead/index.php/.../doc.../2116-interacao-homem-computador

- [21] Acedido a 18 de outubro de 2018.
<https://pt.slideshare.net/janynnegomes/aula-evolucao-das-Interfaces>
- [22] Sears, Andrew e Jacko, Julie A. 19 de setembro de 2007. "The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications, Second Edition". Acedido a 14 de outubro de 2018
https://books.google.pt/books?hl=en&lr=&id=A8TPF_O385AC&oi=fnd&pg=PA413&dq=Interface+Multimodal+evolution&ots=fmVJkOjSZh&sig=X30S6joXYmbWE2ghuvvpX07Lvs&redir_esc=y#v=onepage&q&f=false
- [23] Sears, Andrew e Jacko, Julie A. 2007. "*Multimodal Interfaces*". Em "The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications, Second Edition", 417-420. CRC Press. Acedido a 16 de outubro de 2018.
https://books.google.pt/books?hl=en&lr=&id=A8TPF_O385AC&oi=fnd&pg=PA413&dq=Interface+Multimodal+evolution&ots=fmVJkOjSZh&sig=X30S6joXYmbWE2ghuvvpX07Lvs&redir_esc=y#v=onepage&q&f=false
- [24] Signer, Beat. 2016. "*Multimodal Interaction - Lecture 05 - Next Generation User Interfaces (4018166FNR)*". 25 de outubro de 2016. 16 de outubro de 2018.
<https://www.slideshare.net/signer/Multimodal-interaction-lecture-05-next-generation-user-Interfaces-4018166fnr>
- [25] Lucas Bicudo. 2016. "O futuro da *Interação* humana com a tecnologia será *Multimodal*". Acedido a 17 de outubro de 2018. <https://startse.com/noticia/o-futuro-da-interacao-humana-com-a-tecnologia-sera-Multimodal>
- [26] Leap Motion. 2018. "Reach into the future of virtual and augmented reality". Acedido a 17 de outubro de 2018.
<https://www.leapmotion.com/>
- [27] UOL. 2017. "UOL lança app de realidade virtual com vídeos 360°". Acedido a 17 de outubro de 2018.
<https://tecnologia.uol.com.br/noticias/redacao/2017/10/24/uol-lanca-app-de->

- [28] Shane Wall. 2017. "Fusão dos mundos físico e digital vai inspirar novas tecnologias". Acedido a 17 de outubro de 2018. <http://idgnow.com.br/internet/2017/03/08/artigo-fusao-dos-mundos-fisico-e-digital-vai-inspirar-novas-tecnologias/>
- [29] Lightleak. 2018. Eye tracking technology in the future". Acedido a 19 de outubro de 2018. <https://www.videoblocks.com/video/eye-tracking-technology-in-the-future-voww-o6qneilmqos28>
- [30] Ramond Wong. 2016. "VR is even more real now that you can reach out and touch things with your hands". Acedido a 19 de outubro de 2018. <https://mashable.com/2016/12/05/leap-motion-mobile-vr/?europa=true>
- [31] Lisa Eadicicco. 2016. "This Startup Gave Me a Hand in Virtual Reality". Acedido a 19 de outubro de 2018. https://www.google.com.sg/url?sa=i&source=images&cd=&cad=rja&uact=8&ved=2ahUKEwjsh9fY35LeAhUGaBoKHUfDPoQjRx6BAgBEAU&url=http%3A%2F%2Ftime.com%2F4226031%2Fleap-motion-virtual-reality-orion-hands%2F&psig=AOvVaw3nT_yL6K2o840T5fesWuFY&ust=1540046629293234
- [32] Tap Strap. 2018. Acedido a 25 de outubro de 2018. <https://www.tapwithus.com>
- [33] Miller, Paul. 2016. "Asu Cast One smartwatch with built-in projector is now a real thing you can buy". Acedido a 25 de outubro de 2018. <https://www.theverge.com/circuitbreaker/2016/5/18/11700894/asu-cast-one-smartwatch-projector-released-china>