

Biomecânica: Análise do Movimento Humano

Classificação Temporal de Ações Humanas

Dissertação apresentada para a obtenção do grau de Mestre em
Instrumentação Biomédica

Autora

Jessica Patrícia Madureira Vital

Orientadores

Doutora Fernanda Madureira Coutinho

Doutor Nuno Miguel Fonseca Ferreira

Instituto Superior de Engenharia de Coimbra

Doutor Micael Santos Couceiro

Ingeniarius, Lda.

Coimbra, novembro 2015

Dedico, com imensa honra e gratidão, a subsequente dissertação à minha família e diletos amigos, reconhecendo todo o apoio facultado no decorrer deste percurso.

À minha mãe Elisabete Madureira, à minha irmã Beatriz Vital e ao meu avô Henrique Madureira pelo amor, apoio, confiança e motivação incondicional. Que sempre me incita para a vitória dos meus desafios e para a superação das diversas dificuldades.

À minha avó Júlia que teria muito orgulho de estar presente nesta etapa da minha vida.

“A tarefa não é tanto ver aquilo que ninguém viu, mas pensar o que ninguém ainda pensou sobre aquilo que todo mundo vê.”

Arthur Schopenhauer

AGRADECIMENTOS

Ao Professor Doutor Micael Couceiro, mais do que um orientador, um amigo e uma referência de vida. Todas as palavras são poucas para descrever todo o empenho e confiança que demonstrou para comigo, bem como a dedicação que demonstrou neste trabalho.

À Professora Doutora Fernanda Coutinho, pela disponibilidade e ajuda que manifestou para acompanhar este trabalho.

Ao Professor Doutor Nuno Ferreira, pelo apoio disponibilizado ao longo do meu percurso académico.

Aos meus colegas da Ingeniarius, Samuel Pereira, André Araújo, Sérgio Louros e Gui Mota, pelo apoio incondicional e incentivo ao longo de todo o trabalho.

Ao Doutor Diego Faria, pela partilha de conhecimento e por toda a dedicação que demonstrou.

Ao Professor Doutor Gonçalo Dias, pela disponibilidade, participação e colaboração neste trabalho.

A todos os amigos que me ajudaram na aquisição dos dados, Catarina Páscoa, Ivone Leite, Tiago Pereira, Diogo Robalo e João Pais.

Ao Instituto Superior de Engenharia de Coimbra, onde fiz o meu percurso académico.

A todos os meus amigos que estiveram sempre presentes e a apoiar-me incondicionalmente.

Por último, mas não em último, aos meus pais, irmã e avô, pelo indubitável apoio que disponibilizaram ao longo de todo o meu percurso. À minha avó Júlia que, mesmo não estando presente, me deu forças durante todos os dias deste trabalho.

RESUMO

Este estudo tem como objetivo identificar atividades diárias de diferentes pessoas com recurso a métodos de classificação supervisionados. Neste sentido, começou-se por analisar várias tecnologias associadas à captura e análise do movimento humano, tais como sensores (e.g., inertial measurement unit) e câmaras de filmar (e.g., RGB, infravermelhos e time-of-flight). A revisão da literatura indica claramente que, contrariamente ao uso das câmaras de filmar, a tecnologia wearable tende a ser mais adequada para a análise cinemática de movimentos desportivos. Este tipo de tecnologia permite ainda obter uma estimativa da orientação e produção de movimento dos membros superiores e inferiores com elevado nível de precisão e exatidão, bem como imunidade a ângulos mortos, aumentando deste modo a quantidade e qualidade da informação obtida. Tendo isto presente, este trabalho apresenta uma metodologia para classificar atividades diárias do movimento humano com recurso a um fato sensorial (wearable), Ingeniarius FatoXtract. O desempenho da solução proposta é ainda comparado com a utilização de uma câmara time-of-flight, Microsoft Kinect v2. A metodologia proposta considera a integração probabilística de três classificadores: o Naïve Bayes, as Redes Neurais Artificiais e as Máquinas de Vetor de Suporte. Com vista a alcançar um desempenho superior na classificação geral do movimento, foram consideradas diversas features no domínio do tempo (e.g., velocidade) e no domínio da frequência (e.g., Transformada Rápida de Fourier), combinado com as tradicionais features geométricas (e.g., posição angular das juntas).

Realizou-se a aquisição de dados de cinco atividades comuns do dia-a-dia, realizadas por seis participantes com repetições de 20 ensaios cada, usando o FatoXtract e o Kinect v2. O conjunto de dados foi projetado para ser extremamente desafiador, uma vez que a duração das atividades varia drasticamente e algumas atividades são muito semelhantes (e.g., lavar os dentes e acenar).

Palavras-chave: reconhecimento de padrões; extração de *features*; métodos de classificação; análise do movimento humano; classificação temporal; fato inercial; Naïve Bayes; Redes Neurais Artificiais; Máquinas de Vetor de Suporte.

ABSTRACT

This study aims to classify daily activities of different people, based on supervised classification methods. In this sense, firstly it was analyzed different technologies associated with capture and analysis of human movement such as sensors (e.g., inertial measurement unit, or IMU) and cameras, including cameras RGB, infrared, and time-of-flight. The literature indicate that, contrary to the use of video cameras, wearable technology tends to be more suitable for kinematic analysis of sports movements. This type of technology allows even getting an unlimited estimate orientation and movement of production of arms and legs, as well as immunity to blind spots, thereby increasing the quantity and quality of the information obtained. Having this in mind, this paper presents how to classify daily activities of human movement using a wearable motion capture full - body suit, denoted as FatoXtract. Its performance is compared with a time-of-flight camera, Microsoft Kinect v2. This methodology proposed considers a probabilistic fusion of three classifiers: Naïve Bayes, Artificial Neural Networks and Support Vector Machines. In order to attain an improved performance on the overall classification accuracy, multiple features in time domain (e.g., velocity) and frequency domain (e.g., Fast Fourier Transform), combined with traditional geometrical features (e.g., asngular positions of the joints), were considered.

Data acquisition was collected for five different daily activities performed by six participants with 20 trials, using FatoXtract and Kinect v2. The dataset provided in this work was designed to be extremely challenging since there are high intra-class variations, the duration of the action clips vary dramatically, and some of the actions are quite similar (e.g., brushing teeth and waving).

Keywords: pattern recognition; feature extraction; classification methods; human movement analysis; temporal classification; inertial suit; Naïve Bayes; Artificial Neural Network; Support Vector Machine

ÍNDICE

CAPÍTULO 1. INTRODUÇÃO	1
1.1. Objetivos	2
1.2. Estrutura do Relatório	3
CAPÍTULO 2. ESTADO-DA-ARTE	5
2.1. Enquadramento Histórico	5
2.2. Aquisição do Movimento Humano: Tecnologias	6
2.2.1. Câmaras com Marcadores Ativos <i>versus</i> Passivos	6
2.2.2. Câmaras de Profundidade	9
2.2.3. Câmara <i>Time-of-Flight</i>	10
2.2.4. Sensores Inerciais	11
2.2.5. Comparação das Tecnologias	14
2.3 Análise do Movimento Humano: Métodos de Classificação	17
2.3.1. Redes Neurais Artificiais	17
2.3.2. Redes Bayesianas	19
2.3.3. Máquinas de Vetor de Suporte	21
2.4. Trabalhos Relacionados	22
2.5. Sumário	26
CAPÍTULO 3. ARQUITECTURA PARA CLASSIFICAÇÃO DO MOVIMENTO HUMANO	27
3.1. Pré-Processamento	27
3.2. <i>Features</i>	27
3.2.1. Ângulos de Euler e Quaterniões (<i>fs1</i>)	29
3.2.2. Velocidade Angular (<i>fs2</i>)	31
3.2.3. Aceleração Angular (<i>fs3</i>)	31
3.2.4. Análise da Componente Principal (<i>fs4</i>)	32
3.2.5. Valores Acumulativos (<i>fs5</i>)	33
3.2.6. FFT (<i>fs6</i>)	33
3.3. Arquitetura Geral	35
3.3.1. Atribuição de Pesos para o Modelo de Fusão usando a Entropia	38
3.3.2. Classificadores de Base para a Fusão DBMM	39
3.4. Sumário	40
CAPÍTULO 4. RESULTADOS EXPERIMENTAIS	41
4.1. Descrição do Setup Experimental	41
4.2. Descrição do Procedimento	42
4.3. Descrição da Amostra	42

4.4. Resultados	43
4.4.1. <i>Features</i> do Domínio Temporal	43
4.4.2. <i>Features</i> no Domínio da Frequência	44
4.4.3. Todas as <i>Features</i>	45
4.4.4. Discussão sobre a escolha das <i>Features</i>	46
4.4.5. <i>Benchmarking</i>	48
4.5. Discussão	50
CAPÍTULO 5. CONCLUSÃO E TRABALHO FUTURO	53
5.1. Conclusão	53
5.2. Trabalho Futuro	54
REFERÊNCIAS	55
ANEXO I	61
Publicações Relacionadas	61
ANEXO II	63
Modelo Denavit Hartenberg	63

ÍNDICE DE FIGURAS

Figura 1 - Leonardo Da Vinci (1452-1519).....	5
Figura 2 – Giovanni Borelli (1608-1679).....	5
Figura 3 - Jules E.Marey (1830-1904)	5
Figura 4 - Otto Fischer (1861-1917).....	6
Figura 5 - Christian Braune (1831-1892)	6
Figura 6 - Uso de marcadores para análise do movimento humano.....	7
Figura 7 - Câmara RGB-D (ASUS Xtion).....	9
Figura 8 – IMU (3-Space Embedded)	11
Figura 9 - Representação de um Giroscópio	12
Figura 10 – Características técnicas do <i>FatoXtract</i>	15
Figura 11 – Exemplo de uma ANN de 2 camadas	18
Figura 12 - Redes Bayesianas	19
Figura 13 – Máquinas de Vetor de Suporte.....	21
Figura 14 – Exemplo da concatenação de 20 ensaios	27
Figura 15 - FFT nas frames: a) 100; b) 500; c) 1000; d) 1500.....	34
Figura 16 - Modelo DBMM	36
Figura 17 - Exemplo de um participante a executar a terceira atividade (andar de bicicleta). (a) Participante equipado com <i>FatoXtract</i> ; (b) Avatar 3D obtido do <i>FatoXtract Visualizer</i> ; (c) Numeração dos IMUs.....	41
Figura 18 – Matriz Confusão global usando o classificador DBMM e as as features espaço-temporais propostas (frequência + temporal + geométrica)	47
Figura 19 - Comparação entre os dois tipos de features.....	48
Figura 20 - Matriz Confusão dos dados do Kinect v2.....	49

ÍNDICE DE TABELAS

Tabela 1 - Vantagens e Desvantagens do uso de marcadores	7
Tabela 2 - Vantagens e Desvantagens do uso de câmaras de profundidade	10
Tabela 3 - Vantagens e Desvantagens do uso de Câmaras <i>Time-of-Flight</i>	11
Tabela 4 – Características do Sensor Inercial 3-Space Embedded	13
Tabela 5 - Comparação de produtos das diferentes tecnologias	16
Tabela 6 – <i>Features</i> e métodos de classificação usados na literatura usando sensores inerciais	25
Tabela 7 - <i>Features</i> usadas neste trabalho	29
Tabela 8 - Descrição do número de <i>features</i> usadas por frame	35
Tabela 9 - Atividades	42
Tabela 10 - Precisão e Sensibilidade dos diferentes participantes e atividades utilizando <i>features</i> do domínio temporal	44
Tabela 11 - Precisão e Sensibilidade dos diferentes participantes e atividades utilizando <i>features</i> do domínio da frequência	45
Tabela 12 - Precisão e Sensibilidade dos diferentes participantes e atividades usando os dois tipos de <i>features</i>	46
Tabela 13 - Comparação dos valores Precisão e Sensibilidade do <i>FatoXtract</i> e do Kinect v2 50	

SIMBOLOGIA

t - Instante de tempo t

q_n^t - Quaterniões no instante de tempo t do IMU n , $n = \{1, \dots, 17\}$

θ_n^t - Ângulo de euler no instante de tempo t do IMU n , $n = \{1, \dots, 17\}$

M_e - Matriz de ângulos de euler (3×3)

M_q - Matriz de quaterniões (4×3)

lcM_κ - Vetor que contém os elementos da matriz triangular superior de M_κ , $\kappa = \{e, q\}$

\log - Logaritmo

v_q - Velocidade angular do quaterniões q

ΔT - Intervalo de amostragem

fr - Taxa de aquisição

\bar{X}, μ - Média

σ - Desvio padrão

a_{v_q} - Aceleração da velocidade angular do quaterniões q

C - Matriz resultante da correlação cruzada 2D de dois sinais

$cov_{i,j}$ - Covariância de C

$E_{cov(i,j)}^t$ - Energia da matriz de covariância $cov_{i,j}$ ao longo do tempo t

Y_i^t - Valor acumulado da *feature* i no instante de tempo t

y_i^t - *Feature* i no instante de tempo t

ψ - Dimensão de y_i^t

$A(\theta_n^t)$ - Amplitude dos ângulos de euler

Y_n^k - Resultado obtido pela aplicação da FFT aos dados da junta n

f_H - Frequência máxima

w_i^t - Peso do neurónio i no instante de tempo t

η - Taxa de aprendizagem da ANN

x_i^t - Entrada para o neurónio i no instante de tempo t

e^t - Erro da saída da ANN no instante de tempo t

d^t - Saída esperada da ANN no instante de tempo t

y^t - Resposta calculada pela ANN no instante de tempo t

X_i - Variável de entrada para o nó i

$P(A)$ - Probabilidade de A

$P(B)$ - Probabilidade de B

$P(A|B)$ - Probabilidade de A sabendo B

$P(B|A)$ - Probabilidade de B sabendo A

A_m^t - Modelo com m atributos

\mathcal{M}_{trans} - Modelo para a Probabilidade da Transição de Estado entre as Variáveis Classe

β - Fator de Normalização do Método DBMM

$P(C^t|A)$ - Distribuição de probabilidade do DBMM

$P_i(A|C^t)$ - Resultado *à posteriori* para o classificador base i no instante de tempo t

$P(C^t|C^{t-1})$ - Distribuição de probabilidade de transição entre as variáveis de classe ao longo do tempo

$H_i(\mathcal{L})$ - Nível de confiança utilizado para atribuir pesos

α - Fator de Normalização do Classificador NB

\mathcal{L} - Conjunto com Probabilidades Condicionais

ABREVIATURAS

2D - Duas Dimensões

3D - Três Dimensões

ANN - Rede Neuronal Artificial (*Artificial Neural Network*)

AR - Modelo Autorregressivo (*Autoregressive Model*)

BIC - Critério de Informação Bayesiano (*Bayesian Information Criterion*)

CPT - Tabela de Probabilidade Condicional (*Conditional Probability Table*)

DAG - Grafos Direcionados e Acíclicos (*Directed Acyclic Graph*)

DBMM - *Dynamic Bayesian Mixture Models*

DCT - Transformada Discreta do Cosseno (*Discrete Cosine Transform*)

FFT - Transformada Rápida de Fourier (*Fast Fourier Transform*)

HMM - Modelo Escondido de Markov (*Hidden Markov Model*)

IBL - Aprendizagem Instanciada (*Instance - Based Learning*)

IMU - Unidade de Medida Inercial (*Inertial Measurement Units*)

IQR - Variação Interquartil (*Interquartile Range*)

IR - Infravermelhos (*Infrared*)

k-NN - k- Vizinho Mais Próximo (*k-Nearest Neighbor*)

LDQA - Análise Linear Discriminativa (*Linear Discriminant Analysis*)

LED - Díodo Emissor de Luz (*Light Emitting Diode*)

MAD - Desvio Médio Absoluto (*Mean Absolute Deviation*)

MAP - Máxima Probabilidade à *Posteriori* (*Maximum a Posteriori Probability*)

MEMS - Sistemas Micro-Eletromecânicos (*Microelectromechanical Systems*)

MLE - Estimativa de Probabilidade Máxima (*Maximum-Likelihood Estimation*)

MoCap - Captura de Movimento (*Motion Capture*)

NB - Teorema de Bayes (*Naive Bayes*)

PCA - Análise das Componentes Principais (*Principal Component Analysis*)

Pfd - Função de Densidade de Probabilidade (*Probability Density Function*)

Pre - Precisão

RGB-D - *Red-Green-Blue - Depht*

Sens - Sensibilidade

SVM - Máquinas de Vetor de Suporte (*Support Vector Machine*)

ToF - Tempo de voo (*Time-of - Flight*)

INTRODUÇÃO

A Biomecânica é a ciência que estuda e analisa o movimento humano baseando-se na análise dos parâmetros cinemáticos e dinâmicos do movimento (Amadio & Serrão, 2007). A análise cinemática do movimento humano é deveras importante em várias áreas, tais como no desporto (*e.g.*, para análise do desempenho atlético), na saúde (Aggarwal & Cai, 1999) (*e.g.*, na reabilitação de pessoas com deficiência motora), nas atividades diárias, no reconhecimento de ações através de vigilância, e na assistência à autonomia no domicílio (*ambient assisted living*) (*e.g.*, usando robôs ou casas inteligentes). Através da Biomecânica, torna-se possível analisar as causas e os fenómenos do movimento e, por sua vez, compreender a complexidade do movimento humano. A Biomecânica pode ser subdividida em duas áreas de estudo: biomecânica interna e biomecânica externa (Amadio & Serrão, 2007). A Biomecânica interna consiste nas forças internas, tais como forças musculares, forças nos tendões, ligamentos e ossos. A Biomecânica externa estuda as grandezas observáveis externamente na estrutura do movimento.

O estudo da cinemática do corpo humano envolve vários métodos que recorrem à análise de diversos parâmetros que advêm do movimento. Importantes parâmetros a ter em conta são a aceleração, velocidade e posição (lineares ou angulares) das diversas articulações do corpo humano, que podem ser medidos por sensores ou através da análise de repetidas imagens obtidas por câmaras (Zhou & Hu, 2008). Tal aquisição do movimento humano é uma tarefa difícil, mas, com o avanço da tecnologia, tem-se tornado cada vez mais fácil, mais económico e mais cómodo para o utilizador, possibilitando a aquisição da posição ou rotação das diferentes juntas (*i.e.*, conexão entre dois elos) e/ou elos (corpo rígido que possui, pelo menos, uma junta que permite anexar a outros elos) do corpo humano, mantendo a *validade ecológica*¹ da ação.

Neste relatório é apresentado um levantamento das principais tecnologias e dos produtos existentes no mercado utilizados no âmbito da análise do movimento humano. Existem várias tecnologias, tais como as câmaras de profundidade (RGB-D), câmaras *time-of-flight* (ToF), câmaras de infravermelhos com marcadores passivos e ativos, e sensores vestíveis (normalmente designados por sensores *wearable*). Os sensores *wearable* tendem a ser a tecnologia mais adequada para a análise cinemática do movimento humano devido às diversas vantagens que apresentam, entre as quais se destaca a validade ecológica dos mesmos face aos sistemas de infraestruturas com câmaras (Chen X. , 2013).

¹ A validade ecológica significa que os comportamentos e ações observados em determinado estudo refletem os comportamentos e ações que ocorrem em contexto real, independentemente dos métodos, materiais e *setup* considerados.

Este trabalho apresenta uma metodologia para classificar um conjunto de cinco movimentos humanos que compõem as cinco atividades diárias, usando duas tecnologias diferentes para o efeito: um fato de captura e aquisição de movimento (*Motion Capture*, ou MoCap), denominado de *Ingenarius FatoXtract*, e uma câmara *time-of-flight*, denominado de *Microsoft Kinect v2*. Este trabalho descreve a estrutura de reconhecimento das atividades que dependem de características (doravante designadas por *features*), divididas no domínio do tempo e da frequência. Essas *features* são utilizadas para identificar as atividades com recurso a métodos de classificação supervisionados. Com o intuito de melhorar essa identificação, este trabalho apresenta ainda uma arquitetura de classificação que se baseia na fusão de diferentes classificadores, oferecendo, simultaneamente, vantagens do ponto de vista funcional e computacional, com o intuito de realizar uma análise eficiente da cinemática do movimento humano.

1.1. Objetivos

Este Projeto de Mestrado tem como objetivo principal identificar atividades diárias, de forma autónoma, com recurso a métodos de classificação supervisionados. Para esse efeito, utilizou-se o *FatoXtract* e o *Kinect v2* para aquisição de dados inerentes ao movimento humano. O *FatoXtract* é um equipamento que beneficia de sensores inerciais (*inertial measurement units*, ou IMU), do tipo sensor *wearable*, utilizando tecnologia *wireless* de forma a permitir capturar a cinemática 3D do corpo humano virtualmente em qualquer local, seja em ambientes exteriores ou interiores, e independentemente do tipo de terreno e condições atmosféricas. O equipamento dispensa a utilização de câmaras, marcadores ou outro tipo de infraestrutura. O *Kinect v2* consiste numa câmara *time-of-flight* de baixo custo, tendo sido utilizado neste trabalho apenas como tecnologia de comparação.

Deste modo, de forma a promover a adequada escolha do equipamento, inicialmente é apresentado um levantamento de tecnologias, produtos existentes no mercado e métodos de análise cinemática do movimento humano. De seguida, são apresentadas um conjunto de *features* no domínio do tempo e da frequência, baseadas em ângulos de euler e quatérniões usados para reconhecimento de atividades do movimento humano, adquiridos através do *FatoXtract*, bem como a respetiva arquitetura de classificação adotada. Por último, é apresentada a avaliação e validação do conjunto de *features* propostas, comparando o *FatoXtract* com o *Kinect v2*.

Os casos de estudo abordados encontram aplicações em biomecânica, reabilitação, ergonomia, jogos e ciências do desporto.

1.2. Estrutura do Relatório

Este relatório encontra-se dividido em 5 capítulos e 2 anexos.

O Capítulo 1 (atual) é composto pela introdução, onde é feita uma breve descrição sobre o tema apresentado, bem como dos objetivos deste Projeto de Mestrado.

No Capítulo 2 é realizado o levantamento do estado-da-arte, encontrando-se dividido em cinco subcapítulos. Inicialmente, é feito um breve enquadramento histórico, sendo seguido por um estudo de várias tecnologias existentes para captura do movimento humano. Posteriormente, é apresentada uma comparação das diferentes tecnologias e produtos existentes no mercado. Este capítulo apresenta também uma breve explicação acerca dos diferentes métodos de classificação utilizados neste trabalho. Isto é procedido de uma apresentação dos diversos métodos de classificação e *features* existentes na literatura para análise cinemática do movimento humano com o uso de sensores *wearable*.

No Capítulo 3 é apresentada a arquitetura usada para a classificação do movimento humano. Está dividido em quatro subcapítulos, descrevendo o pré-processamento dos dados, as diferentes *features* implementadas e a arquitetura geral do método aplicado.

O Capítulo 4 consiste na apresentação e discussão dos resultados experimentais, encontrando-se dividido em 5 subcapítulos. Inicialmente é realizada uma descrição do *setup* experimental, do procedimento e da amostra. De seguida, são discutidos os resultados obtidos, sendo também realizada uma comparação entre as duas tecnologias utilizadas para a aquisição do movimento humano (o *FatoXtract* e o *Kinect v2*). Por último, é feita uma discussão geral dos resultados obtidos utilizando a metodologia proposta.

O Capítulo 5, e último capítulo deste Relatório de Projeto de Mestrado, apresenta as conclusões e propostas de trabalho futuro.

Por último, são apresentados dois anexos, o primeiro referente às publicações relacionadas com este trabalho, e o segundo contém a aplicação do modelo Denavit Hartenberg.

ESTADO-DA-ARTE

Neste capítulo é apresentado o estado-da-arte sobre as tecnologias e métodos para a aquisição, análise e classificação do movimento humano. Está dividido em cinco secções.

A Secção 2.1 faz um breve enquadramento histórico sobre a análise do movimento humano. A Secção 2.2 descreve e compara as tecnologias mais relevantes para a aquisição do movimento humano. A Secção 2.3 apresenta alguns dos métodos de classificação mais usados, nomeadamente as Redes Neurais, as Redes Bayesianas e as Máquinas de Vetor de Suporte. A Secção 2.4 apresenta um levantamento do estado-da-arte e, por último, a Secção 2.5 apresenta um breve sumário do capítulo.

2.1. Enquadramento Histórico

A análise biomecânica do movimento humano é um tema que já vem a ser discutido desde antes de Cristo, nomeadamente por filósofos como Aristóteles, Arquimedes, Galeno, Da Vinci, Galileo, Newton, Giovanni Borelli, Jules E. Marey, Braune e Fischer.

Leonardo da Vinci (Figura 1) realizou o primeiro estudo sobre a anatomia humana, partindo do contexto mecânico. Estudou a função das articulações, analisou as forças musculares e de que forma estas atuam ao longo das linhas que fazem as ligações às origens e às intersecções. Assim, Da Vinci descreveu as atividades de “caminhar”, “saltar” e “ficar em pé” a partir da posição “sentada” (Pinheiro, 2008).



Figura 1 - Leonardo Da Vinci (1452-1519)²



Figura 2 - Giovanni Borelli (1608-1679)³

Anos mais tarde, Giovanni Borelli (Figura 2) tornou-se no primeiro cientista a aplicar os princípios da Física aos estudos dos movimentos do ato de deslocação nos animais. É considerado o pai da Biomecânica. Borelli foi ainda o responsável por ter definido o centro de gravidade do corpo (Magarreiro, 2015).

Jules E. Marey (Figura 3) iniciou o desenvolvimento de instrumentos para análise do movimento, usando métodos gráficos e fotográficos.



Figura 3 - Jules E. Marey (1830-1904)⁴

² <http://www.biography.com/people/leonardo-da-vinci-40396>

³ https://en.wikipedia.org/wiki/Giovanni_Alfonso_Borelli

⁴ https://fr.wikipedia.org/wiki/%C3%89tienne-Jules_Marey

Fischer (Figura 4) e Braunne (Figura 5) foram os responsáveis pelo primeiro estudo que retratou a análise matemática a três dimensões (3D) da marcha humana (Schepers, 2009).



Figura 4 -
Otto Fischer
(1861-1917)⁵



Figura 5 -
Christian
Braune (1831-
1892)⁶

2.2. Aquisição do Movimento Humano: Tecnologias

Na análise cinemática do movimento humano devem ser considerados os seguintes parâmetros: aceleração, velocidade, tempo de duração e posição dos segmentos corporais (Tözeren, 2000). Estes podem ser medidos através de sensores ou recorrendo à análise de imagens obtidas por câmaras de filmar.

Nesta secção é feito o levantamento das principais tecnologias para a captura do movimento humano, nomeadamente os marcadores ativos e passivos (Secção 2.2.1), as câmaras de profundidade (Secção 2.2.2), as câmaras *time-of-flight* (Secção 2.2.3) e os sensores inerciais (Secção 2.2.4). Por último, é realizada uma análise comparativa entre os vários produtos existentes no mercado que usam estas tecnologias (Secção 2.2.5).

2.2.1. Câmaras com Marcadores Ativos versus Passivos

Na análise biomecânica do movimento utilizando câmaras é usual colocarem-se marcadores retrorrefletores nas extremidades dos segmentos corporais que se pretendem analisar, permitindo, assim, a identificação da posição e do movimento desses segmentos.

Os marcadores podem ser definidos como objetos revestidos de material refletor que, atados ou fixados ao corpo da pessoa (Figura 6), facilitam o rastreamento das imagens utilizando métodos de segmentação e deteção adequados (Couceiro, et al., 2013). Normalmente, é necessário mais do que uma câmara para obter a localização dos marcadores, pois alguns podem ficar ocultos em determinados movimentos. Cada segmento do corpo humano deve ser definido, pelo menos, com três marcadores, o que permite obter movimentos 3D reduzindo o risco de oclusão causada por outras partes do corpo.

⁵ <http://slideplayer.com.br/slide/4937261/>

⁶ https://www.uni-leipzig.de/unigeschichte/professorenkatalog/leipzig/Braune_751/

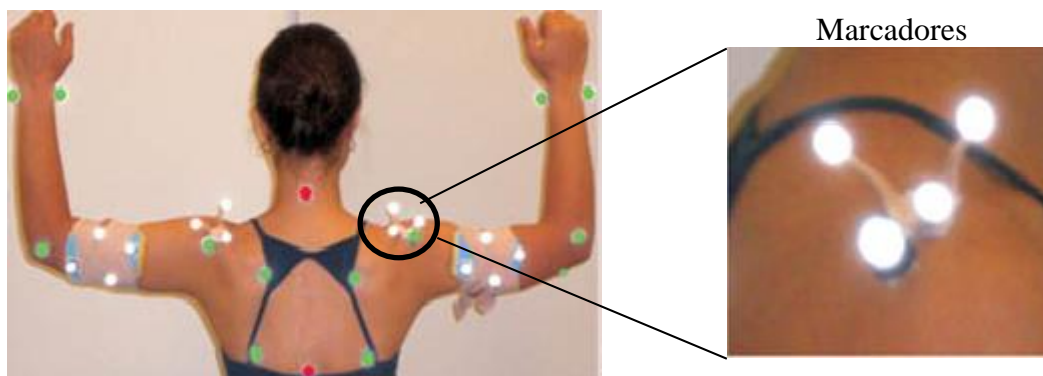


Figura 6 - Uso de marcadores para análise do movimento humano⁷

A Tabela 1 apresenta um levantamento das principais vantagens e desvantagens do uso de marcadores.

Tabela 1 - Vantagens e Desvantagens do uso de marcadores

Vantagens	Desvantagens
Possibilitam altas velocidades de captura e altos níveis de precisão e exatidão	Pode ser necessário um pós-processamento intenso dos dados
Os movimentos de múltiplos sujeitos podem ser capturados simultaneamente	Possibilidade de oclusão dos marcadores, resultando em perda de dados
Um alto número de marcadores pode ser usado e a configuração dos mesmos pode ser alterada facilmente	<i>Hardware</i> e <i>software</i> geralmente mais caros do que os de outros sistemas
O objeto ou indivíduo pode-se movimentar livremente dentro do volume de captura, não sofrendo obstrução por fios ou equipamentos que limitam os movimentos	A maioria dos sistemas necessita que a luz do ambiente seja controlada
O volume de captura pode ser grande o suficiente para vários tipos de aplicações	Tem de se recorrer à ajuda da colocação de marcadores no corpo da pessoa
	Pouco prático pois tem de se montar um sistema de câmaras no local onde se pretende utilizar

Os marcadores dividem-se em dois tipos: ativos e passivos. Em seguida, são descritos cada um dos dois tipos de marcadores.

⁷ http://www.moreirajr.com.br/revistas.asp?fase=r003&id_materia=4103

- **Marcadores Ativos**

Os marcadores ativos emitem luz através de díodos emissores de luz (LED) que operam na frequência do infravermelho (IR). O facto de não ocorrer interferências entre os marcadores constitui uma vantagem deste método, permitindo assim que os marcadores possam ser colocados juntos. A conjugação deste tipo de marcadores com câmaras IR permite rapidamente conhecer a localização de cada marcador através de funções básicas de processamento de imagem. No entanto, são impostas restrições no número, proximidade e tamanho dos marcadores necessários para a aquisição dos dados, bem como constrangimentos adicionais associados à luz ambiente, campo de visão das câmaras e tipos de movimentos a serem analisados. Uma desvantagem no uso das câmaras IR é que estas registam somente as coordenadas dos marcadores e não toda a imagem do movimento (Suaide, 2010). O facto deste tipo de estratégia exigir mais equipamento e manutenção torna-o pouco atrativo (Araújo, Andrade, & Barros, 2005). No entanto, existem várias soluções no mercado baseadas neste tipo de tecnologia tais como o *Selpot*, *OptiTrack*⁸ e o *Colstel* (Macellari, 1983).

- **Marcadores Passivos**

Os marcadores passivos são apenas refletores de luz e, por isso, os sistemas que os usam não necessitam de condutores elétricos, baterias ou manutenção periódica na sua utilização. Os LEDs dispostos à volta de cada câmara enviam pulsos de radiação IR que são posteriormente refletidos pelos marcadores para a lente da câmara. As lentes possuem filtros IR e sistemas de *threshold* que identificam os marcadores mais brilhantes e suprimem a informação dos objetos menos brilhantes. Este método diminui a validade ecológica do movimento da pessoa quando comparado com o uso de marcadores ativos (Silva A. F., 2011). Cada câmara grava uma imagem bidimensional, aparecendo os marcadores na imagem como pontos brilhantes. Posteriormente, o sistema de processamento de imagem isola os pontos dos marcadores na imagem e grava a sua posição. Apesar de ser possível implementar estes sistemas de marcadores passivos com apenas três câmaras, seis é o número recomendado. Existem várias soluções baseadas em marcadores passivos tais como o *Elite* (Ferrigno & Pedotti, 1985), *MaxReflex* (Josefsson, Nordh, & Eriksson, 1996) e o *Vicon*⁹.

O uso de marcadores, em conjunto com os sistemas de vídeo, permitem obter a posição absoluta de um segmento no espaço. Estes dados podem ainda combinar-se com os valores de momentos

⁸ www.optitrack.com

⁹ www.vicon.com

de massa de inércia e com as velocidades e acelerações de cada segmento e articulação, com o intuito de avaliar os momentos das articulações e também a energia associada a rotações instantâneas das articulações (Barbosa, 2011).

2.2.2. Câmaras de Profundidade

A câmara de profundidade, tipicamente identificada como RGB-D (Figura 7), consiste numa câmara de cores (*Red-Green-Blue*) com adição de um sensor de profundidade (*Depth*). Estas câmaras, associadas a um módulo de processamento, são capazes de extrair informação sobre o posicionamento de um sujeito. Este processo apenas permite a identificação dos padrões, existindo por isso a necessidade de um treino *offline*, constituindo isto uma desvantagem face às alternativas anteriores. Os dados de profundidade são fornecidos tipicamente através de um sensor de infravermelhos (Litomisky, 2012). Existem atualmente diversos sistemas RGB-D de baixo custo, nomeadamente a primeira versão do *Microsoft Kinect*¹⁰ e o *ASUS Xtion*¹¹.



Figura 7 - Câmara RGB-D (ASUS Xtion)¹²

Na Tabela 2 são apresentadas as vantagens e desvantagens do uso de câmaras de profundidade (Dubois & Charpille, 2013).

¹⁰ <https://en.wikipedia.org/wiki/Kinect>

¹¹ https://www.asus.com/3D-Sensor/Xtion_PRO/

¹² https://www.asus.com/3D-Sensor/Xtion_PRO_LIVE/

Tabela 2 - Vantagens e Desvantagens do uso de câmaras de profundidade

Vantagens	Desvantagens
Sistema fácil de distribuir e instalar	Difícil a utilização em outdoor, principalmente em espaços amplos
Baixo custo	Necessidade de treino offline para reconhecer movimentos do humano
	Pouco prático pois tem de se montar a câmara sempre no local onde se pretende utilizar

2.2.3. Câmara *Time-of-Flight*

As câmaras *time-of-flight* (*tempo de voo*, ou ToF), à semelhança das anteriores, são uma tecnologia que oferece uma vasta informação sensorial sobre o cenário, tendo um sistema de instalação não invasivo. No entanto, contrariamente à tecnologia anterior, as câmaras ToF proporcionam imagens 3D numa *frame rate* elevada e, simultaneamente, fornecem uma medida de profundidade em cada pixel do cenário. O intervalo de dados fornecido permite fazer a segmentação do corpo humano e também distinguir os diferentes movimentos, que normalmente são indistinguíveis em sistemas monoculares (Ganapathi, Plagemann, Koller, & Thrun, 2010).

Basicamente, o princípio de funcionamento das câmaras ToF baseiam-se na medição do tempo que a luz demora a viajar a partir de uma fonte de luz para os objetos no campo de visão e de volta para o sensor. A luz é emitida através de LEDs IR sendo, em seguida, refletida de volta para o sensor. Os sistemas convencionais ToF usam um detetor de ponto e digitalizam um feixe de *laser* modulado através do cenário, afim de adquirir uma imagem 3D (Lange & Seitz, 2001). As câmaras ToF são usadas em várias áreas de pesquisa tais como a computação gráfica, visão artificial e robótica (Uebersax, 2010). O *Kinect v2*¹³ é um exemplo deste tipo de tecnologia.

As câmaras ToF têm vantagens (Foix, Alenyà, & Torras, 2011) e desvantagens (Cui, Schuon, Chan, Thrun, & Theobalt, 2010) que são resumidas na Tabela 3.

¹³ www.microsoft.com/en-us/kinectforwindows/purchase/

Tabela 3 - Vantagens e Desvantagens do uso de Câmaras *Time-of-Flight*

Vantagens	Desvantagens
Profundidade em cada pixel com um elevado frame rate	Suscetibilidade à iluminação
Baixo peso, pequeno e design compacto	Interferências: se várias câmaras ToF estiverem em execução em simultâneo, estas podem perturbar as medições de cada uma, dependendo do tempo de multiplexação e das frequências de modulação
Auto iluminação	Múltiplas reflexões: as câmaras ToF iluminam um cenário inteiro.
Reduzido consumo de energia.	

2.2.4. Sensores Inerciais

Um método alternativo para analisar o movimento humano é através de sensores inerciais (Figura 8), também conhecidos como unidades de medição inerciais (*inertial measurement unit*, ou IMU), colocados sobre o corpo. Estes sensores permitem que as medições sejam feitas fora do ambiente de laboratório, o que é muito vantajoso (Takeda, Tadano, Natorigawa, Todoh, & Yoshinari, 2009). Os sensores inerciais são transdutores que exploram a propriedade da inércia, nomeadamente a resistência a alterações de movimento para avaliar o movimento angular e as variações em movimento linear (Chen X. , 2013). Os IMU dão a informação da aceleração linear, da velocidade de rotação e das medidas de campo magnético, a partir da sua orientação (Miezal, Bleser, Schmitz, & Stricker, 2013).

**Figura 8** – IMU (3-Space Embedded)¹⁴

Com o avanço da tecnologia surgiram os sistemas micro-eletromecânicos (*Microelectromechanical Systems*, ou MEMS). Estes equipamentos permitem que as medições

¹⁴ <https://www.yeitechnology.com/productdisplay/3-space-embedded-0>

sejam feitas fora do contexto laboratorial, o que é de extrema importância para a análise do movimento humano (Takeda, Tadano, Natorigawa, Todoh, & Yoshinari, 2009). Os sensores inerciais sem fios são mais leves que outros equipamentos similares, o que faz com que seja possível realizar o rastreamento do movimento humano a qualquer instante e em qualquer lugar. Por outro lado, quando comparado com as alternativas anteriores, os sensores inerciais são caracterizados por terem baixo consumo de energia e baixo custo, proporcionando ao utilizador a monitorização precisa das atividades diárias sem necessidade de se deslocar para um cenário com infraestrutura de câmaras instalada. Tal como o *FatoXtract*, o *Xsens MVN* é um exemplo de um fato *wearable* para a análise de movimento humano que utiliza sensores inerciais (Roetenberg, Luinge, & Slycke, 2009).

Tipicamente, os IMU compreendem um giroscópio, um acelerómetro e um magnetómetro (Kok, Hol, & Schön, 2014). Estes três sensores inerciais foram usados neste trabalho e de seguida vai ser feita uma breve descrição sobre cada um deles:

- **Giroscópio**

O giroscópio (Figura 9) é um sensor que permite medir a orientação dos segmentos corporais no espaço 3D, bem como efetuar medições de velocidade de rotação de um objeto em torno de um centro de rotação. Ou seja, o giroscópio consegue mensurar a velocidade angular de um determinado objeto, servindo apenas como referência de direção. Com a utilização do giroscópio é possível calcular, continuamente, a orientação do sensor no espaço 3D mesmo quando este se encontra sujeito a aceleração (Carvalho, 2011). Todavia, a calibração do giroscópio é muitas vezes difícil de operacionalizar face ao seu *delay*, causando um efeito conhecido como *drift*. Para tal, uma polarização DC não-zero é introduzida na leitura do sensor, produzindo assim uma medição de velocidade angular diferente de zero, *i.e.*, quando nenhum movimento está realmente a ocorrer (Lin, 2012).

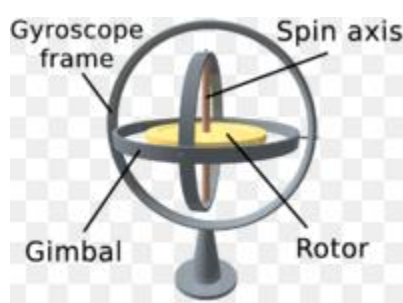


Figura 9 - Representação de um Giroscópio¹⁵

¹⁵ <https://pt.wikipedia.org/wiki/Girosc%C3%B3pio>

- **Acelerómetro**

Os acelerómetros têm vindo a ser usados em estudos da cinemática humana desde 1990 (Fong & Chan, 2010). Os acelerómetros são dispositivos eletromecânicos capazes de medir a aceleração e a velocidade própria de objeto, calculando assim as forças exercidas sobre ele (Yang & Hsu, 2010). Tipicamente, contêm uma pequena massa conectada a uma mola que incorpora sensores que permitem medir a deflexão da mesma quando esta é acelerada (Gomes, 2012). Este dispositivo oferece um método prático e relativamente económico de captura do movimento humano. Sistemas baseados em acelerómetros são capazes de adquirir e estimar vários valores cinemáticos como a velocidade e o deslocamento, quando ligados a vários segmentos do corpo do sujeito (Lin, 2012). O acelerómetro pode ser usado para compensar o desvio do giroscópio sobre os eixos do plano horizontal quando o objeto não está a sofrer nenhum tipo de aceleração.

- **Magnetómetro**

O magnetómetro é usado juntamente com o giroscópio e o acelerómetro para análise cinemática do movimento humano, permitindo obter a orientação do sensor em relação ao norte magnético (Garofalo, 2010). Os magnetómetros magneto indutivos e os magnetómetros de efeito Hall são os mais conhecidos devido à sua sensibilidade e baixo consumo de energia (Lin, 2012).

A Tabela 4 apresenta as características principais de cada um dos sensores descritos anteriormente para o caso do *3-Space Embedded* desenvolvido pela *Yei Technology*.

Tabela 4 – Características do Sensor Inercial 3-Space Embedded

	Giroscópio	Acelerómetro	Magnetómetro
Opções de Faixa de Leitura	$\pm 250 / \pm 500 / \pm 2000$ °/s	$\pm 2g / \pm 4g / \pm 8g$	± 0.88 Ga até ± 8.1 Ga
Resolução	16 bit	14 bit	12 bit
Sensibilidade	0.00833°/s /digito para ± 250 °/s	0.00024°/digito -0.00096°/digito	0.73 mGa/digito
Taxa de Comunicação	250 Hz com Kalman AHRS		
Temperatura de operação	-40°C até +85°C		
Opções de Comunicação	USB 2.0, SPI , série assíncrona		

2.2.5. Comparação das Tecnologias

Nesta secção são apresentados exemplos de produtos disponíveis no mercado e associados a cada uma das tecnologias apresentadas nas secções anteriores, nomeadamente:

- **Marcador Ativo - *Optitrack***

O *Optitrack* é um sistema de câmaras que se baseia no uso de marcadores ativos. Este sistema tem várias desvantagens, tais como a baixa portabilidade e o seu custo elevado. O processamento é feito em tempo real. Este tipo de sistemas só pode ser usado em ambiente de laboratório e com o máximo de 12 indivíduos.

- **Marcador Passivo - *Vicon***

O *Vicon* é um sistema de câmaras que se baseia no uso de marcadores passivos. Tal como o *Optitrack*, este tem baixa portabilidade e é dispendioso. Este sistema tem um espaço de trabalho mais reduzido, o que possibilita um máximo de 6 indivíduos.

- **Câmara de Profundidade - *Microsoft Kinect v1***

O *Microsoft Kinect v1* usa o princípio das câmaras de profundidade RGB-D. Este consegue detetar diversas articulações do corpo humano, de maneira não invasiva, o que pode trazer diversas vantagens. É eficiente para deteção e identificação dos ângulos de diferentes movimentos (Júnior, 2013). O sensor é capaz de capturar imagens com 640×480 pixels a 30 fps. Juntamente com a informação de profundidade, o sensor é capaz de produzir o modelo virtual do esqueleto da pessoa que está a ser capturada. O sensor de profundidade consiste num emissor de infravermelhos que cria um padrão estruturado de luz infravermelha, e uma câmara de infravermelho que lê a reflexão desses feixes. A câmara interpreta a deformação da projeção e converte essa informação em valores de profundidade, medindo a distância entre o objeto e o sensor (Correia, 2013). O *Microsoft Kinect v1* tem um espaço de trabalho de $2 \times 3.5 \times 2$ m o que apenas possibilita um máximo de 2 indivíduos.

- **Câmara ToF - *Microsoft Kinect v2***

O *Microsoft Kinect v2* usa o princípio das câmaras ToF. A taxa de aquisição é igual à da versão anterior (*Microsoft Kinect v1*), mas o número de indivíduos que é capaz de analisar simultaneamente já é superior, podendo ser, no máximo, de 6. O espaço de trabalho também aumentou relativamente à versão anterior (ver Tabela 5). Com o aparecimento do *Kinect v2*, o reconhecimento do movimento humano é realizado com uma resolução superior ao *Kinect v1*. O *Kinect v1* consegue detetar 20 juntas do corpo humano, enquanto que o *Kinect v2* permite detetar 26 juntas.

- **Sensor Inercial - *XSENS MVN***

O *XSENS MVN* é um fato de captura de movimento que recorre ao uso de sensores inerciais. Este não precisa de câmaras, emissores ou marcadores externos, podendo

assim ser utilizado no exterior e no interior. Não tem restrições de iluminação nem constrangimentos associados às oclusões. Este fato possui uma taxa de aquisição de 60 Hz e um custo de aproximadamente 60 000 €.

- **Sensor Inercial - *FatoXtract***

O *FatoXtract* (Figura 10) é um equipamento desenvolvido para a aquisição do movimento humano (*motion capture*, ou *MoCap*), não necessitando de câmaras, marcadores ou infraestruturas. Tendo a vantagem de possuir 17 IMU e comunicação *ZigBee*, o *FatoXtract* permite obter a configuração cinemática do corpo humano a uma frequência de 100 Hz, nomeadamente a posição, a velocidade e a aceleração angular das 17 juntas mais importantes do corpo humano. O *FatoXtract* consiste num fato *MoCap* de baixo custo que permite adquirir e analisar a informação cinemática do utilizador em tempo real, independentemente do género e/ou das suas propriedades morfológicas. Este tem um custo de 22 875 €, inferior ao *XSENS MVN*.

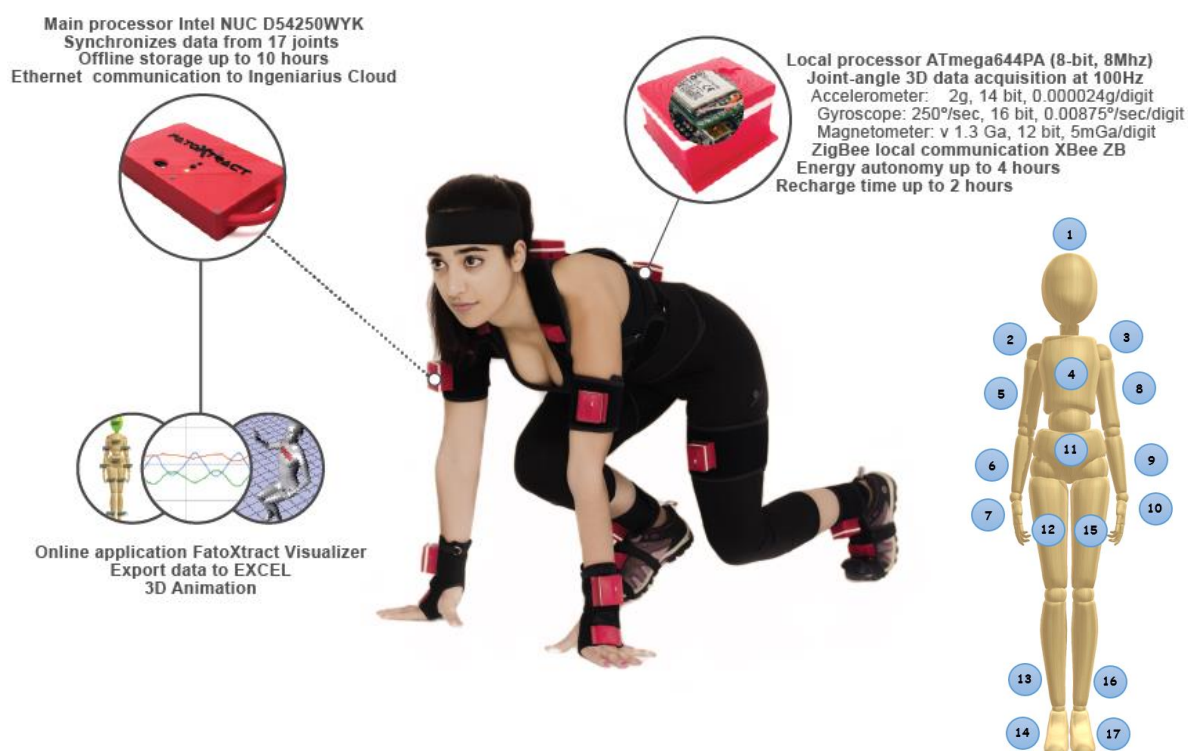

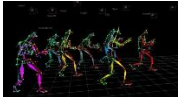

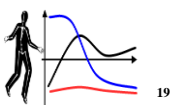




Figura 10 – Características técnicas do *FatoXtract*

Em suma, tanto o uso de marcadores como os sistemas de vídeo permitem uma análise eficaz do movimento humano. Porém, as câmaras de filmar são muito dispendiosas e apresentam várias limitações, como a portabilidade e o tempo elevado necessário para o processamento e análise de dados. A Tabela 5 apresenta uma análise comparativa destes produtos relativamente à taxa de aquisição, o número de indivíduos, o espaço de trabalho e o custo.

Tabela 5 - Comparação de produtos das diferentes tecnologias

Produtos	Tipo	Taxa de Aquisição (Hz)	Nº Indivíduos	Espaço de Trabalho	Custo	Representação Gráfica
Optitrack (p/ 72 câmaras)	Marcadores Ativos	180	até 12	26×26×3 m	~300 000 €	 16
Vicon (T-Series160 – p/ 24 câmaras)	Marcadores Passivos	120	até 6	9×9×5 m	~150 000 €	 17
Microsoft Kinect V1	Câmara de Profundidade	30	até 2	2×3.5×2 m	~150 €	 18
Microsoft Kinect v2 (iPi Biomech Add-On)	Câmara Time-Of-Flight	30	até 6	3×4×2 m	~200 € (+iPi Biomech Add-On 400€/ano)	 19
XSens MVN Awinda	Sensores Inerciais	60	1	-	~60 000 €	 20
FatoXtract	Sensores Inerciais	100	1	-	22 875 €	 21

Ao analisar a Tabela 5 podemos ver que o *Optitrack*, o *Vicon* e o *FatoXtract* são os que apresentam um valor maior de taxa de aquisição. Tanto o *Vicon* como o *Optitrack* apresentam a grande restrição a nível do espaço de trabalho, enquanto o *FatoXtract* não apresenta essa restrição, ou seja, pode ser usado em qualquer local. O *FatoXtract* apresenta também o custo mais baixo relativamente aos produtos que têm melhores características. Pode-se então concluir que, tendo por base a taxa de aquisição, espaço de trabalho e custo, o uso de sensores inerciais tende a ser mais adequado para a análise cinemática de movimentos humanos. Neste trabalho é usado um fato *MoCap*, o *FatoXtract*, e uma câmara ToF, o *Kinect v2*, com o propósito de

¹⁶ <https://www.naturalpoint.com/optitrack/systems/#motive-body/prime-41/72>

¹⁷ <http://www.creativebloq.com/audiovisual/how-track-motion-capture-data-10134857>

¹⁸ <http://www.fastmocap.com/>

¹⁹ <http://ipisoft.com/software/biomech-add-on/>

²⁰ <https://www.xsens.com/products/mvn-biomech-awinda/>

²¹ <http://www.ingenarius.pt/?page=fatoxtract>

comparar a eficiência de ambos, seja no que se refere à tecnologia como à seleção de *features*, para adequada classificação do movimento humano.

2.3 Análise do Movimento Humano: Métodos de Classificação

Existem vários métodos de classificação para reconhecer padrões do movimento humano, tais como Redes Neurais, Máquinas de Vetor de Suporte, Redes Bayesianas e Modelo Escondido de Markov. O reconhecimento de padrões pode ser realizado através de dois tipos de aprendizagem: supervisionada e não supervisionada.

A aprendizagem supervisionada é feita a partir de exemplos, em que o projetista ajuda o sistema a construir o modelo, através da definição das classes e dos exemplos em cada classe. O sistema tem que determinar a descrição para cada classe, ou seja, o conjunto de *features* comuns nos exemplos que lhe são fornecidos. Estando a descrição determinada, é possível formular a regra de classificação. Por sua vez, esta pode ser utilizada para prever a classe de um objeto que não tenha sido considerado aquando da aprendizagem. De forma simplificada, sabendo o que aconteceu no passado, pode prever o que se segue.

A aprendizagem não supervisionada é efetuada com base na observação. Não são definidas classes, pelo que o sistema necessita de observar os exemplos e reconhecer os padrões por si próprio. Daqui resulta um conjunto de descrições de classes, uma para cada classe descoberta na base de dados. Ou seja, dados vários padrões o sistema descobre semelhanças entre os mesmos e agrupa-os.

Neste trabalho optou-se por recorrer à classificação supervisionada dado que o objetivo consiste em identificar um conjunto de atividades do dia-a-dia previamente definidas. São apresentados de seguida alguns dos classificadores mais comuns que podem ser utilizados para este fim.

2.3.1. Redes Neurais Artificiais

As redes neurais artificiais (ANN) são modelos matemáticos ou computacionais que são inspirados no sistema nervoso central do ser humano (cérebro). Uma ANN é composta por grafos, sendo estes formados pela combinação de diversos neurónios artificiais (Brumatti, 2005). O neurónio artificial é uma estrutura lógica e matemática com um comportamento e com funções similares a um neurónio biológico. Numa ANN, os nós do grafo são os neurónios e as ligações entre os vários nós fazem a função das sinapses (Graupe, 2007). A interligação entre os neurónios de uma ANN é feita em diferentes camadas de cada sistema. A primeira camada é constituída pelos neurónios de entrada, que através das sinapses enviam dados para a segunda camada de neurónios, e assim sucessivamente, até chegar à última camada de neurónios que contém a saída (Yadav, Yadav, & Jain, 2014). Tipicamente quanto mais complexo for o sistema, mais camadas terá. As sinapses armazenam os parâmetros, também designados por pesos, que manipulam os dados nos cálculos (Figura 11) (Priddy, 2005).

Uma ANN é definida por três níveis:

1. Modelo de interligação entre os diferentes níveis de neurónios;
-

2. Processo de aprendizagem para atualizar os pesos das interligações;
3. A função de ativação que converte a entrada numa saída (Kumar & Sharma, 2014).

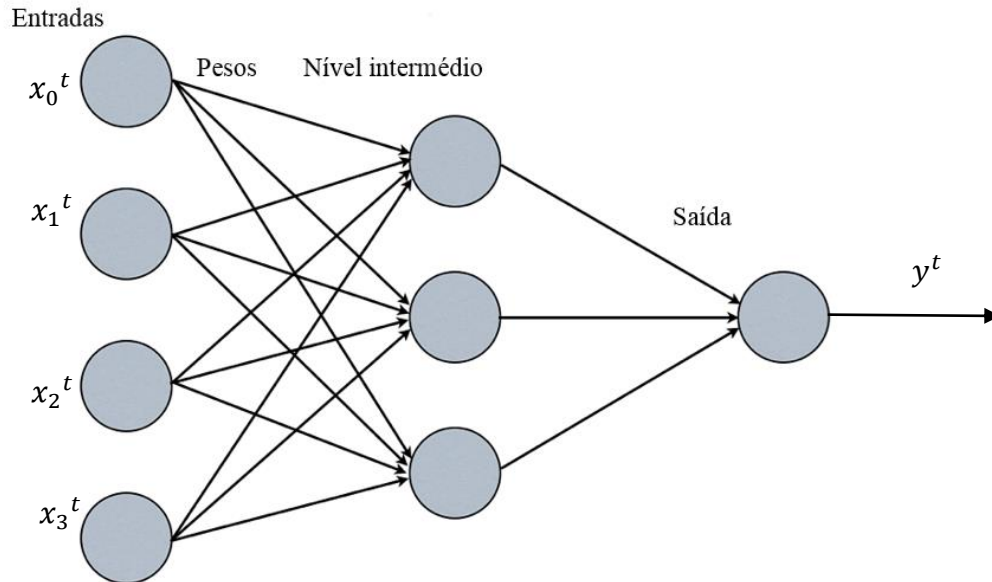


Figura 11 – Exemplo de uma ANN de 2 camadas²²

A forma de aprendizagem das ANNs pode ser supervisionada ou não supervisionada (Li, 1994). Na aprendizagem supervisionada há um agente externo que confere o quanto a rede está próxima de uma solução aceitável, adaptando no treino os pesos entre os neurónios para que a classificação desejada seja obtida. Pelo facto dos pesos serem ajustados durante o processo de aprendizagem, estes são os responsáveis pela memorização do padrão.

A aprendizagem não supervisionada não utiliza um agente externo, não possuindo conhecimento *à priori* das saídas da rede, funcionando assim de forma a distinguir classes de padrões diferentes apresentados à rede através de algoritmos de aprendizagem baseados em conceitos de vizinhança e agrupamento. Neste tipo de aprendizagem, a rede é ajustada de acordo com as regularidades estatísticas dos dados de entrada.

As ANNs possuem, portanto, a capacidade de aprender através exemplos. Um algoritmo de aprendizagem é um conjunto de procedimentos usados para adaptar os parâmetros da ANN e produzir uma saída esperada. A convergência do processo de aprendizagem supervisionada por

²² <http://www.cureus.com/articles/2260-classification-of-brain-metastases-prognostic-groups-utilizing-artificial-neural-network-approaches.html>

correção de erros é a diferença mínima entre a soma ponderada das entradas pelo peso e a saída desejada.

A forma genérica para alteração dos pesos por correção de erros é definida por:

$$w_i^{t+1} = w_i^t + \eta e^t x_i^t, \quad (1)$$

onde η é a taxa de aprendizagem, e x_i^t representa a entrada para o neurónio i no instante tempo t . O ajuste de pesos é feito para $w_i^{(t+1)}$, sendo w_i^t o peso atual.

O termo e^t representa o erro e é dado por:

$$e^t = d^t - y^t, \quad (2)$$

onde d^t é a saída esperada e y^t representa a resposta atual da ANN.

O ajuste de pesos w_i^t deve ser proporcional ao produto do erro entre a saída calculada pela rede e a saída esperada (Jain, Mao, & Mohiuddin, 1996; Cintra, Velho, & Todling, 2011).

Hoje em dia, as ANNs são muito utilizadas na análise e processamento de sinais, em robótica, no reconhecimento de padrões e na análise de voz. A classificação e reconhecimento de padrões é um exemplo de aplicação que tipicamente usa ANNs com aprendizagem supervisionada.

2.3.2. Redes Bayesianas

As Redes Bayesianas (Figura 12) são Grafos Direcionados Acíclicos (DAG), conectados e que representam dependências entre variáveis num modelo probabilístico (Heckerman, 1998).

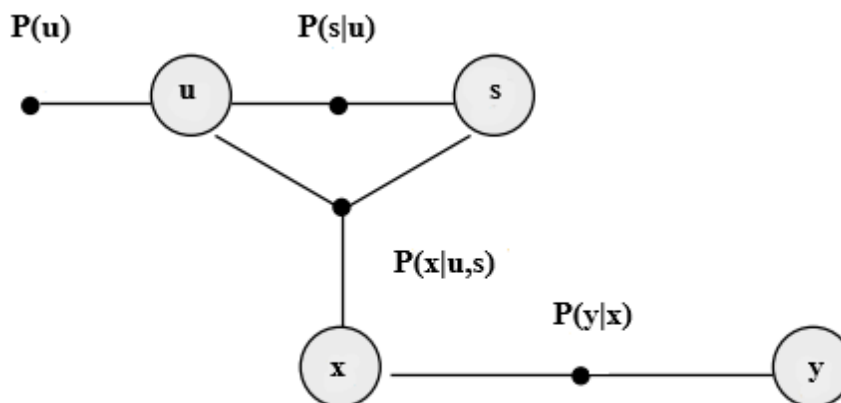


Figura 12 - Redes Bayesianas ²³

²³ <http://www.lx.it.pt/~afred/docencia/Percepcao/acetatos/rb.pdf>

Estas representam de forma simples as relações de causalidade das variáveis de um sistema. Os nós contêm a informação probabilística de cada evento e os arcos representam uma relação de dependência (condicional) entre esses eventos. Um exemplo de Rede Bayesiana é uma árvore genealógica, onde os nós representam as pessoas e os filhos recebem setas (arcos) dos seus pais.

Uma DAG é uma Rede Bayesiana se, por definição,

$$P(X_1, X_2, \dots, X_N) = \prod_{i=1}^N P(X_i | \text{Pais}(X_i)), \quad (3)$$

onde P é a probabilidade de X_i sabendo $\text{Pais}(X_i)$, X_i representa as variáveis de entrada e N o número de entradas. Se um nó i é fonte, ou seja não tem pais, então,

$$P(X_i | \text{Pais}(X_i)) = P(X_i). \quad (4)$$

Um elemento importante dentro da estrutura deste tipo de redes é a Tabela de Probabilidade Condicional (CPT) (da Cunha, 2009). Na CPT podemos visualizar os parâmetros de probabilidade condicional da variável condicionada a seu(s) pai(s).

As Redes Bayesianas são modelos de representação do conhecimento que trabalham com o conhecimento incerto e incompleto tendo por base o Teorema de Bayes (Ivansson, 2002). O Teorema de Bayes mostra a relação entre uma probabilidade condicionada e a sua inversa. Dados dois eventos A e B ,

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}, \quad (5)$$

sendo $P(A)$ e $P(B)$ as probabilidades *à priori* de A e de B , respetivamente. $P(A|B)$ e $P(B|A)$ representa a probabilidade condicionada (ou *à posteriori*) de A e B , respetivamente (Marques & Dutra, 2002).

O procedimento para construção de uma Rede Bayesiana é o que se segue:

1. Escolher um conjunto de variáveis X_i que descrevam o problema;
2. Escolher uma ordem para as variáveis;
3. Para todas as variáveis:
 - a) Escolher uma variável X_i e adicionar um nó na rede;
 - b) Determinar os nós $\text{Pais}(X_i)$ entre os nós que já estão na rede e que satisfaçam a equação:

$$P(X_i | X_j, \text{Pais}(X_i)) = P(X_i | \text{Pais}(X_i)); \quad (6)$$

- c) Definir a CPT para X_i .

O facto de que cada nó estar conectado aos nós mais antigos na rede, garante que o grafo será sempre acíclico. Uma das principais vantagens das Redes Bayesianas é o facto de permitirem

encontrar probabilidades *à posteriori*. Por outro lado, têm a desvantagem de que os algoritmos têm normalmente uma complexidade exponencial que é proporcional ao número de nós na rede (Silva R. , 2006).

As técnicas Bayesianas podem encontrar aplicações na área da estatística e da inteligência artificial.

2.3.3. Máquinas de Vetor de Suporte

Uma Máquina de Vetor de Suporte (*Support Vector Machine*, ou SVM) é um classificador não probabilístico no qual o reconhecimento de padrões é realizado através de aprendizagem supervisionada (Byun & Lee, 2003). Constrói um hiper-plano, ou um conjunto de hiper-planos, num espaço, onde se separam as classes de dados distintas (Figura 13). Diz-se que o SVM é um classificador linear binário não probabilístico quando os dados têm duas classes. Uma SVM classifica os dados ao encontrar o melhor hiper-plano que separa todos os pontos de uma classe de dados dos da outra classe. O melhor hiper-plano é aquele com a maior margem entre as duas classes (Osuna, Freund, & Girosi, 1997).

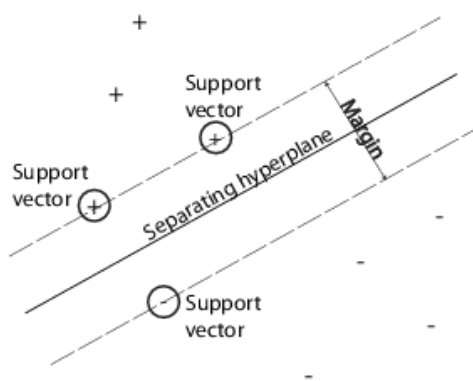


Figura 13 – Máquinas de Vetor de Suporte²⁴

Os vetores de suporte são os pontos de dados que estão mais próximos do hiper-plano de separação. Quando os dados têm mais do que duas classes, o classificador linear não é capaz de realizar uma boa separação das classes. De forma a contornar esta situação, é possível realizar um mapeamento do espaço de entrada para um espaço k -dimensional, onde as classes podem ser separadas por um hiper-plano. Para que este mapeamento seja possível, usa-se a função de *Kernel* (Ben-Hur & Weston, 2010). A escolha do *Kernel* é crucial para incorporar o conhecimento *à priori* sobre a aplicação.

²⁴ <http://www4.stat.ncsu.edu/~post/reading/scheduleFall2014>

Os primeiros *Kernel* a serem investigados para o problema de reconhecimento de padrões foram os seguintes: Linear, Polinomial, Gaussiano, Fisher (Ayat, Cheriet, & Suen, 2002; Hussain, Wajid, Elzaart, & Berbar, 2011). O Kernel Linear tem sido o mais usado e, como tal, será considerado neste trabalho.

As SVM têm várias áreas de aplicação tais como na categorização de textos, na análise de imagens e em Bioinformática.

2.4. Trabalhos Relacionados

Muitos autores têm vindo a estudar e a avaliar diferentes abordagens para detetar e reconhecer atividades diárias (Bao & Intille, 2004; Huynh T. , 2008; Huynh & Schiele, 2005; Suutala, Pirttikangas, & Roning, 2007).

O reconhecimento do movimento é muito importante em diferentes situações, como na área do desporto ou no reconhecimento de comportamentos suspeitos de pessoas em diferentes locais com grande afluência. Por exemplo, no futuro espera-se que as redes de câmaras instaladas em aeroportos possam, de forma inteligente, detetar e reconhecer automaticamente acontecimentos complexos e de risco envolvendo vários atores que interagem uns com os outros (Fusier, et al., 2007; Chen & Hauptmann, 2009). No desporto, o reconhecimento da atividade humana pode fornecer ao usuário um *feedback* sobre o seu estilo de vida em relação à atividade física e assim promover um estilo de vida mais ativo (Ermes & Parkka, 2008). As taxas de reconhecimento podem ser melhoradas através da seleção cuidadosa de *features* individuais para cada atividade. No entanto, tal seleção depende principalmente da tecnologia usada para adquirir os dados. Por exemplo, hoje em dia, a literatura explora muito o uso de sensores inerciais para análise do movimento humano (Vital, Couceiro, Dias, & Ferreira, 2015).

Em Lara e Labrador (2013), os autores fazem uma revisão sobre os diferentes métodos de extração de *features*, divididos em termos de domínio do tempo, domínio da frequência e outros. No domínio do tempo, os autores usam as *features* mais tradicionais, tais como a média, o desvio padrão, a variância, a variação interquartil (*Interquartile Range*, ou IQR), o desvio médio absoluto (*Mean Absolute Deviation*, ou MAD), a correlação entre eixos, a entropia e a curtose. No domínio da frequência, os autores referem a Transformada Rápida de Fourier (*Fast Fourier Transform*, ou FFT) e a Transformada Discreta do Cosseno (*Discrete Cosine Transform*, ou DCT), embora outros também tenham sido discutidos, tais como a análise das componentes principais (*Principal Component Analysis*, ou PCA), a análise linear discriminativa (*Linear Discriminant Analysis*, ou LDQA), o modelo autorregressivo (*Autoregressive Model*, ou AR) e os filtros de Haar.

Bao e Intille (2004) usaram várias *features*, tais como a média, a energia, a entropia, e a correlação. Os autores testaram vários classificadores, como a tabela de decisão, a aprendizagem instanciada (*Instance-based Learning*, ou IBL), a árvore de decisão C4.5, e o classificador Naïve Bayes (NB), que podem ser encontrados no *Weka Machine Learning*

Algorithms Toolkit ²⁵. Neste artigo, os autores concluíram que a árvore de decisão apresentou uma classificação com o valor de precisão mais elevado para o reconhecimento das atividades humanas escolhidas. Da mesma forma, o autor Huynh (2008) recorreu ao uso das mesmas *features*, demonstrando que estas são suficientes para reconhecer atividades básicas como “caminhar”, “levantar” ou “sentar”. Este autor escolheu como classificadores as SVM, o Modelo Escondido de Markov (*Hidden Markov Model*, ou HMM) e o classificador Naïve Bayes.

Para o reconhecimento da atividade humana, Zhang e Sawchuk (2013) usaram um telemóvel com uma tecnologia sensorial, considerando como *features* a média, a variância, a correlação e a entropia. Os autores concluíram que os piores métodos de classificação foram o *k* vizinhos mais próximos (*k-Nearest Neighbor*, ou *k-NN*) e o classificador NB. Como em muitos outros trabalhos, o SVM foi o método que apresentou os melhores resultados. O trabalho de Suutala *et al.* (2007) utilizou quatro acelerómetros localizados na coxa, no punho direito, no pulso esquerdo e no pescoço. Um conjunto de 17 atividades diárias foram consideradas pelos autores. Foram utilizadas diferentes janelas temporais, concluindo que com uma janela de tempo menor, e apenas utilizando *features* como a média e o desvio padrão, a taxa de reconhecimento foi mais elevada. Os autores concluíram também que o SVM foi o método que apresentou melhores resultados face ao uso do HMM.

Huynh e Schiele (2005) usaram sensores inerciais para classificar o movimento humano. Os autores escolheram como *features* a média, a variância, a energia, a entropia e a FFT. Os autores concluíram que a FFT foi a *feature* que apresentou melhores resultados a nível da precisão, mas que os seus valores são diferentes para cada atividade. Os autores concluíram também que o reconhecimento do movimento humano pode ser melhorado através da seleção de *features* separadamente para cada atividade.

Faria *et al.* (2015) refere outras *features* extraídas dos dados do sensor RGB-D, como o *Kinect v1*, que são modeladas de forma a caracterizar as atividades diárias, tais como *features* baseadas em energia usando a velocidade das juntas, o logaritmo da energia, a entropia baseada nas *features* usando a pose do esqueleto e a auto correlação baseada nas *features* que usam as distâncias das poses do esqueleto em diferentes instantes de tempo.

Khoshhal *et al.* (2010) apresentou as Redes Bayesianas como classificador da ação e comportamento em dados espaciais 3D. Os autores usaram como característica a FFT e o espectro de potência.

²⁵ <http://www.cs.waikato.ac.nz/ml/weka/>

A Tabela 6 apresenta uma súmula das *features* e dos métodos de classificação mais adotados na literatura. De realçar que todos os trabalhos citados usam soluções de aquisição *wearable*.

A Tabela 6 mostra que as *features* mais utilizadas no reconhecimento de atividades diárias são:

- **Média** (Bao & Intille, 2004; Huynh T. , 2008; Suutala, Pirttikangas, & Roning, 2007; Ermes & Parkka, 2008; Zhang & Sawchuk, 2013; Avci, Bosch, Marin-Perianu, & Havinga, 2010; Zhu & Sheng, 2009);
- **Variância** (Bao & Intille, 2004; Huynh T. , 2008; Suutala, Pirttikangas, & Roning, 2007; Ermes & Parkka, 2008; Zhang & Sawchuk, 2013; Avci, Bosch, Marin-Perianu, & Havinga, 2010; Zhu & Sheng, 2009);
- **Entropia** (Bao & Intille, 2004; Huynh T. , 2008; Zhang & Sawchuk, 2013; Avci, Bosch, Marin-Perianu, & Havinga, 2010; Hong, Kim, Ahn , & Kim , 2008);
- **FFT** (Avci, Bosch, Marin-Perianu, & Havinga, 2010; Bao & Intille, 2004; Zhang & Sawchuk, 2013; Huynh T. , 2008).

A Tabela 6 mostra que os métodos de classificação mais utilizados no reconhecimento de atividades diárias são:

- **SVM** (Huynh T. , 2008; Suutala, Pirttikangas, & Roning, 2007; Zhang & Sawchuk, 2013; Avci, Bosch, Marin-Perianu, & Havinga, 2010);
- **HMM** (Huynh T. , 2008; Suutala, Pirttikangas, & Roning, 2007; Zhang & Sawchuk, 2013; Avci, Bosch, Marin-Perianu, & Havinga, 2010; Zhu & Sheng, 2009);
- **k-NN** (Bao & Intille, 2004; Huynh T. , 2008; Zhang & Sawchuk, 2013; Avci, Bosch, Marin-Perianu, & Havinga, 2010);
- **NB** (Bao & Intille, 2004; Huynh T. , 2008; Zhang & Sawchuk, 2013; Avci, Bosch, Marin-Perianu, & Havinga, 2010);
- **ANN** (Avci, Bosch, Marin-Perianu, & Havinga, 2010; Ermes & Parkka, 2008; Zhu & Sheng, 2009).

A escolha das *features* tem uma enorme importância e é através destas que se pode otimizar o classificador de forma a melhorar a classificação do movimento humano. Apesar de diferentes métodos de classificação poderem apresentar diferentes níveis de exatidão e precisão no reconhecimento do movimento humano, é através da escolha cuidada das *features* que se pode impulsionar de forma significativa a precisão geral da classificação.

Tabela 6 – Features e métodos de classificação usados na literatura usando sensores inerciais

Classificadores Features	SVM	HMM	ANN	k-NN	NB	Árvore de Decisão	Tabela de Decisão	Modelo Híbrido	
<i>Média</i>	(Zhang & Sawchuk, 2013) (Suutala, Pirttikangas, & Roning, 2007) (Huynh T. , 2008) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Zhang & Sawchuk, 2013) (Huynh T. , 2008) (Avci, Bosch, Marin-Perianu, & Havinga, 2010) (Zhu & Sheng, 2009)	(Avci, Bosch, Marin-Perianu, & Havinga, 2010) (Zhu & Sheng, 2009) (Ermes & Parkka, 2008)	(Zhang & Sawchuk, 2013) (Bao & Intille, 2004) (Huynh T. , 2008) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Zhang & Sawchuk, 2013) (Bao & Intille, 2004) (Huynh T. , 2008) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Zhang & Sawchuk, 2013) (Bao & Intille, 2004) (Huynh T. , 2008) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Bao & Intille, 2004) (Avci, Bosch, Marin-Perianu, & Havinga, 2010) (Ermes & Parkka, 2008) (Hong, Kim, Ahn , & Kim , 2008)	(Bao & Intille, 2004) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Ermes & Parkka, 2008)
<i>Desvio Padrão</i>	(Suutala, Pirttikangas, & Roning, 2007) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Suutala, Pirttikangas, & Roning, 2007) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Avci, Bosch, Marin-Perianu, & Havinga, 2010)		
<i>Variância</i>	(Zhang & Sawchuk, 2013) (Huynh T. , 2008) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Huynh T. , 2008) (Avci, Bosch, Marin-Perianu, & Havinga, 2010) (Zhu & Sheng, 2009)	(Avci, Bosch, Marin-Perianu, & Havinga, 2010) (Zhu & Sheng, 2009) (Ermes & Parkka, 2008)	(Zhang & Sawchuk, 2013) (Huynh T. , 2008) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Zhang & Sawchuk, 2013) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Avci, Bosch, Marin-Perianu, & Havinga, 2010) (Ermes & Parkka, 2008)	(Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Ermes & Parkka, 2008)	
<i>Correlação entre eixos</i>	(Zhang & Sawchuk, 2013) (Huynh T. , 2008) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Huynh T. , 2008) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Zhang & Sawchuk, 2013) (Bao & Intille, 2004) (Huynh T. , 2008) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Zhang & Sawchuk, 2013) (Bao & Intille, 2004) (Huynh T. , 2008) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Bao & Intille, 2004) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Bao & Intille, 2004) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)		
<i>Entropia</i>	(Zhang & Sawchuk, 2013) (Huynh T. , 2008) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Huynh T. , 2008) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Zhang & Sawchuk, 2013) (Bao & Intille, 2004) (Huynh T. , 2008) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Zhang & Sawchuk, 2013) (Bao & Intille, 2004) (Huynh T. , 2008) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Bao & Intille, 2004) (Avci, Bosch, Marin-Perianu, & Havinga, 2010) (Hong, Kim, Ahn , & Kim , 2008)	(Bao & Intille, 2004) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)		
<i>Curtose</i>			(Ermes & Parkka, 2008)			(Ermes & Parkka, 2008)		(Ermes & Parkka, 2008)	
<i>FFT</i>	(Huynh T. , 2008) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Huynh T. , 2008) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Zhang & Sawchuk, 2013) (Huynh T. , 2008) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Huynh T. , 2008) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Avci, Bosch, Marin-Perianu, & Havinga, 2010) (Hong, Kim, Ahn , & Kim , 2008)	(Avci, Bosch, Marin-Perianu, & Havinga, 2010)		
<i>Energia</i>	(Huynh T. , 2008) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Huynh T. , 2008) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Bao & Intille, 2004) (Huynh T. , 2008) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Bao & Intille, 2004) (Huynh T. , 2008) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)	(Bao & Intille, 2004) (Avci, Bosch, Marin-Perianu, & Havinga, 2010) (Hong, Kim, Ahn , & Kim , 2008)	(Bao & Intille, 2004) (Avci, Bosch, Marin-Perianu, & Havinga, 2010)		

2.5. Sumário

Neste capítulo foi realizado um levantamento das diferentes tecnologias usadas para a captura do movimento humano e dos diferentes métodos usados para a classificação do movimento humano. Existem vários tipos de tecnologias, que vão desde câmaras até ao uso de sensores colocados no próprio sujeito. Todas as tecnologias apresentam vantagens e desvantagens, porém as tecnologias que usam câmaras apresentam mais desvantagens, tais como, a infraestrutura, a iluminação, o uso exclusivo em contexto laboratorial e a oclusão de juntas. Hoje em dia, o uso de sensores inerciais é mais comum, tendo estes a grande vantagem de se poderem usar em qualquer local e não terem problemas de oclusão de juntas.

Existem vários métodos de classificação, tais como SVM, ANNs e NB aplicados ao reconhecimento do movimento humano. A seleção de *features* a serem usadas pelos métodos de classificação é um passo muito importante. As mais adotadas são a média, o desvio padrão, e a FFT.

ARQUITECTURA PARA CLASSIFICAÇÃO DO MOVIMENTO HUMANO

Neste capítulo é descrita a arquitetura proposta para a classificação do movimento humano.

Na Secção 3.1 é feita uma breve descrição acerca do pré-processamento dos dados. Na Secção 3.2 são apresentadas as *features* utilizadas. Na Secção 3.3 é descrita a arquitetura geral do método de classificação usado, o qual é baseado na combinação probabilística de três métodos de classificação descritos anteriormente. Na Secção 3.4 é feito um breve sumário do capítulo.

3.1. Pré-Processamento

De forma a classificar as diferentes atividades é preciso gerar uma série temporal que representa a atividade ao longo do tempo de todos os ensaios. Por outras palavras, depois de obter os dados de cada um dos 20 ensaios é necessário concatenar os dados necessários.

Figura 14 exemplifica 20 ensaios de um indivíduo de uma determinada atividade concatenada numa série temporal.

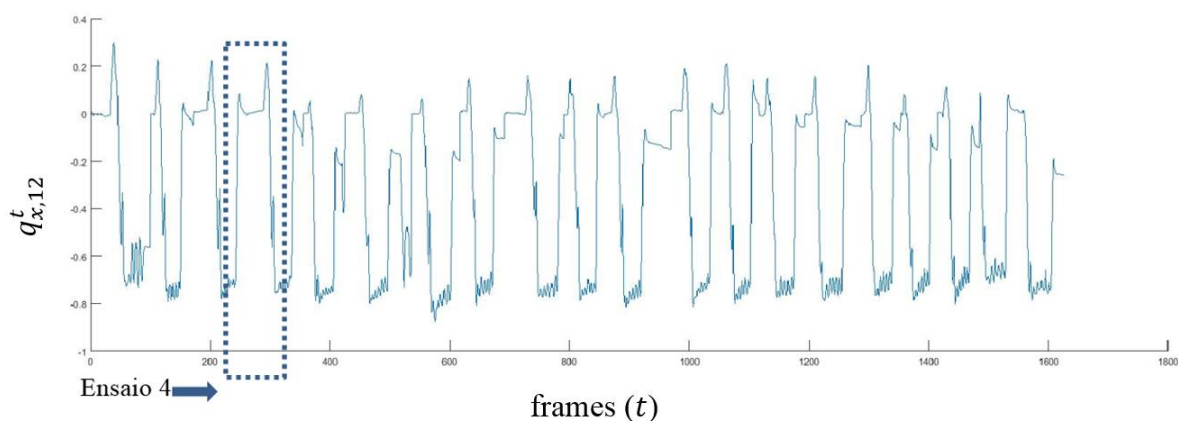


Figura 14 – Exemplo da concatenação de 20 ensaios

3.2. Features

Esta secção descreve o conjunto de *features* que foram utilizadas neste trabalho para classificar atividades diárias recorrendo ao uso de tecnologia *wearable*. Estas *features* dependem da existência de relações entre as diferentes partes do corpo monitorizadas (*e.g.*, juntas). As *features* foram extraídas a partir de um conjunto de 17 sensores IMU integrados no *FatoXtract*, dispostos como mostra a

Figura 10, sendo baseadas em ângulos de euler e em quatérnios.

Um quatérnio q_n no instante de tempo t , q_n^t , pode ser representado por:

$$q_n^t = (q_{x,n}^t i, q_{y,n}^t j, q_{z,n}^t k, q_w^t), \quad n = \{1, \dots, 17\}, \quad (7)$$

onde i , j e k designam a base ortogonal para o espaço 3D e w é a parte real.

Considerando os 17 IMUs do fato com quatro quatérniões por junta, obtém-se:

$$\text{número de features (quatérniões)/ frame} = 4 * 17 = 68. \quad (8)$$

Um ângulo euler θ_n no instante de tempo t , θ_n^t , pode ser representado por:

$$\theta_n^t = (\theta_{x,n}^t, \theta_{y,n}^t, \theta_{z,n}^t), \quad (9)$$

o qual fornece as suas rotações em relação a x , y e z para cada junta do corpo humano. De forma semelhante, para 17 IMUs e três ângulos de euler por junta, obtém-se:

$$\text{número de features (eulers)/ frame} = 3 * 17 = 51. \quad (10)$$

Tendo em conta os ângulos de euler e os quatérniões de cada sensor IMU em cada *frame*, um conjunto de diferentes *features* F é definido compreendendo seis subgrupos diferentes

$$F = \{f_{s1}, f_{s2}, f_{s3}, f_{s4}, f_{s5}, f_{s6}\}.$$

É de notar que esta divisão em subgrupos de *features* serve meramente para facilitar a descrição com base no tipo que estas representam.

O primeiro subconjunto de *features* f_{s1} é baseado em *features* geométricas onde os ângulos de euler e os quatérniões são usados diretamente (Secção 3.1.1). O segundo subconjunto de *features* f_{s2} é baseado em informações temporais e primeira derivada (*i.e.*, velocidade angular) (Secção 3.1.2). O terceiro subconjunto f_{s3} é baseado em *features* da segunda derivada (*i.e.*, aceleração angular) (Secção 3.1.3). O quarto subconjunto f_{s4} é constituído pela análise da componente principal numa matriz de covariância (Secção 3.1.4). O quinto subconjunto f_{s5} utiliza a energia das *features* dos subconjuntos anteriores, onde é calculada a diferença quadrática entre as *frames*, *i.e.*, iterações ou instantes de tempo, e é acumulando esse valor ao longo do tempo (Secção 3.1.5). Por fim, o sexto subconjunto f_{s6} é baseado em *features* no domínio da frequência, nomeadamente a FFT (Secção 3.1.6).

A Tabela 7 apresenta um resumo das *features* usadas neste trabalho e cuja obtenção será descrita nas secções que se seguem.

Tabela 7 - *Features* usadas neste trabalho

Subconjunto	<i>Features</i>
f_{s1}	Básicos
	Log-Covariância
f_{s2}	Velocidade
	Média da Velocidade
	Desvio Padrão da Velocidade
f_{s3}	Aceleração
f_{s4}	Energia da Autocorrelação
	Energia da Autocorrelação da Velocidade
	Energia da Autocorrelação da Aceleração
f_{s5}	Básicos VA ²⁶
	Média da Velocidade VA
	Desvio Padrão da Velocidade VA
	Log-Covariância VA
	Velocidade VA
	Aceleração VA
f_{s6}	FFT VA

3.2.1. Ângulos de Euler e Quaterniões (f_{s1})

O primeiro subconjunto f_{s1} considera os ângulos de euler e os quaterniões medidos por cada um dos 17 IMUs do fato. De acordo com as equações (8) e (10) obtém-se 119 *features* (68 dos quaterniões e 51 dos ângulos de euler).

Duas matrizes foram construídas: uma contendo os quaterniões para as 17 juntas, formando uma matriz M_q (17×4):

²⁶ Valor Acumulado

$$M_q^t = \begin{bmatrix} q_{x,0}^t & q_{y,0}^t & q_{z,0}^t & q_{w,0}^t \\ q_{x,1}^t & q_{y,1}^t & q_{z,1}^t & q_{w,1}^t \\ \vdots & \vdots & \vdots & \vdots \\ q_{x,17}^t & q_{y,17}^t & q_{z,17}^t & q_{w,17}^t \end{bmatrix}_{(17 \times 4)}, \quad (11)$$

e outra contendo os ângulos de euler, formando assim uma matriz M_e 17×3 :

$$M_e^t = \begin{bmatrix} \theta_{x,0}^t & \theta_{y,0}^t & \theta_{z,0}^t \\ \theta_{x,1}^t & \theta_{y,1}^t & \theta_{z,1}^t \\ \vdots & \vdots & \vdots \\ \theta_{x,17}^t & \theta_{y,17}^t & \theta_{z,17}^t \end{bmatrix}_{(17 \times 3)}. \quad (12)$$

Posteriormente, foi aplicada a log-covariância a cada matriz para obter uma matriz 4×4 para M_q e uma matriz 3×3 para M_e , da seguinte forma:

$$lcM_\kappa = U(\log(\text{cov}(M_\kappa)))^2, \kappa = \{e, q\}, \quad (13)$$

onde lcM_κ é um vetor contendo os elementos da componente triangular superior da matriz do logaritmo da covariância da matriz M_κ , sendo determinado, para os quaterniões, da seguinte forma:

$$lcM_q = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{bmatrix}_{(4 \times 4)} = [a \ b \ c \ d \ f \ g \ h \ k \ l \ p]_{(1 \times 10)}, \quad (14)$$

E, para os ângulos de euler, como:

$$lcM_e = \begin{bmatrix} q & r & s \\ t & u & v \\ x & z & w \end{bmatrix}_{(3 \times 3)} = [q \ r \ s \ u \ v \ w]_{(1 \times 6)}, \quad (15)$$

A lógica por detrás da log-covariância é o mapeamento do cone convexo de uma matriz de covariância para o espaço vetorial usando o logaritmo da matriz, como proposto por Arsigny *et al.* (2006). Uma matriz de covariância forma um cone convexo, de modo que não se encontra no espaço euclidiano, isto é, o espaço da matriz covariância não é fechado sob multiplicação de escalares negativos. A ideia de log-covariância é baseada no trabalho de Guo (2012), em que foram utilizados métricas de Riemannian múltiplas e log-covariância aplicadas em *features* de imagem 2D para o reconhecimento de atividades. Aqui, neste trabalho, considera-se o uso da log-covariância como um conjunto de *features* aplicadas nas rotações das articulações humanas. Assim, com a log-covariância empregue em ângulos de euler e quaterniões, obtêm-se 16 *features* adicionais (6 dos euler e 10 dos quaterniões).

3.2.2. Velocidade Angular (f_{s2})

O segundo subconjunto f_{s2} considera a variação angular das juntas ao longo do tempo. A velocidade angular dos quaterniões v_q representa a rotação das juntas do corpo e pode ser calculada da seguinte forma:

$$v_q = \frac{M_q^t - M_q^{t-1}}{\Delta T}, \quad (16)$$

onde o intervalo de amostragem ΔT depende da taxa de aquisição escolhido, aqui definido como $fr = 5 \text{ Hz}$, o que perfaz $\Delta T = \frac{1}{fr} = 0.2 \text{ s}$. A matriz v_q representa a velocidade angular usando a variação dos quaterniões ao longo do tempo. Tendo por base a Equação (8), obtém-se assim 68 *features* adicionais.

O mesmo procedimento pode ser realizado utilizando ângulos de euler em vez de quaterniões. No entanto, em vez de utilizar diretamente as velocidades dos ângulos de euler, que, em certa medida, podem ser redundantes com a velocidade angular dos quaterniões dada a relação que existe entre ambos, calculámos dois parâmetros estatísticos para representar a distribuição de rotações sobre o corpo humano: a média (μ) e o desvio padrão (σ). Para considerar a relação temporal de cada velocidade angular individual, o valor de μ e de σ de cada velocidade angular foram obtidos pela acumulativa integração das velocidades angulares de cada articulação do corpo sobre *frames* consecutivas. As velocidades angulares de euler representam a pose do corpo humano, em que os dois parâmetros (μ, σ), a partir de uma distribuição normal, foram obtidos como uma estimativa de probabilidade máxima (*Maximum-Likelihood Estimation* ou MLE) para aproximar a distribuição das rotações das juntas do corpo. Afim de contar com estes parâmetros estatísticos (μ, σ), torna-se necessário ter um bom estimador, sendo que o MLE tem sido comumente utilizado como tal (Uddin, Thang, Kim, & Kim, 2011). No entanto, pode ser parcial, o que significa que o valor esperado do parâmetro pode não ser igual ao parâmetro a ser estimado, por exemplo, ao estimar a variância de uma distribuição normal. Um estimador imparcial é o estimador de variância imparcial mínimo, o que tem a variância mínima de todos os estimadores de um parâmetro.

Da Equação (16) obtém-se 51 *features* para μ e 51 para σ , o que perfaz 102 *features* adicionais.

3.2.3. Aceleração Angular (f_{s3})

O terceiro subconjunto f_{s3} contém *features* baseadas na segunda derivada de M_q , a_{vq} , dada por:

$$a_{vq} = \frac{v_q^t - v_q^{t-1}}{\Delta T}, \quad (17)$$

Estas *features* foram obtidas usando o mesmo método descrito na Secção 3.2.2, com a diferença de que, em vez de se ter considerado os quaterniões como entradas, foi considerada a velocidade angular dos quaterniões. Tendo por base a Equação (8), resultam assim 68 *features* adicionais.

3.2.4. Análise da Componente Principal (f_{s4})

O quarto subconjunto f_{s4} é constituído pela aplicação da análise da componente principal. A análise da componente principal (*Principal Component Analysis*, ou PCA) é uma maneira de identificar a relação entre as *features* extraídas dos dados. A aplicação da PCA numa matriz de covariância (identificada como PCA-cov) é capaz de reduzir a dimensão dos dados, com o propósito de extrair a informação discriminativa da pose e do movimento do corpo humano (Jackson, 2005; Jolliffe, 2002; Krzanowski, 2000; Seber, 2009). Essencialmente, a PCA de uma matriz de covariância $n \times n$ retorna os coeficientes, também conhecidos como pesos (*loadings*), representado por uma matriz P ($n \times n$), em que cada coluna contém os coeficientes para uma componente principal. As colunas são classificadas segundo a ordem decrescente da variância da componente.

Neste trabalho, usaram-se os ângulos de euler (representado como uma matriz M_e de 17×3), os quaterniões (representados por uma matriz M_q de 17×4) e derivadas dos ângulos de euler e quaterniões, também representados como matrizes de 17×3 e 17×4 , respetivamente.

Dada uma matriz de entrada M genérica, primeiro calcula-se a correlação cruzada 2D de dois sinais, neste caso entre M em diferentes instantes de tempo, por exemplo, M^t e M^{t-1} , em que cada elemento (i, j) de uma matriz resultante C é calculado como se segue:

$$C_{i,j} = \sum_k^K \sum_l^L M_{k,l}^t \times M_{k-i,l-j}^{t-1},$$

$$\text{com } 0 \leq i \leq k + p - 1$$

$$0 \leq j \leq l + q - 1$$
(18)

em que k e l são índices de M^t , p e q são índices de M^{t-1} , e i, j são índices da matriz resultante C do tamanho $k + p - 2$ por $l + q - 2$. É de notar que $k = p$ e $l = q$ dado que a dimensão da matriz M não varia ao longo do tempo, sendo que os índices i, j acabam por variar entre valores negativos e positivos.

A covariância de C é calculada através de:

$$cov_{i,j} = \frac{1}{N} \sum_{k=1}^N (C_{i,k} - \bar{X}_i)(C_{k,j} - \bar{X}_j),$$
(19)

sendo a entrada da PCA-cov. Em seguida, o modelo de energia ao longo dos elementos da matriz resultante cov ($n \times n$) para o instante t é calculada como se segue:

$$E_{cov_{i,j}}^t = \sum_{i=0}^{k+p-2} \sum_{j=0}^{l+q-2} cov_{i,j}^2.$$
(20)

Assim, obtêm-se 4 *features* a partir da energia da PCA-cov sobre a correlação cruzada dos quaterniões, da velocidade dos ângulos de euler, da velocidade dos quaterniões e da aceleração dos ângulos de euler.

3.2.5. Valores Acumulativos (f_{s5})

O quinto subconjunto f_{s5} consiste no modelo de energia utilizado como um valor acumulado de cada *feature* calculado de forma recursiva em cada *frame*. Ou seja, cada *feature* na *frame* t é o valor acumulado Y_i^t (*i.e.*, energia) calculada desde a *frame* anterior até à *frame* atual, como é representado em baixo:

$$\begin{cases} Y_i^t = y_i^t & , t = 1 \\ Y_i^t = \sum_{k=2}^t (y_i^k - y_i^{k-1})^2 & , t > 1 \end{cases}, i = \{1, \dots, \psi\}, \quad (21)$$

onde y_i^t representa o conjunto de *features*, no instante de tempo t , nomeadamente:

- os ângulos de euler e os quaterniões ($51 + 68 = 119$ *features*);
- a velocidade e a aceleração dos quaterniões ($68 + 68 = 136$ *features*);
- a média e o desvio padrão da velocidade dos ângulos de euler ($51 + 51 = 102$ *features*);
- a log-covariância dos euler e dos quaterniões ($6 + 10 = 16$ *features*).

ψ representa a dimensão de y_i que, neste caso, corresponde a 373 *features*.

3.2.6. FFT (f_{s6})

O sexto subconjunto f_{s6} consiste no uso da FFT. A FFT incide na análise do espectro de uma série temporal, que neste caso é representada pela amplitude dos ângulos de euler. Considerando (9), a Equação (22) calcula a amplitude angular para a junta n no instante t :

$$A(\theta_n^t) = \sqrt{\theta_{x,n}^t{}^2 + \theta_{y,n}^t{}^2 + \theta_{z,n}^t{}^2}. \quad (22)$$

A FFT é calculada para os ângulos de euler pela equação (23),

$$Y_n^k = \sum_{t=1}^T A(\theta_n^t) e^{-\frac{i2\pi}{T}(t-1)(k-1)}, \quad 0 \leq k < f_H, \quad (23)$$

em que, neste caso, f_H corresponde à frequência máxima do movimento de 22 Hz, com uma discretização de 0.2 Hz.

Ao calcular a FFT para ângulos de Euler de cada junta obtém-se a matriz Y_n^k . De seguida, é calculado o valor acumulado de Y_n^k e obtém-se 1870 *features* adicionais por cada *frame* t .

A Figura 15 é constituída por quatro gráficos, os quais demonstram a aplicação da FFT ao longo do tempo de uma determinada atividade. A atividade aqui considerada é a de lavar os dentes, ilustrando apenas a junta do pulso. Esta atividade (ver Secção 3.1) tem um total de 1626 *frames* (*i.e.*, 325 segundos). Na Figura 15 podemos visualizar a aplicação da FFT na *frame* 100, 500, 1000 e 1500, podendo-se assim verificar que, ao longo do tempo, obtém-se informação mais

detalhada sobre a atividade, apresentando assim uma variabilidade incremental de sinais em frequências distintas.

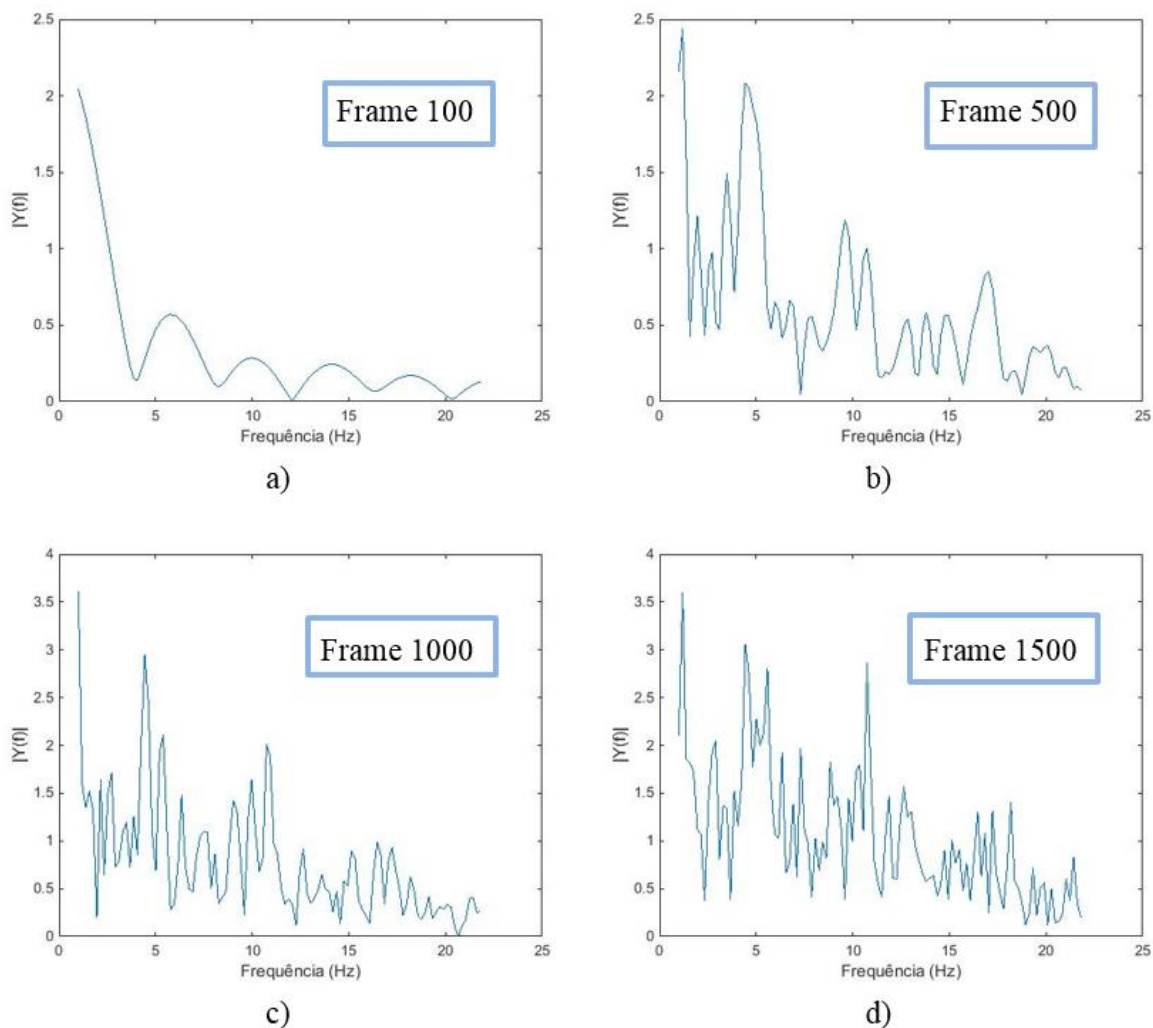


Figura 15 - FFT nas frames: a) 100; b) 500; c) 1000; d) 1500

Com o intuito de resumir as secções anteriores, a Tabela 8 apresenta as 2620 *features* consideradas neste trabalho para a classificação do movimento humano, destacando os dados que serviram como base para cálculo das mesmas (ângulos de euler e quaterniões). Como é possível verificar das 2620 *features* utilizadas, 2190 *features* são extraídas com base em ângulos de euler e 430 *features* com base em quaterniões.

Tabela 8 - Descrição do número de *features* usadas por frame

Subconjunto	Features	Eulers	Quaterniões
f_{s1}	<i>Básicos</i>	51	68
	<i>Log-Covariância</i>	6	10
f_{s2}	<i>Velocidade</i>	-	68
	<i>Média da Velocidade</i>	51	-
	<i>Desvio Padrão da Velocidade</i>	51	-
f_{s3}	<i>Aceleração</i>	-	68
f_{s4}	<i>Energia da Autocorrelação</i>	-	1
	<i>Energia da Autocorrelação da Velocidade</i>	1	1
	<i>Energia da Autocorrelação da Aceleração</i>	1	-
f_{s5}	<i>Básicos VA²⁷</i>	51	68
	<i>Média da Velocidade VA</i>	51	-
	<i>Desvio Padrão da Velocidade VA</i>	51	-
	<i>Log-Covariância VA</i>	6	10
	<i>Velocidade VA</i>	-	68
	<i>Aceleração VA</i>	-	68
f_{s6}	<i>FFT VA</i>	1870	-
	Total	2190	430

Após a extração de *features*, o passo seguinte incide na classificação dos movimentos registados. A fusão de diferentes métodos de classificação com diferentes *features* tem como objetivo conduzir a uma classificação com desempenho superior ao que cada método de classificação individual poderia conduzir. Neste trabalho, são conjugados três métodos: NB, SVM e ANN.

3.3. Arquitetura Geral

Com o propósito de aumentar o desempenho da classificação, alguns autores utilizam uma combinação probabilística de diferentes classificadores. O método de fusão dinâmica Bayesiana de modelos (*Dynamic Bayesian Mixture Models*, ou DBMM), introduzido em (Faria,

²⁷ VA- Valor Acumulado

Premebida, & Nunes, 2014) e (Faria, Vieira, Premebida, & Nunes, 2015), beneficia do conceito de mistura de classificadores de forma dinâmica, de modo a combinar as saídas das probabilidades condicionais dos diferentes classificadores. Esta secção apresenta uma breve descrição da classificação DBMM. Para uma descrição mais detalhada, consulte Faria, Premebida e Nunes (under review).

De forma sucinta, no DBMM é atribuído um peso a cada classificador, de acordo com o conhecimento prévio (processo de aprendizagem), com uma medida de incerteza com um nível de confiança, a qual pode ser atualizada localmente durante a classificação *online*. A atualização do peso atribuído prioridade a cada classificador com mais confiança ao longo do tempo, uma vez que podem variar ao longo das diferentes classificações por *frame*.

A Figura 16 demonstra a arquitetura DBMM adotada para o reconhecimento de atividades, onde os classificadores base são integrados como distribuições *à posteriori* ponderadas. O DBMM assume uma propriedade de Markov de primeira ordem que inclui a informação temporal como um ciclo probabilístico dinâmico simples. No entanto, de forma semelhante ao modelo escondido de Markov (*Hidden Markov Model*, ou HMM), o modelo aceita a matriz de probabilidade de transição de estado.

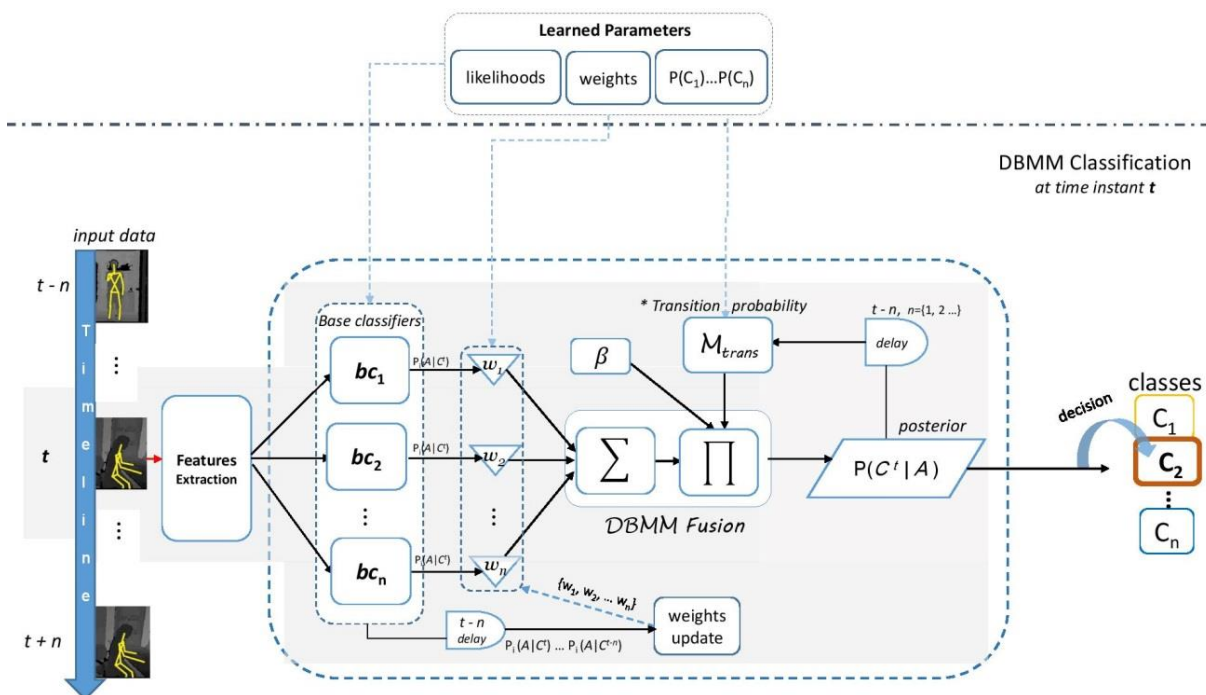


Figura 16 - Modelo DBMM

A Figura 16 ilustra a arquitetura geral DBMM para a classificação no instante de tempo t . O conjunto de classificadores probabilísticos dinâmicos considera a probabilidade Bayesiana como meio de classificação, em que cada modelo contribui para o resultado da inferência na

proporção da sua probabilidade. Posto de outro modo, o modelo de fusão é apresentado diretamente como uma soma ponderada das distribuições, onde é possível obter a combinação de diferentes modelos num só.

O DBMM compreende assim um conjunto de modelos $A = \{A_m^1, A_m^2, \dots, A_m^T\}$, onde A_m^t é um modelo com m atributos, por exemplo, variáveis ou *features* espaço-temporais observadas, num instante de tempo $t = \{1, 2, \dots, T\}$. O modelo geral do DBMM é dado por:

$$P(C^t|A) = \underbrace{\beta \times \mathcal{M}_{trans}}_{\text{Transações Dinâmicas}} \times \underbrace{\sum_{i=1}^N w_i^t \times P_i(A|C^t)}_{\text{Modelo com } w \text{ dinâmico}}, \quad (24)$$

Onde:

- \mathcal{M}_{trans} é o modelo para a probabilidade da transição de estado entre as variáveis classe, ou estados, ao longo do tempo. Este modelo pode representar a *a priori* no DBMM, como um ciclo probabilístico dinâmico, onde a *a posteriori* atual se torna a nova *a priori*, ou pode representar a matriz de transição de estado;
- $P_i(A|C^t)$ é o resultado a *a posteriori* para o classificador base i no instante de tempo t , que se torna a probabilidade i do modelo de mistura, com $i = \{1, \dots, N\}$ e N sendo o número total de classificadores considerados no modelo;
- O peso w_i^t é estimado usando uma medida de confiança baseada na entropia (ver próxima secção);
- $\beta = \frac{1}{\sum_j (P(C_j^t|C_j^{t-1}) \times \sum_{i=1}^N w_i^t \times P_i(A|C_j^t))}$ é o fator de normalização, necessário devido à contínua atualização da confiança.

Neste trabalho, calcula-se \mathcal{M}_{trans} como uma informação dinâmica *a priori* do modelo. Em seguida, a função geral da distribuição de probabilidade do DBMM, $P(C^t|A) = \prod_{t=1}^T P(C^t|C^{t-1}) \times \sum_{i=1}^N w_i^t \times P_i(A|C^t)$, é calculada para cada classe C , assumindo a propriedade Markov, tomando a *a posteriori* do instante de tempo anterior $t - 1$ como o *a priori* para o instante de tempo presente t , podendo formalizar o DBMM reescrevendo o modelo geral representado na Equação (24) da seguinte forma:

$$P(C^t|A) = \underbrace{\beta \times P(C^t|C^{t-1})}_{\text{Transações Dinâmicas}} \times \underbrace{\sum_{i=1}^N w_i^t \times P_i(A|C^t)}_{\text{Modelo com } w \text{ dinâmico}}, \quad (25)$$

$$\text{com } \begin{cases} P(C^t|C^{t-1}) = \frac{1}{c} (\text{uniform}), & t = 1 \\ P(C^t|C^{t-1}) = P(C^{t-1}|A), & t > 1 \end{cases},$$

onde $P(C^t|C^{t-1})$ representa a distribuição de probabilidade de transição entre as variáveis de classe ao longo do tempo. A classe C no tempo t está condicionada à classe no tempo $t - 1$. Este passo descreve o comportamento não estacionário aplicado de forma recursiva, onde o posterior do tempo anterior de cada classe torna-se o corrente prior para reforçar a classificação no instante de tempo t , usando a informação do tempo $t - 1$. Este modelo geral é o único testado em datasets do estado-da-arte para classificação temporal do movimento humano, tal como apresentado na secção de resultados experimentais.

3.3.1. Atribuição de Pesos para o Modelo de Fusão usando a Entropia

Na estrutura do DBMM, $H_i(\mathcal{L})$ é visto como um nível de confiança utilizado para atribuir pesos w_i^t aos classificadores de base e , consequentemente, para atualizar o modelo probabilístico global. Os pesos são calculados com base na entropia para cada classificador base, $H_i(\mathcal{L})$, através da análise dos seus resultados anteriores.

É possível calcular a entropia das probabilidades posteriores previamente observadas da seguinte forma:

$$H_i(\mathcal{L}) = - \sum_j^s \mathcal{L}_j \times \log(\mathcal{L}_j), \quad (26)$$

onde \mathcal{L} representa um conjunto com probabilidades condicionais $P_i(C|A)$ dadas pelo classificador base i (*e.g.*, os seus posteriores), e j é o índice para o conjunto de posteriores de um classificador base i específico. Sabendo H_i , o peso w_i^t para cada classificador base i é obtido por dois passos, sendo necessário, num primeiro passo, calcular o valor global do peso w_i^t :

$$\forall w_i, w_i^t = \left[1 - \left(\frac{H_i}{\sum_{i=1}^N H_i} \right) \right], \quad (27)$$

onde $H_i \equiv H_i(\mathcal{L})$ é o corrente valor de entropia dado por (26). De seguida, o peso w_i^t é normalizado da seguinte forma:

$$w_i^t = \frac{w_i^t}{\sum_i w_i^t}. \quad (28)$$

Este segundo passo garante que $\sum_i w_i = 1$.

Durante o processo de classificação, os classificadores base podem alterar o desempenho ao longo do tempo (*i.e.*, *frame a frame*). Assim, uma atualização local dos pesos durante a classificação irá produzir uma convicção superior, atribuindo prioridade aos classificadores de base com mais confiança sobre a classificação anterior. Assumindo uma memória do sistema como propriedade de Markov durante a classificação *online*, obtém-se uma informação temporal a partir do conjunto de posteriores para cada classificador base $\mathcal{L}_{\{i \dots s\}} = \{P(C_i^t|C_i^{t-1}); P(C_i^{t-1}|C_i^{t-2}); P(C_i^{t-2}|C_i^{t-3}) \dots P(C_i^{t-s}|C_i^{t-(s-1)})\}$. Esta informação é usada

juntamente com os pesos no instante de tempo anterior w_i^{t-1} de forma a actualizar os pesos de cada classificador base i na classificação de cada *frame*, como se segue:

$$w_i^t = \frac{w_i^{t-1} \times P(w_{i_{new}} | H_i(\mathcal{L}))}{\sum_{i=1}^n w_i^{t-1} \times P(w_{i_{new}} | H_i(\mathcal{L}))}, \quad (29)$$

onde w_i^t é o peso estimado que é actualizado por cada classificador base em cada instante de tempo e w_i^{t-1} é dado pelo peso anterior calculado em $t - 1$. A fim de se obter $P(w_{i_{new}} | H_i(\mathcal{L}))$, calcula-se (27) e (28) sobre o novo conjunto de posteriores de base de cada classificador base.

3.3.2. Classificadores de Base para a Fusão DBMM

Diferentes classificadores base podem ser usados no DBMM uma vez que as saídas são convertidas em probabilidades (*i.e.*, estimativas posteriores). Neste trabalho, utilizaram-se os seguintes classificadores apresentados na secção 2.3, nomeadamente o NB, SVM, e ANN.

O classificador NB assume que as *features* são independentes entre si, dada a variável classe. Assim, uma função de densidade de probabilidade (*probability density function*, ou pdf) para cada modelo das *features* foi considerada, obtendo a seguinte expressão:

$$P(C_i | A) = \alpha P(C_i) \prod_{j=1}^m P(A_j | C_i), \quad (30)$$

onde $\alpha = \frac{1}{\sum_i P(A | C_j) P(C_i)}$ é o fator de normalização e m é o número de modelos independentes das *features*. Em seguida, a inferência é geralmente obtida através da utilização da estimativa máxima *a posteriori* (*maximum a posteriori*, ou MAP). Para o classificador NB, um conjunto de parâmetros válidos de distribuições de probabilidade podem ser testados usando o critério de informação Bayesiano (*Bayesian Information Criterion*, ou BIC) (Schwarz, 1978) para seleccionar a melhor distribuição para os dados, como descrito em (Faria, Premevida, & Nunes, 2014).

Um exemplo de um classificador discriminativo usado no DBMM é uma multiclasse SVM que utiliza um Kernel linear. Detalhes sobre a teoria dos SVM e a sua implementação podem ser encontrados em (Cortes & Vapnik, 1995) e (Vapnik, 1998).

Neste trabalho, para a implementação do Kernel linear no SVM multiclasse, adotou-se por beneficiar da biblioteca *LibSVM* (Chang & Lin, 2011). Os SVMs foram treinados de acordo com a estratégia ‘um contra um’ (*one-vs-one*), com margem suave, parâmetro definido como 1.0, e as saídas da classificação foram dadas em termos das estimativas de probabilidade.

Usou-se também o método de classificação ANN nesta fusão DBMM, usando-se a *MatLab Neural Network toolbox*²⁸. No caso da ANN, a função de transferência da camada oculta é uma função de transferência sigmoide tangente (*tansig*), e a função de transferência de saída é uma normalizada exponencial (*softmax*). As saídas podem ser interpretadas como estimativas de probabilidade posterior da classe, condicionadas na entrada, como $\sum_{i=1}^n P(C_i|x) = 1$.

3.4. Sumário

A arquitetura DBMM implementada para a classificação do movimento humano, baseada no método DBMM, utilizou as seguintes *features*:

- Ângulos de euler e quaterniões;
- Velocidade angular dos quaterniões;
- Aceleração dos quaterniões;
- Média da velocidade dos ângulos de euler;
- Desvio padrão da velocidade dos ângulos de euler;
- Log-covariância dos ângulos de euler e dos quaterniões;
- Energia da PCA-cov:
 - Quaterniões;
 - Velocidade dos ângulos de euler e dos quaterniões;
 - Aceleração dos ângulos de euler;
- Valores acumulados:
 - FFT dos ângulos de euler;
 - Ângulos de euler e quaterniões;
 - Média da velocidade dos ângulos de euler;
 - Desvio padrão da velocidade dos ângulos de euler;
 - Log-covariância dos ângulos de euler e dos quaterniões;
 - Velocidade dos quaterniões;
 - Aceleração dos quaterniões.

Esta, por outro lado, considerou a fusão dos métodos de classificação ANN, SVM, e NB.

²⁸ <http://www.mathworks.com/products/neural-network/>

RESULTADOS EXPERIMENTAIS

Neste capítulo são apresentados os resultados experimentais e está dividido em cinco secções.

Na Secção 4.1 é descrito o setup experimental, de seguida é feita uma descrição do procedimento realizado para as diferentes atividades adquiridas (Secção 4.2), e por fim é feita uma descrição da amostra (Secção 4.3). Na Secção 4.4 são apresentados os resultados obtidos, recorrendo só ao uso de *features* temporais, recorrendo só ao uso de *features* no domínio da frequência, e por fim, resultados com todas as *features*. Ainda nesta secção são apresentados resultados que comparam o *FatoXtract* e o Kinect v2. Por último, é feita uma breve discussão acerca dos resultados obtidos (Secção 4.5).

4.1. Descrição do Setup Experimental

Neste trabalho usou-se o *FatoXtract* e o *Kinect v2* para aquisição, em simultâneo, das diferentes atividades desempenhadas por diversos participantes. A Figura 17 (a) mostra um dos participantes equipado com o *FatoXtract* a realizar a terceira atividade (andar de bicicleta) e com o *Kinect v2* a capturar o movimento em simultâneo. A Figura 17 (b) mostra a representação correspondente do movimento com um avatar 3D obtido do *FatoXtract Visualizer* (aplicação *online* do equipamento). A Figura 17 (c) representa a numeração de todos os 17 IMUs do *FatoXtract*.

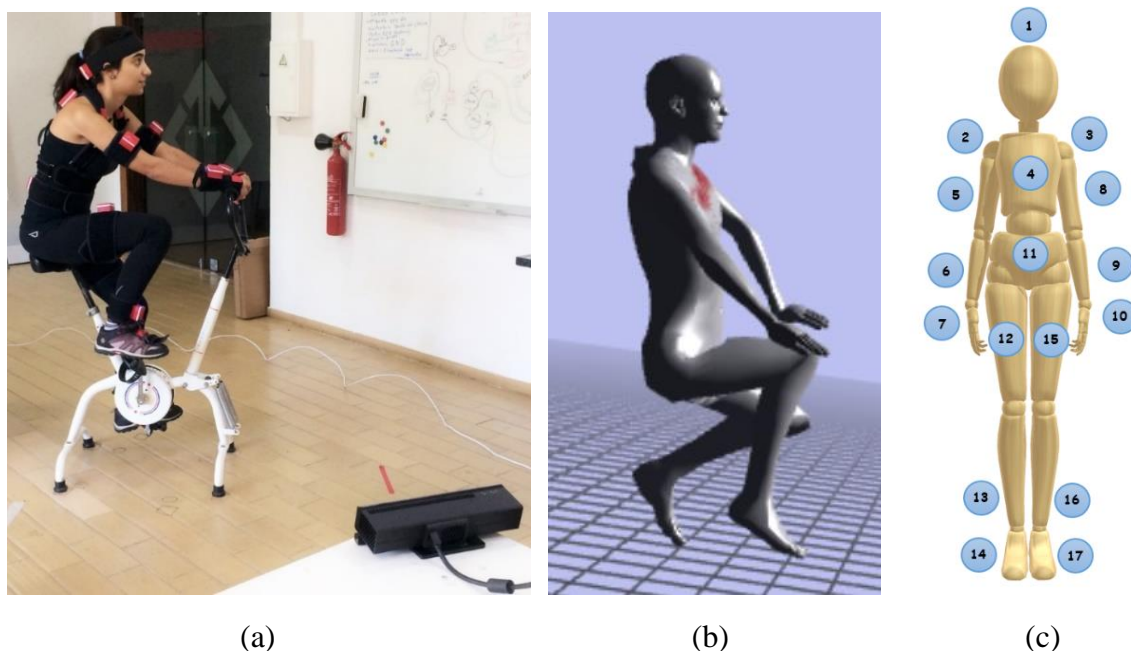


Figura 17 - Exemplo de um participante a executar a terceira atividade (andar de bicicleta). (a) Participante equipado com *FatoXtract*; (b) Avatar 3D obtido do *FatoXtract Visualizer*; (c) Numeração dos IMUs.

Através da aplicação *FatoXtract Visualizer* alojado na *Cloud Ingeniarius*, torna-se possível extrair os dados adquiridos pelo *FatoXtract* em formato *Excel*. Porém, foi desenvolvido um

script em *MatLab* de forma a extrair os dados diretamente da aplicação por intermédio de uma estrutura JSON, utilizando a função *webread* para esse efeito. O *dataset* do *FatoXtract* foi organizado por *frames*, cada uma contendo 119 dados (quatro valores de quaterniões e três valores de Euler por cada um dos 17 IMUs).

Os dados obtidos pelo *Kinect v2* foram processados através dos scripts disponíveis *online*²⁹. A câmara capturou a atividade realizada pelo participante, beneficiando de uma rotina desenvolvida em *Microsoft Visual Studio* de forma a identificar as diferentes juntas em cada *frame* e gravá-las num ficheiro de texto. De seguida, o ficheiro de texto é interpretado pela rotina desenvolvida em *MatLab*, sendo possível gravar os dados de forma desejada para posterior pós-processamento. O *dataset* do *Kinect v2* foi organizado por *frames*, cada uma contendo 75 dados (três valores (*x*, *y* e *z*) por cada uma das 25 juntas).

4.2. Descrição do Procedimento

Cada participante realizou seis atividades diárias, repetindo cada uma vinte vezes. As atividades realizadas por cada participante foram atividades do dia-a-dia, nomeadamente: lavar os dentes, acenar, andar de bicicleta, caminhar e step (Tabela 9).

Tabela 9 - Atividades

A	1	2	3	4	5
Atividades	Lavar os dentes	Acenar	Andar de Bicicleta	Caminhar	Step

Os participantes realizaram as duas primeiras atividades (lavar os dentes e acenar) sentados, de forma a demonstrar o movimento da parte superior do corpo humano, tendo essas atividades sido realizadas com o braço direito. As últimas três atividades (andar de bicicleta, caminhar e o step) dão ênfase ao movimento dos membros inferiores de cada participante.

4.3. Descrição da Amostra

A atividade experimental foi realizada por uma amostra de seis participantes: três do sexo feminino e três do sexo masculino ($19,67 \pm 1,37$ anos). É um estudo piloto com uma amostra de dimensão reduzida que visa apenas demonstrar o desempenho do equipamento *FatoXtract* e da metodologia aqui apresentada para classificação do movimento humano. Todos os participantes apresentam uma atividade física normal e são saudáveis.

²⁹ https://github.com/ininex/KinectV2_to_MATLAB

4.4. Resultados

Nesta secção são apresentados os resultados obtidos na classificação das diferentes atividades. Inicialmente, usaram-se apenas *features* no domínio temporal para classificar as diferentes atividades (Secção 4.4.1). De seguida, classificaram-se as diferentes atividades apenas recorrendo a *features* no domínio da frequência (Secção 4.4.2.). A Secção 4.4.3 apresenta os resultados com ambos os tipos de *features*. Por fim, realiza-se um *benchmarking* do *FatoXtract* comparando-o com os resultados obtidos usando o *Kinect v2*.

Os resultados são apresentados ao nível da precisão e de sensibilidade. A precisão está relacionada com reprodutibilidade e repetibilidade. É o grau em que medidas repetidas sob condições inalteradas mostram os mesmos resultados, ou seja, a divisão dos verdadeiros positivos (VP), que corresponde às atividades que foram classificadas corretamente, pela soma dos verdadeiros positivos com os falsos positivos (FP), que corresponde às atividades que foram classificadas como positivas de forma incorreta:

$$\text{Precisão} = \frac{VP}{VP + FP}. \quad (31)$$

A sensibilidade é o grau de proximidade das medições de uma quantidade para o verdadeiro valor, ou seja, é a divisão dos VP pela soma dos VP com os falsos negativos (FN), que corresponde às atividades que foram classificadas incorretamente apesar de serem positivas:

$$\text{Sensibilidade} = \frac{VP}{VP + FN}. \quad (32)$$

É de notar que, em todas as experiências realizadas, se utilizou sempre a técnica do deixar um de fora (*leave-one-out*). Por exemplo, na classificação associada ao primeiro participante, apenas os dados associados aos restantes participantes foram usados para treino, e a mesma metodologia foi seguida para os restantes casos.

4.4.1. Features do Domínio Temporal

Recorrendo ao uso de *features* no domínio temporal para classificar as diferentes atividades obtiveram-se os resultados apresentados na Tabela 10. Esta apresenta os resultados, em percentagem, obtidos ao nível da precisão (Prec) e da sensibilidade (Sens). O valor de “*Total Atividades*” representa a percentagem obtida pela média entre os valores obtidos por todas os participantes em cada atividade. O valor “*Total Participantes*” representa a percentagem obtida pela média calculada entre os valores obtidos em todas as atividades por cada participante.

Tabela 10 - Precisão e Sensibilidade dos diferentes participantes e atividades utilizando *features* do domínio temporal

		P - Participantes			A – Atividades			Total Atividades
		P1	P2	P3	P4	P5	P6	
A1	Pre	39.72	53.83	59.09	55.39	73.80	99.21	63.51
	Sens	39.36	99.04	96.88	99.30	3.86	92.16	71.77
A2	Pre	33.69	90.61	87.90	99.08	50.88	92.75	75.82
	Sens	28.96	15.08	35.48	20.78	98.47	98.94	49.62
A3	Pre	85.14	99.90	99.75	98.56	99.45	98.46	96.88
	Sens	99.03	99.44	97.24	62.40	99.71	99.35	92.86
A4	Pre	98.74	99.90	99.67	60.33	99.42	98.24	92.72
	Sens	93.63	98.86	97.44	99.65	99.22	98.40	97.87
A5	Pre	95.33	98.37	98.54	98.78	98.12	99.28	98.07
	Sens	98.95	99.22	98.99	70.40	99.26	98.65	94.25
Total Participantes	Pre	70.52	88.52	88.99	82.43	84.33	97.59	85.40
	Sens	71.99	82.33	85.21	70.51	80.10	97.50	81.27

Ao analisar a Tabela 10 pode-se verificar que P1 é o que apresenta piores resultados relativamente aos restantes. Este apresenta uma precisão total de 70.52 % e uma sensibilidade total de 71.99%. Por outro lado, P6 é o que apresenta melhores resultados, com uma precisão total de 97.59 % e uma sensibilidade total de 97.50 %. Relativamente às atividades, verifica-se que as últimas três apresentam melhores valores de precisão e de sensibilidade, pois são atividades que beneficiam de mais juntas para classificar o movimento, enquanto que as primeiras duas atividades apenas usam as juntas do braço direito.

Ao recorrer ao uso de *features* no domínio temporal conseguiu-se obter uma taxa global de precisão de 85.4 % e uma taxa global de sensibilidade de 81.27 %. Para tentar melhorar o resultado da classificação, recorreu-se ao uso de *features* no domínio da frequência.

4.4.2. Features no Domínio da Frequência

Na Tabela 11 podemos visualizar os resultados obtidos para a classificação das diferentes atividades apenas recorrendo ao uso de *features* no domínio da frequência. Nesta tabela serão apresentados também os resultados, em percentagem, da precisão e da sensibilidade de cada atividade realizada pelos diferentes participantes, assim como as percentagens globais.

Tabela 11 - Precisão e Sensibilidade dos diferentes participantes e atividades utilizando *features* do domínio da frequência

		P - Participantes			A – Atividades			Total Atividades
		P1	P2	P3	P4	P5	P6	
A1	Pre	22.26	60.63	53.47	0.81	48.88	92.85	46.48
	Rec	10.09	72.10	26.77	0.06	84.55	52.53	41.02
A2	Pre	31.99	47.85	29.92	43.70	43.73	73.23	45.07
	Rec	68.75	39.69	23.48	91.34	23.44	80.42	54.52
A3	Pre	67.09	74.04	46.40	73.48	72.49	71.37	67.48
	Rec	21.06	84.61	86.20	39.29	64.33	86.08	63.60
A4	Pre	66.37	86.20	79.60	85.53	86.96	88.75	82.24
	Rec	55.27	86.14	80.24	48.07	51.11	93.09	68.99
A5	Pre	58.76	88.08	81.19	51.77	62.18	77.00	69.83
	Rec	73.51	73.93	68.89	89.82	78.27	83.23	77.94
Total Participantes	Pre	49.29	71.36	58.12	51.06	62.85	80.64	62.22
	Rec	45.74	70.69	57.12	53.72	60.34	79.07	61.21

Ao analisar a Tabela 11 pode-se concluir que P1 foi, uma vez mais, aquele que obteve piores percentagens de classificação, apresentando valores de precisão de 49.29 % e de sensibilidade de 45.74 %. Por outro lado, P6 apresenta os melhores valores relativamente aos restantes participantes, com valores de precisão de 80.64 % e de sensibilidade de 79.07 %. Ao recorrer apenas ao uso de *features* no domínio da frequência para classificar as diferentes atividades, obteve-se uma taxa global de precisão de 62.22 % e uma taxa global de sensibilidade de 61.21 %.

Estes valores foram inferiores aos valores obtidos na classificação recorrendo apenas ao uso de *features* no domínio temporal. Posto isto, optou-se por juntar os dois tipos de *features* de forma a tentar melhorar os resultados.

4.4.3. Todas as *Features*

Esta secção apresenta os resultados quando se decide juntar os dois tipos de *features*: as *features* no domínio temporal e as *features* no domínio da frequência. Na Tabela 12 pode-se visualizar os resultados em percentagem dos valores de precisão e de sensibilidade obtidos aquando do uso destes dois tipos de *features* em simultâneo.

Tabela 12 - Precisão e Sensibilidade dos diferentes participantes e atividades usando os dois tipos de *features*

		P - Participantes			A - Atividades			Total Atividades
		P1	P2	P3	P4	P5	P6	
A1	Pre	69.25	50.62	65.60	72.32	91.71	99.13	74.77
	Rec	11.93	99.14	96.05	91.90	44.61	85.38	71.50
A2	Pre	46.46	50.73	91.46	79.93	64.67	86.63	69.98
	Rec	63.43	3.24	52.75	65.11	98.47	99.87	63.81
A3	Pre	64.20	95.61	99.55	99.78	99.60	100.00	93.12
	Rec	99.18	99.02	97.17	82.65	89.81	99.00	94.47
A4	Pre	99.47	99.71	99.70	91.02	97.77	98.28	97.66
	Rec	89.12	94.52	97.93	97.49	98.31	99.38	96.12
A5	Pre	97.19	99.60	98.90	98.46	91.60	100.00	97.62
	Rec	99.30	99.03	98.99	99.95	99.26	98.48	99.17
Total Participantes	Pre	75.32	79.25	91.04	88.30	89.07	96.81	86.63
	Rec	72.59	78.99	88.58	87.42	86.09	96.42	85.01

Verificou-se uma melhoria no desempenho da classificação das atividades, tanto a nível da precisão como da sensibilidade. P6 continua a ser o participante com valores mais elevados, com 96.81 % de precisão e 96.42 % de sensibilidade, e P1 mantém-se o participante com valores mais baixos de classificação, apresentando 75.32 % de precisão e 72.59 % de sensibilidade. Como a Tabela 11 demonstra, A5 (step) é a atividade que apresenta uma melhor classificação, com uma precisão de 97.62 % e com uma sensibilidade de 99.17 %.

A classificação apresenta uma taxa global de precisão 86.63 % e uma taxa global de sensibilidade de 85.01 %. Ao juntar as *features* do domínio da frequência às *features* do domínio temporal, fez com que o valor de precisão total tivesse um aumento de 1,23 % e a sensibilidade total um aumento de 3,74 %, relativamente aos valores obtidos quando a classificação foi realizada apenas recorrendo ao uso de *features* do domínio temporal. Pode-se então concluir que, ao usar os dois tipos de *features* em simultâneo, obtém-se melhores resultados de classificação.

4.4.4. Discussão sobre a escolha das *Features*

Na Figura 18 podemos visualizar a matriz confusão dos resultados da classificação obtidos através do *FatoXtract*. A matriz confusão quantifica a percentagem de dados que foram classificados de forma correta e incorreta pelo modelo construído. Através desta figura podemos concluir que A1 e A2 são atividades semelhantes, apresentando valores de sensibilidade inferiores às restantes atividades. A1 contém uma percentagem de VP de 71.50 % e 25.31% de FP, confundindo-se com A2 em várias ocasiões. A2 contém uma percentagem de VP de 63.81 % e FP de 30.89 %, sendo que esta se confunde com A1 em cerca de 1/3 das situações. As restantes atividades apresentam valores de VP superiores a 94 %, ou seja, pode-

se concluir que as últimas três atividades foram classificadas com um desempenho elevado utilizando as *features* propostas e recorrendo ao classificador DBMM apresentado.

A1- Lavar os Dentes	71.50	25.31	2.76	0.09	0.35
A2 – Acenar	30.89	63.81	5.25	0.05	
A3- Andar de Bicicleta	0.80	1.57	94.47	1.95	1.20
A4 – Caminhar	0.19	0.75	1.98	96.12	0.96
A5 – Step	0.17	0.16	0.15	0.35	99.17
	A1 - Lavar os Dentes	A2 – Acenar	A3- Andar de Bicicleta	A4 – Caminhar	A5 – Step

Figura 18 – Matriz Confusão global usando o classificador DBMM e as as *features* espaço-temporais propostas (frequência + temporal + geométrica)

Tal como resumido na Figura 18, pode-se concluir que ao combinar os dois tipos de *features* é possível obter um melhor desempenho na classificação.

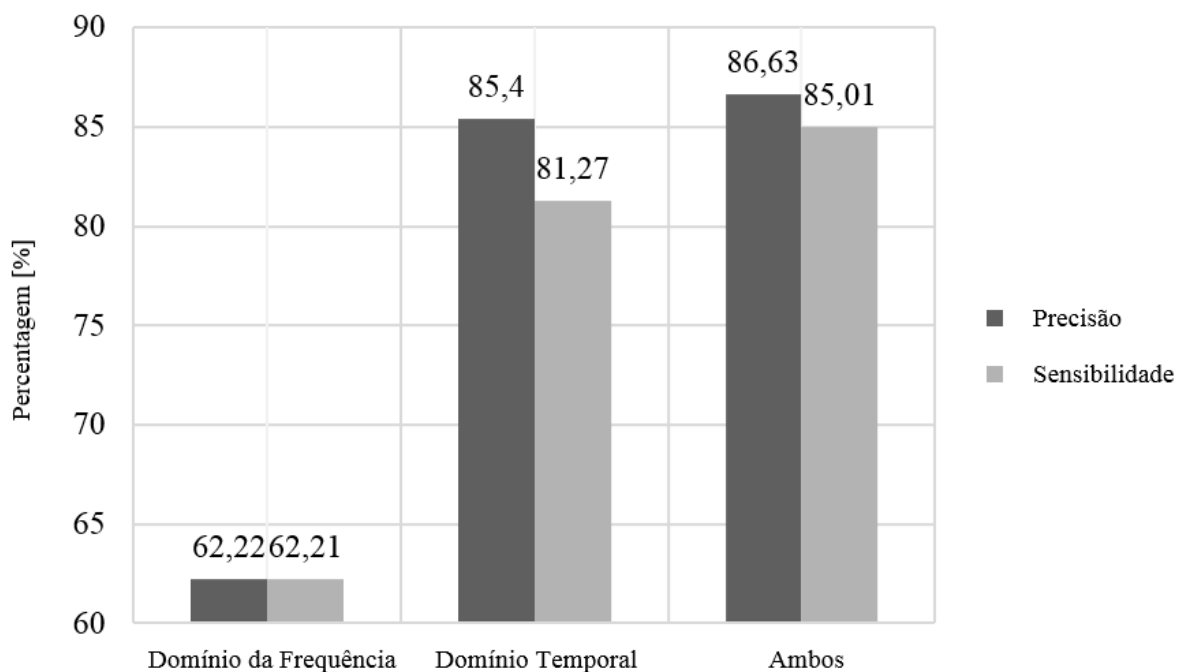


Figura 19 - Comparação entre os dois tipos de features

Como se pode observar na Figura 19, o conjunto proposto de *features* foi capaz de lidar com situações difíceis, especialmente quando se consideram todas as *features*, distinguindo, assim, atividades similares na maioria das situações, tais como lavar os dentes e acenar. As restantes atividades foram mais fáceis de distinguir, apresentando uma taxa de erro de classificação inferior. Isto pode ser explicado devido ao facto de estes requererem a maior parte das 17 juntas do corpo, o que torna mais fácil identificá-los quando comparados com as atividades diárias representados por um menor número de juntas (*e.g.*, lavar os dentes ou acenar).

4.4.5. Benchmarking

No caso do *Kinect v2* foram utilizadas *features* que dependem das relações existentes entre as diferentes partes do corpo para capturar movimentos com características significativas de uma pessoa que exerça uma atividade. Estas *features* foram extraídas a partir do esqueleto 3D, em particular, as coordenadas 3D de cada articulação do esqueleto. Dadas as coordenadas 3D do esqueleto em cada *frame*, um conjunto de diferentes *features* foi definido em três diferentes subconjuntos por $\varpi = \{\omega_{s1}, \omega_{s2}, \omega_{s3}\}$, somando um total de 108 features.

- O primeiro subconjunto de *features* é baseado na distância euclidiana entre as juntas do esqueleto e os ângulos formados entre as juntas que são calculados como *features* geométricas;
- O segundo subconjunto de *features* é baseado em valores derivados e acumulativos empregues no primeiro subconjunto;

- O terceiro subconjunto é baseado nas diferentes *features* espaço-temporais, tais como, modelos de energia e velocidades, velocidade angulares e auto correlação das poses do esqueleto, modelo de log-energia das *features* geométricas, log-distância entre as poses do esqueleto, log-covariância sobre as diferenças de distâncias entre as juntas do esqueleto em diferentes *frames* e outras *features* estatísticas foram calculadas, estimativa da probabilidade máxima.

Mais detalhes sobre as *features* usadas com tecnologia *Kinect* podem ser encontrados em Faria, Premebida e Nunes (under review).

A Figura 20 ilustra a matriz confusão originada pelos dados do *Kinect v2*. Nesta figura podemos visualizar que A1 e A2 são as atividades que apresentam percentagens mais baixas de classificação. A1 apresenta 77.73 % de VP, e esta confunde-se com A2 em 21.69 % das situações. A2 apresenta 71.31 % de VP, e esta confunde-se com A1 em 14.92 %. As restantes atividades apresentam valores acima dos 87 %, não se confundido significativamente com as restantes atividades.

1- Lavar os Dentes	77.73	21.69	0.27	0.05	0.26
2 – Acenar	14.92	71.31	10.18	0.26	3.33
3- Andar de Bicicleta	0.38	0.11	98.60	0.14	0.78
4 – Caminhar	0.20	0.32	1.41	87.20	10.87
5 – Step	0.90	0.75	6.92	3.10	88.34
	1 - Lavar os Dentes	2 – Acenar	3- Andar de Bicicleta	4 – Caminhar	5 – Step

Figura 20 - Matriz Confusão dos dados do *Kinect v2*

De seguida apresenta-se uma breve comparação entre os dados obtidos através do *FatoXtract* e o *Kinect v2*. Na Tabela 13 podemos visualizar os valores de precisão e de sensibilidade obtidos usando o *FatoXtract* e o *Kinect v2* nas diferentes atividades.

Tabela 13 - Comparação dos valores Precisão e Sensibilidade do *FatoXtract* e do Kinect v2

	A1		A2		A3		A4		A5		Total	
	Prec	Sens	Prec	Sens	Prec	Sens	Prec	Sens	Prec	Sens	Prec	Sens
FatoXtract	74.77	71.50	69.98	63.81	93.12	94.47	97.66	96.12	97.62	99.17	86.63	85.01
Kinect v2	82.57	77.73	75.73	71.31	84.00	98.60	96.10	87.20	85.29	88.34	84.74	84.64

Ao observar a Tabela 13 pode-se concluir que, para a classificação do movimento humano, o uso do *FatoXtract* apresenta melhores resultados, de forma geral, quando comparado com o uso do *Kinect v2*. O *FatoXtract* apresenta um valor total de precisão de 86.63%, enquanto o *Kinect v2* apresenta um valor de 84.74 %. O mesmo acontece no caso do valor total da sensibilidade que, para o *FatoXtract*, apresenta um valor de 85.01%, enquanto que para o *Kinect v2* este apresenta um valor inferior de 84.64%.

Todas as atividades são melhor classificadas aquando o uso do *FatoXtract*, à exceção das duas primeiras atividades (A1 e A2), que apresentam melhores resultados com o uso do *Kinect v2*. Tal pode ser explicado por se tratarem de duas atividades que não têm juntas que sofram de oclusão, pois as atividades são apenas realizado com o braço direito, estando as poucas juntas necessárias para classificar o movimento no campo de visão do *Kinect v2*.

4.5. Discussão

Neste trabalho apresentamos uma arquitetura que recorre ao uso de múltiplos classificadores para a classificação do movimento humano. Os dados foram adquiridos a partir de uma tecnologia que recorre ao uso de IMUs – *FatoXtract* – tendo sido feita uma análise comparativa com o *Kinect v2*. Em geral, todas as atividades foram bem classificadas.

A classificação depende do número de juntas que são necessárias para classificar um movimento. Por exemplo, se estamos a lavar os dentes com o braço direito, apenas serão relevantes as juntas do braço direito para classificar esse movimento. No entanto, se estivermos a andar de bicicleta, todas as juntas do corpo são necessárias para identificar esse movimento, sobretudo os membros inferiores. O facto de haver atividades muito similares torna o processo de classificação mais exigente e difícil. Devido a isso, teve de ser feita uma escolha de diferentes *features* dependendo da tecnologia utilizada que, juntamente com o grande conjunto de *frames*, fez com que todo o processamento fosse moroso.

Em primeiro lugar foram usadas *features* no domínio temporal para classificar as diferentes atividades. Depois, foram usadas *features* no domínio da frequência, e, por último, juntaram-se estes dois tipos de *features*, tendo-se observado uma melhoria significativa no desempenho da classificação das atividades realizadas pelos participantes. A primeira atividade (lavar os dentes) e a segunda (acenar) eram muito semelhantes, necessitando as duas atividades apenas das juntas do braço direito para classificar o movimento. Isto fez com que a classificação das

mesmas apresentasse piores resultados relativamente as outras, dado serem atividades mais difíceis de distinguir, o que se verificou pela análise percentual dos FP e FN. Por esta razão, os valores de precisão e de sensibilidade foram piores do que nas restantes atividades. A primeira atividade teve uma precisão global de 74,77% e uma sensibilidade global de 71,50%. A segunda atividade obteve a precisão global de 69,98% e uma sensibilidade global de 63,81%. As restantes atividades demonstraram uma precisão global e uma sensibilidade global superior a 90 %, pois estas atividades recorrem a todas as juntas do corpo para serem devidamente classificadas.

O uso do *FatoXtract* foi comparado com o uso de outra tecnologia que recorre ao uso de câmaras; o *Kinect v2*. O *Kinect v2* só apresentou melhores resultados de classificação nas duas primeiras atividades. Tal pode ser explicado por se tratarem de duas atividades que não têm um número tão elevado de juntas relevantes e, por outro lado, as poucas relevantes não sofrem de oclusões. Pode-se assim concluir que a tecnologia *wearable*, usada pela solução *FatoXtract*, apresenta um desempenho superior para a devida identificação do movimento humano relativamente à tecnologia que recorre ao uso de câmaras, nomeadamente utilizando o *Kinect v2*. O *FatoXtract* apresentou uma precisão total de 86.63 % enquanto que o *Kinect v2* ficou pelos 84.74%. Relativamente à sensibilidade, o mesmo voltou a acontecer, sendo que o *FatoXtract* apresentou 85.01 %, sendo superior à sensibilidade apresentada pelo *Kinect v2* que ficou nos 84.64 %.

Assim, considerando os resultados obtidos, verifica-se que o uso da tecnologia *wearable* apresenta grandes vantagens na análise do movimento humano quando comparado com outras tecnologias, principalmente por não necessitar de qualquer câmara, marcadores ou infraestrutura. Estes aspetos são importantes para uma melhor compreensão do reconhecimento de padrões do movimento humano, e como estes emergem no contexto de laboratório e em situações diárias (Chen & Hauptmann, 2009). Isto tem aplicações práticas que abrange não só a área de Engenharia, mas também a Biomecânica e as Ciências do Desporto. Aggarwal e Cai (1999) afirmam que, por todas estas razões, considera-se que a classificação de movimentos padrão, como as atividades do dia-a-dia, são muito mais complexas do que alguma vez se imaginou. Neste caso, os resultados mostram como diferentes pessoas podem ser diferentes no desempenho do mesmo movimento e como esta diferença se acaba por verificar no desempenho do classificador. Este tipo de variabilidade intrínseca do sistema de movimento humano, que é uma espécie de “assinatura” única para cada pessoa, é algo que a tecnologia *wearable* pode identificar em futuras investigações associadas a movimentos mais complexos. A possibilidade em combinar métodos de classificação distintos, como o ANN, SVM e NB, permite convergir no sentido de apresentar uma arquitetura geral de classificação que se proponha a resolver problemas relacionados com o desporto, saúde, biomecânica e engenharia. Por exemplo, na área da biomecânica, estes três classificadores, quando combinados com a tecnologia *wearable* e a respetiva metodologia aqui apresentada, vai eventualmente desvendar alguns dos “segredos” por trás da execução de diversos movimentos, tais como, por exemplo, tocar bateria; algo que quase não tem sido estudado na literatura.

CONCLUSÃO E TRABALHO FUTURO

5.1. Conclusão

Este trabalho de mestrado incidiu sobre a análise do movimento humano, mais especificamente na classificação de atividades do dia-a-dia. Para tal, esta classificação apoiou-se na escolha de um conjunto de *features* e numa metodologia de classificação adequados, apoiados na utilização de tecnologia *wearable*.

Realizou-se, em primeiro lugar, uma revisão da literatura sobre as tecnologias e produtos existentes no mercado, chegando-se assim à conclusão que a tecnologia *wearable*, nomeadamente a solução *FatoXtract*, seria a mais adequada por demonstrar várias vantagens relativamente às outras tecnologias. Como tecnologia de comparação, usou-se o *Kinect v2*. De seguida, foi feito um estudo aprofundado de diversos conjuntos de *features* e de métodos de classificação utilizados por diferentes autores. Com base na revisão da literatura, pode-se concluir que a seleção de *features* tem uma enorme importância e é através desta que se pode otimizar o classificador com o intuito em melhor identificar as atividades humanas.

O presente estudo baseou-se na escolha do conjunto de *features* adequadas, perante o facto do conjunto de dados ter sido projetado para ser extremamente desafiador uma vez que existem grandes variações intra-classe, a duração das atividades varia drasticamente, e algumas atividades são muito semelhantes. Usaram-se *features* no domínio temporal (como, por exemplo, a velocidade) e *features* no domínio da frequência (como, por exemplo, a FFT). De seguida, foi utilizado o modelo de classificação que combina as probabilidades de três classificadores diferentes, ou seja, Naïve Bayes, Redes Neurais Artificiais e Máquinas de Vetor de Suporte.

O conjunto de *features* e o método de classificação escolhidos demonstraram terem sido adequados para classificar as cinco atividades diárias escolhidas. Conseguiu-se obter uma taxa total de precisão de 86,63 % e uma taxa total de sensibilidade de 85,01 % através do uso da tecnologia *wearable*. Obtiveram-se valores inferiores recorrendo ao uso de câmaras *Kinect v2*, com uma taxa total de precisão de 84,74 % e uma taxa total de sensibilidade de 84,64 %. Pode-se assim concluir que o uso da tecnologia *wearable* será mais adequada para a captura de atividades diárias do que o uso de tecnologias que recorrem ao uso de câmaras.

Apesar dos diferentes métodos de classificação poderem apresentar diferentes níveis de exatidão e precisão no reconhecimento do movimento humano, a escolha das *features* mais adequadas podem impulsionar significativamente a precisão global da classificação para um nível superior.

5.2. Trabalho Futuro

Como trabalho futuro pode-se aumentar o leque de atividades diárias e aplicar a mesma extração de *features* e o mesmo método de classificação utilizado neste trabalho. Para além de atividades diárias, pode-se aplicar esta abordagem a outro tipo de atividades, como, por exemplo, atividades na área da saúde (Aggarwal & Cai, 1999) (nomeadamente na reabilitação de pessoas com deficiências motoras), atividades para análise e controlo de diferentes pessoas no seu local de trabalho, ou mesmo atividades de reconhecimentos de ações através de vigilância (Santos L. C., 2013).

O uso de outros classificadores, como o Modelo Escondido de Markov e a Lógica Difusa podem ser considerados para fazer a fusão com os métodos já existentes no classificador DBMM, de forma a fazer o *benchmarking* entre as diversas arquiteturas de fusão de classificação.

Também será possível identificar a componente latente do *dataset*, de modo a utilizar só a informação relevante do movimento a estudar. Ou seja, aplicar uma janela temporal dinâmica de forma a obter o conjunto de *frames* mais relevantes para a classificação do movimento.

Futuras investigações poderiam determinar o melhor conjunto de *features* para classificar as diferentes atividades. Para este efeito, poderá recorrer-se a um método de otimização, tal como o *fractional order darwinian particle swarm optimization* (Couceiro, Rocha, Ferreira, & Machado, 2012), com o intuito em maximizar medidas de desempenho associadas à classificação ao longo do tempo (*e.g.*, maximizar sensibilidade e precisão).

Por último, poder-se-á fazer a classificação em tempo real com duas finalidades. Uma poderia ser como método de prevenção/predição do movimento, ou seja, prever o movimento da pessoa de modo a precaver situações menos desejadas. Por exemplo, suponhamos que a trajetória efetuada por um idoso pode resultar na queda do mesmo, o método poderia conseguir prever essa situação e evitá-la conjugando isso com um sistema de alarme. Por outro lado a classificação em tempo real poderia ter como finalidade a classificação de vários movimentos num determinado intervalo de tempo.

REFERÊNCIAS

- Aggarwal, J. K., & Cai, Q. (1999). Human motion analysis: A review. *Computer vision and image understanding*, 73(3), 428-440.
- Amadio, A. C., & Serrão, J. C. (2007). Contextualização da biomecânica para a investigação do movimento: fundamentos, métodos e aplicações para análise da técnica esportiva. pp. 61-85.
- Araújo, A. G., Andrade, L. M., & Barros, R. M. (2005). Sistema para análise cinemática da marcha humana baseado em videogrametria. *Fisioterapia e Pesquisa*, 11(1), 3-10.
- Arsigny, V., Fillard, P., Pennec, X., & Ayache, N. (2006). Log-euclidean metrics for fast and simple calculus on diffusion tensors. *Magnetic Resonance in Medicine*, 56(2), 411-421.
- Avcı, A., Bosch, S., Marin-Perianu, M., & Havinga, P. (2010). Activity recognition using inertial sensing for healthcare, wellbeing and sports applications: A survey. *International Conference in Architecture of Computing Systems (ARCS), 2010*, (pp. 1-10).
- Ayat, N. E., Cheriet, M., & Suen, C. Y. (2002). KMOD - A Two-parameter SVM Kernel for Pattern Recognition. *International Conference on Pattern Recognition* (pp. 331-334). IEEE.
- Bao, L., & Intille, S. S. (2004). Activity Recognition from User-Annotated Acceleration Data. (pp. 1-17). Springer Berlin Heidelberg.
- Barbosa, C. (2011). *Modelação biomecânica do corpo humano: aplicação na análise da marcha*. Universidade do Minho. Dissertação Mestrado.
- Ben-Hur, A., & Weston, J. (2010). A User's Guide to Support Vector Machines. Em *Data mining techniques for the life sciences* (pp. 223-239). Humana Press.
- Brumatti, M. (2005). *Redes Neurais Artificiais*. Espírito Santo: Vitória.
- Byun, H., & Lee, S.-W. (2003). A survey on pattern recognition applications of support vector machines. *International Journal of Pattern Recognition and Artificial Intelligence*, 17(3), 459-486.
- Carvalho, A. G. (2011). *Influência da Modelagem dos Componentes de Bias Instabilidade dos Sensores Inerciais no Desempenho do Navegador Integrado SNI/GPS*.
- Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3), 27.
- Chen, M.-Y., & Hauptmann, A. (2009). *Mosift: Recognizing Human Actions in Surveillance Videos*.
-

-
- Chen, X. (2013). *Human Motion Analysis with Wearable Inertial*. University of Tennessee. Knoxville: PhD Thesis.
- Cintra, R. S., Velho, H. F., & Todling, R. (2011). Redes Neurais Artificiais na Melhoria de Desempenho de Métodos de Assimilação de Dados: Filtro de Kalman. *Trends in Applied and Computational Mathematics*, 11(1), 29-39.
- Correia, M. M. (2013). *Reconhecimento de Elementos Gestuais com Kinect*.
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273-297.
- Couceiro, M. S., Portugal, D., Gonçalves, N., Rocha, R., Luz, J. M., Figueiredo, C. M., & Dias, G. (2013). A Methodology for Detection and Estimation in the Analysis of the Golf Putting. *Pattern Analysis and Applications*, 16(3), 459-474.
- Couceiro, M. S., Rocha, R. P., Ferreira, N. M., & Machado, J. A. (2012). Introducing the fractional-order Darwinian PSO. *Signal, Image and Video Processing*, 6(3), 343-350.
- Cui, Y., Schuon, S., Chan, D., Thrun, S., & Theobalt, C. (2010). 3D Shape Scanning with a Time-of-Flight Camera. In *Computer Vision and Pattern Recognition, 2010 IEEE Conference on*, (pp. 1173-1180).
- da Cunha, A. M. (2009). *Movimentos de Cabeça guiados pela Voz*. Tese de Doutorado.
- Dubois, A., & Charpille, F. (2013). Human Activities Recognition with RGB-Depth camera using HMM. In *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE* (pp. 4666-4669). IEEE.
- Ermes, M., & Parkka, J. (2008). Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions. *IEEE Transactions on Information Technology in Biomedicine*, 12(1), 20-26.
- Faria, D. R., Premevida, C., & Nunes, U. (2014). A probabilistic approach for human everyday activities recognition using body motion from RGB-D images. *The 23rd IEEE International Symposium on Robot and Human Interactive Communication, 2014 RO-MAN*., 732-737.
- Faria, D. R., Premevida, C., & Nunes, U. (under review). Dynamic Bayesian Mixture Model: Probabilistic Classification for Human Daily Activity Recognition. *IEEE, Transactions on Pattern Analysis and Machine Intelligence*.
- Faria, D. R., Vieira, M., Premevida, C., & Nunes, U. (2015). Probabilistic Human Daily Activity Recognition towards Robot-assisted Living. *Proceedings of IEEE RO-MAN'15: IEEE International Symposium on Robot and Human Interactive Communication*.
- Ferrigno, G., & Pedotti, A. (1985). ELITE: a digital dedicated hardware system for movement analysis via real-time TV signal processing. *IEEE Transactions in Biomedical Engineering*, 943-950.
-

-
- Foix, S., Alenyà, G., & Torras, C. (2011). Lock-in Time-of-Flight (ToF) Cameras: A Survey. *IEEE Sensors Journal*, 11(9), 1917-1926.
- Fong, D. T.-P., & Chan, Y.-Y. (2010). The Use of Wearable Inertial Motion Sensors in Human Lower Limb Biomechanics Studies: A Systematic Review. *Sensors* 10(12), 11556-11565.
- Fusier, F., Valentin, V., Bremond, F., Thonnat, M., Borg, M., Thirde, D., & Ferryman, J. (2007). Video Understanding for Complex Activity Recognition. *Machine Vision and Applications*, 18(3-4), 167-188.
- Ganapathi, V., Plagemann, C., Koller, D., & Thrun, S. (2010). Real Time Motion Capture Real Using a Single Time-Of-Flight Camera. *IEEE Conference in Computer Vision and Pattern Recognition (CVPR)* (pp. 755-762). IEEE.
- Garofalo, P. (2010). *Development of motion analysis protocols based on inertial sensors*. University of Bologna. PhD thesis.
- Gomes, A. F. (2012). *Metodologias e Sistemas Activos Baseados em Fusão Sensorial para Análise de Movimento- Aplicações em Biomecânica*. Universidade do Porto.
- Graupe, D. (2007). *Principles of Artificial Neural Networks* (Vol. 6). World Scientific.
- Guo, K. (2012). *Action recognition using log-covariance matrices of silhouette and optical-flow features*. Boston University, College of Engineering. PhD. Thesis.
- Heckerman, D. (1998). A Tutorial on Learning With Bayesian Networks. Em *Springer Netherlands* (pp. 301-354).
- Hong, Y.-J., Kim, I.-J., Ahn, S. C., & Kim, H.-G. (2008). Activity recognition using wearable sensors for elder care. *International Conference on In Future Generation Communication and Networking, 2008 FGCN'08*. 2, pp. 302-305. IEEE.
- Hussain, M., Wajid, S. K., Elzaart, A., & Berbar, M. (2011). A comparison of SVM kernel functions of breast cancer detection. *International Conference in Computer Graphs, Imaging and Visualization* (pp. 145-150). IEEE.
- Huynh, T. (2008). *Human Activity Recognition with Wearable Sensors*. Technische Universität Darmstadt. PhD Thesis.
- Huynh, T., & Schiele, B. (2005). Analyzing Features for Activity Recognition. In *Proceedings of the 2005 Joint Conference on Smart objects and ambient intelligence: innovative context-aware services: usages and technologies*, (pp. 159-163).
- Ivansson, J. (2002). *Situation assessment in a stochastic environment using Bayesian Networks*. Linköping University. Master Thesis.
- Jackson, J. E. (2005). *A user's guide to principal components* (Vol. 587). John Wiley & Sons.
- Jain, A. K., Mao, J., & Mohiuddin, K. M. (1996). Artificial neural networks: A tutorial. *Computer*, 29(3), 31- 44.
-

-
- Jolliffe, I. T. (2002). *Principal Component Analysis*. John Wiley & Sons, Ltd.
- Josefsson, T., Nordh, E., & Eriksson, P.-O. (1996). A flexible high-precision video system for digital recording of motor acts through lightweight reflex markers. *Computer methods and programs in biomedicine*, 49(2), 119-129.
- Júnior, R. C. (2013). *Estudo e Desenvolvimento de um Sistema Baseado em Jogos para Gerenciamento e Acompanhamento Remoto de Terpaia para Reabilitação Motora*. Universidade Federal da Paraíba.
- Khoshhal, K., Aliakbarpour, H., Quintas, J., Drews, P., & Dias, J. (2010). Probabilistic LMA-based classification of human behaviour understanding using power spectrum technique. *Conference in Information Fusion (FUSION)* (pp. 1-7). IEEE.
- Kok, M., Hol, J. D., & Schön, T. B. (2014). An optimization-based approach to human body motion capture using inertial sensors. *In 19th World Congress of the International Federation of Automatic Control (IFAC)* (pp. 79-85). South Africa: International Federation of Automatic Control.
- Krzanowski, W. (2000). *Principles of multivariate analysis*. Oxford University Press.
- Kumar, E. P., & Sharma, E. P. (2014). Artificial Neural Networks - A Study. *International Journal of Emerging Engineering Research and Technology*, 2(2), 143-148.
- Lange, R., & Seitz, P. (2001). Solid-State Time-of-Flight Range Camera. *IEEE Journal of Quantum Electronics*, 37(3), 390-397.
- Lara, Ó. D., & Labrador, M. A. (2013). A survey on human activity recognition using wearable sensors. *Communications Surveys & Tutorials, IEEE*, 15(3), 1192-1209.
- Li, E. Y. (1994). Artificial neural networks and their business applications. *Information & Management*, 27(5), 303-313.
- Lin, J. F.-S. (2012). *Automated Rehabilitation Exercise Motion Tracking*. University of Waterloo. MSc Thesis.
- Litomisky, K. (2012). *Consumer RGB-D Cameras and their Applications*. Rapport technique, University of California.
- Macellari, V. (1983). CoSTEL: a peripheral remote sensing device for 3-dimensional monitoring of human motion. *Medical and Biological Engineering and Computing*, 21(3), 311-318.
- Magarreiro, J. (2015). *Protocolo de recolha e processamento de dados 3D cinemáticos e cinéticos da marcha em regime laboratorial e ambulatório*. Instituto Politécnico de Setúbal. Master Thesis.
- Marques, R. L., & Dutra, I. (2002). Redes Bayesianas: o que são, para que servem, algoritmos e exemplos de aplicações. *Coppe Sistemas – Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil*.
-

-
- Mieзал, M., Bleser, G., Schmitz, N., & Stricker, D. (2013). A generic approach to inertial tracking of arbitrary Kinematic chains. *In Proceedings of the 8th International Conference on Body Area Networks*, (pp. 189-192).
- Osuna, E., Freund, R., & Girosi, F. (1997). Training support vector machines: an application to face detection. *Computer Vision and Pattern Recognition* (pp. 130-136). IEEE, Computer Society Conference on.
- Pinheiro, A. P. (2008). *Desenvolvimento de um sistema de captura e análise de movimentos baseados em técnicas de visão computacional*. Universidade Federal de Uberlândia. MSc Thesis.
- Priddy, K. L. (2005). *Artificial neural networks: an introduction* (Vol. 68). SPIE Press.
- Roetenberg, D., Luinge, H., & Slycke, P. (2009). *Xsens MVN: full 6DOF human motion tracking using miniature inertial sensors*. Xsens Motion Technologies BV.
- Santos, L. C. (2013). *Laban Movement Analysis: A Bayesian Computational Approach to Hierarchical Motion Analysis and Learning*. University of Coimbra. PhD Thesis.
- Schepers, M. (2009). *Ambulatory assessment of human body kinematics and kinetics*. University of Twente. PhD Thesis.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6(2), 461- 464.
- Seber, G. A. (2009). *Multivariate observations* (Vol. 252). John Wiley & Sons.
- Silva, A. F. (2011). *Projecto de uma clínica de análise biomecânica do movimento*. Faculdade de Engenharia da Universidade do Porto.
- Silva, R. (2006). *Um modelo de redes bayesianas aplicado a sistemas de realidade aumentada*. PhD Thesis, Universidade Federal do Rio de Janeiro.
- Suaide, A. (2010). *Desenvolvimento e Validação de uma Ferramenta Computacional para Mensuração das Curvaturas da Coluna Vertebral*. Universidade de São Paulo. PhD Thesis.
- Suutala, J., Pirttikangas, S., & Roning, J. (2007). Discriminative temporal smoothing for activity recognition from wearable sensors. (pp. 182-195). Springer Berlin Heidelberg.
- Takeda, R., Tadano, S., Natorigawa, A., Todoh, M., & Yoshinari, S. (2009). Gait posture estimation using wearable acceleration and gyrosensors. *Journal of Biomechanics*, 42(15), 2486-2494.
- Tözeren, A. (2000). *Human body dynamics: classical mechanics and human movement*. Springer Science & Business Media.
- Uddin, M. Z., Thang, N. D., Kim, J. T., & Kim, T.-S. (2011). Human activity recognition using body joint-angle features and hidden Markov model. *Etri Journal*, 33(4), 569-579.
- Uebersax, D. (2010). *Gesture Recognition with a Time-of-Flight Camera*. Swiss Federal Institute of Technology in Zurich. MSc Thesis.
-

-
- Vapnik, V. (1998). *Statistical Learning Theory* (Vol. 1). New York: Wiley.
- Vital, J. P., Couceiro, M. S., Dias, G., & Ferreira, N. M. (2015). Tecnologias para a análise do movimento humano. Em R. Ruben, M. Vieira, C. Campos, H. Almeida, J. Siopa, P. Bártolo, & J. Folgado (Edits.), *6º Congresso Nacional de Biomecânica* (pp. 1-6). Instituto Politécnico de Leiria.
- Yadav, J. S., Yadav, M., & Jain, A. (2014). Artificial Neural Network. *International Journal of Scientific Research And Education*, 1(6), 108-118.
- Yang, C.-C., & Hsu, Y.-L. (2010). A Review of Accelerometry-Based Wearable Motion Detectors for Physical Activity Monitoring. *Sensors*, 10(8), 7772-7788.
- Zhang, M., & Sawchuk, A. A. (2013). Human Daily Activity Recognition With Sparse Representation Using Wearable Sensors. *IEEE Journal of Biomedical and Health Informatics*, 17(3), 553-560.
- Zhou, H., & Hu, H. (2008). Human motion tracking for rehabilitation - A survey. *Biomedical Signal Processing and Control*, 3(1), 1-18.
- Zhu, C., & Sheng, W. (2009). Human daily activity recognition in robot-assisted living using multi-sensor fusion. *IEEE International Conference In Robotics and Automation*, (pp. 2154-2159).

ANEXO I

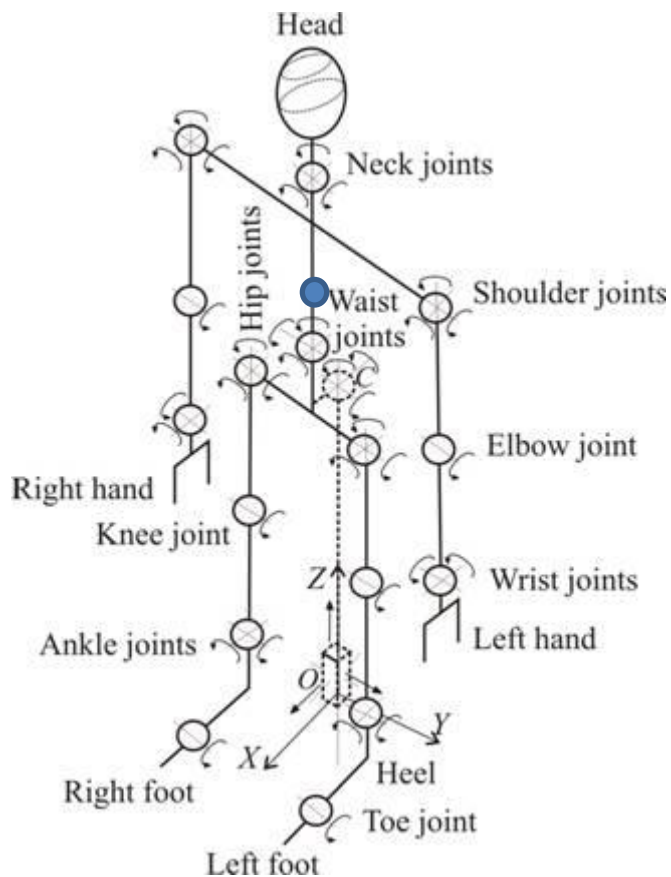
Publicações Relacionadas

Vital, J. P., Couceiro, M. S., Dias, G., & Ferreira, N. M. (2015). Tecnologias para a análise do movimento humano. Em R. Ruben, M. Vieira, C. Campos, H. Almeida, J. Siopa, P. Bártolo, & J. Folgado (Edits.), *6º Congresso Nacional de Biomecânica* (pp. 1-6). Instituto Politécnico de Leiria.

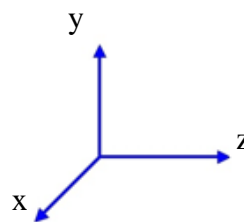
Vital, J. P., Faria, D. R., Dias, G., Couceiro, M. S., Coutinho, F., & Ferreira, N. M. (under review). Combining discriminative spatio-temporal features for daily life activity recognition using wearable motion sensing suit. *Pattern Analysis and Applications*.

ANEXO II

Modelo Denavit Hartenberg



Sistema de eixos no ponto inicial assinalado



- Braço Direito

	X		Z	
	Rotação	Translação	Rotação	Translação
	α_i	a_i	θ_i	d_i
1	θ_x	0	$\theta_z + 90^\circ$	0
2	θ_y	0	0	0
3	0	$\frac{l_{y1}}{2}$	0	$\frac{l_{z1}}{2}$
4	θ_{x1}	0	$\theta_{z1} - 90^\circ$	0
5	θ_{y1}	0	$+90^\circ$	0
6	θ_{x2}	Rx	$\theta_{z2} - 90^\circ$	$\frac{R_{z1}}{2}$
7	θ_{y2}	0	$+90^\circ$	0
8	θ_{x3}	0	$\theta_{z3} - 90^\circ$	$-\frac{R_{z1}}{2} - \frac{R_{z2}}{2}$
9	θ_{y3}	0	$+90^\circ$	0
10	θ_{x4}	0	$\theta_{z4} - 90^\circ$	$-\frac{R_{z2}}{2} - \frac{R_{z3}}{2}$
11	θ_{y4}	0	0	0

Matrizes de Translação e Rotação

$$T_0^1 = \begin{bmatrix} \cos(\theta_z + 90) & -\sin(\theta_z + 90) & 0 & 0 \\ \cos(\theta_x) \sin(\theta_z + 90) & \cos(\theta_x) \cos(\theta_z + 90) & -\sin(\theta_x) & 0 \\ \sin(\theta_x) \sin(\theta_z + 90) & \sin(\theta_x) \cos(\theta_z + 90) & \cos(\theta_x) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_1^2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta_y) & -\sin(\theta_y) & 0 \\ 0 & \sin(\theta_y) & \cos(\theta_y) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_2^3 = \begin{bmatrix} 1 & 0 & 0 & \frac{l_{y1}}{2} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \frac{l_{z1}}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_3^4 = \begin{bmatrix} \cos(\theta_{z1} - 90) & -\sin(\theta_{z1} - 90) & 0 & 0 \\ \cos(\theta_{x1}) \sin(\theta_{z1} - 90) & \cos(\theta_{x1}) \cos(\theta_{z1} - 90) & -\sin(\theta_{x1}) & 0 \\ \sin(\theta_{x1}) \sin(\theta_{z1} - 90) & \sin(\theta_{x1}) \cos(\theta_{z1} - 90) & \cos(\theta_{x1}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_4^5 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ \cos(\theta_{y1}) & 0 & -\sin(\theta_{y1}) & 0 \\ \sin(\theta_{y1}) & 0 & \cos(\theta_{y1}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_5^6 = \begin{bmatrix} \cos(\theta_{z2} - 90) & -\sin(\theta_{z2} - 90) & 0 & R_x \\ \cos(\theta_{x2}) \sin(\theta_{z2} - 90) & \cos(\theta_{x2}) \cos(\theta_{z2} - 90) & -\sin(\theta_{x2}) & \sin(\theta_{x2}) * \frac{R_{z1}}{2} \\ \sin(\theta_{x2}) \sin(\theta_{z2} - 90) & \sin(\theta_{x2}) \cos(\theta_{z2} - 90) & \cos(\theta_{x2}) & -\cos(\theta_{x2}) * \frac{R_{z1}}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_6^7 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ \cos(\theta_{y2}) & 0 & -\sin(\theta_{y2}) & 0 \\ \sin(\theta_{y2}) & 0 & \cos(\theta_{y2}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_7^8 = \begin{bmatrix} \cos(\theta_{z3} - 90) & -\sin(\theta_{z3} - 90) & 0 & R_x \\ \cos(\theta_{x3}) \sin(\theta_{z3} - 90) & \cos(\theta_{x3}) \cos(\theta_{z3} - 90) & -\sin(\theta_{x3}) & -\sin(\theta_{x3}) * \frac{R_{z1}}{2} - \frac{R_{z2}}{2} \\ \sin(\theta_{x3}) \sin(\theta_{z3} - 90) & \sin(\theta_{x3}) \cos(\theta_{z3} - 90) & \cos(\theta_{x3}) & \cos(\theta_{x3}) * \frac{R_{z1}}{2} - \frac{R_{z2}}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_8^9 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ \cos(\theta_{y3}) & 0 & -\sin(\theta_{y3}) & 0 \\ \sin(\theta_{y3}) & 0 & \cos(\theta_{y3}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_9^{10} = \begin{bmatrix} \cos(\theta_{z4} - 90) & -\sin(\theta_{z4} - 90) & 0 & 0 \\ \cos(\theta_{x4}) \sin(\theta_{z4} - 90) & \cos(\theta_{x4}) \cos(\theta_{z4} - 90) & -\sin(\theta_{x4}) & -\sin(\theta_{x4}) * \frac{R_{z2}}{2} - \frac{R_{z3}}{2} \\ \sin(\theta_{x4}) \sin(\theta_{z4} - 90) & \sin(\theta_{x4}) \cos(\theta_{z4} - 90) & \cos(\theta_{x4}) & \cos(\theta_{x4}) * \frac{R_{z2}}{2} - \frac{R_{z3}}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_{10}^{11} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta_{y4}) & -\sin(\theta_{y4}) & 0 \\ 0 & \sin(\theta_{y4}) & \cos(\theta_{y4}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

• Braço Esquerdo

	X		Z	
	Rotação	Translação	Rotação	Translação
	α_i	a_i	θ_i	d_i
1	θ_x	0	$\theta_z + 90^\circ$	0
2	θ_y	0	0	0
3	0	$-\frac{l_{y1}}{2}$	0	$\frac{l_{z1}}{2}$
4	θ_{x5}	0	$\theta_{z5} - 90^\circ$	0
5	θ_{y5}	0	$+90^\circ$	0
6	θ_{x6}	$-lx$	$\theta_{z6} - 90^\circ$	$-\frac{l_{z2}}{2}$
7	θ_{y6}	0	$+90^\circ$	0
8	θ_{x7}	0	$\theta_{z7} - 90^\circ$	$-\frac{l_{z2}}{2} - \frac{l_{z3}}{2}$
9	θ_{y7}	0	$+90^\circ$	0
10	θ_{x8}	0	$\theta_{z8} - 90^\circ$	$-\frac{l_{z3}}{2} - \frac{l_{z4}}{2}$
11	θ_{y8}	0	0	0

Matrizes de Translação e Rotação

$$T_0^1 = \begin{bmatrix} \cos(\theta_z + 90) & -\sin(\theta_z + 90) & 0 & 0 \\ \cos(\theta_x) \sin(\theta_z + 90) & \cos(\theta_x) \cos(\theta_z + 90) & -\sin(\theta_x) & 0 \\ \sin(\theta_x) \sin(\theta_z + 90) & \sin(\theta_x) \cos(\theta_z + 90) & \cos(\theta_x) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_1^2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta_y) & -\sin(\theta_y) & 0 \\ 0 & \sin(\theta_y) & \cos(\theta_y) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_2^3 = \begin{bmatrix} 1 & 0 & 0 & -\frac{l_{y1}}{2} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \frac{l_{z1}}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_3^4 = \begin{bmatrix} \cos(\theta_{z5} - 90) & -\sin(\theta_{z5} - 90) & 0 & 0 \\ \cos(\theta_{x5}) \sin(\theta_{z5} - 90) & \cos(\theta_{x5}) \cos(\theta_{z5} - 90) & -\sin(\theta_{x5}) & 0 \\ \sin(\theta_{x5}) \sin(\theta_{z5} - 90) & \sin(\theta_{x5}) \cos(\theta_{z5} - 90) & \cos(\theta_{x5}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_4^5 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ \cos(\theta_{y5}) & 0 & -\sin(\theta_{y5}) & 0 \\ \sin(\theta_{y5}) & 0 & \cos(\theta_{y5}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_5^6 = \begin{bmatrix} \cos(\theta_{z6} - 90) & -\sin(\theta_{z6} - 90) & 0 & -l_x \\ \cos(\theta_{x6}) \sin(\theta_{z6} - 90) & \cos(\theta_{x6}) \cos(\theta_{z6} - 90) & -\sin(\theta_{x6}) & \sin(\theta_{x6}) * \frac{l_{z2}}{2} \\ \sin(\theta_{x6}) \sin(\theta_{z6} - 90) & \sin(\theta_{x6}) \cos(\theta_{z6} - 90) & \cos(\theta_{x6}) & -\cos(\theta_{x6}) * \frac{l_{z2}}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_6^7 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ \cos(\theta_{y6}) & 0 & -\sin(\theta_{y6}) & 0 \\ \sin(\theta_{y6}) & 0 & \cos(\theta_{y6}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_7^8 = \begin{bmatrix} \cos(\theta_{z7} - 90) & -\sin(\theta_{z7} - 90) & 0 & 0 \\ \cos(\theta_{x7}) \sin(\theta_{z7} - 90) & \cos(\theta_{x7}) \cos(\theta_{z7} - 90) & -\sin(\theta_{x7}) & -\sin(\theta_{x7}) * \frac{l_{z2}}{2} - \frac{l_{z3}}{2} \\ \sin(\theta_{x7}) \sin(\theta_{z7} - 90) & \sin(\theta_{x7}) \cos(\theta_{z7} - 90) & \cos(\theta_{x7}) & \cos(\theta_{x7}) * \frac{l_{z2}}{2} - \frac{l_{z3}}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_8^9 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ \cos(\theta_{y7}) & 0 & -\sin(\theta_{y7}) & 0 \\ \sin(\theta_{y7}) & 0 & \cos(\theta_{y7}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_9^{10} = \begin{bmatrix} \cos(\theta_{z8} - 90) & -\sin(\theta_{z8} - 90) & 0 & 0 \\ \cos(\theta_{x8}) \sin(\theta_{z8} - 90) & \cos(\theta_{x8}) \cos(\theta_{z8} - 90) & -\sin(\theta_{x8}) & -\sin(\theta_{x8}) * \frac{l_{z3}}{2} - \frac{l_{z4}}{2} \\ \sin(\theta_{x8}) \sin(\theta_{z8} - 90) & \sin(\theta_{x8}) \cos(\theta_{z8} - 90) & \cos(\theta_{x8}) & \cos(\theta_{x8}) * \frac{l_{z3}}{2} - \frac{l_{z4}}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_{10}^{11} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta_{y8}) & -\sin(\theta_{y8}) & 0 \\ 0 & \sin(\theta_{y8}) & \cos(\theta_{y8}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- Perna Direita

	X		Z	
	Rotação	Translação	Rotação	Translação
	α_i	a_i	θ_i	d_i
1	θ_x	0	$\theta_z + 90^\circ$	0
2	θ_y	0	0	0
3	0	$\frac{l_{y1}}{2}$	0	$-\frac{l_{z1}}{2}$
4	θ_{x9}	0	$\theta_{z9} - 90^\circ$	0
5	θ_{y9}	0	$+90^\circ$	0
6	θ_{x10}	$\frac{l_{y1}}{2} - R_{px}$	$\theta_{z10} - 90^\circ$	$-\frac{R_{pz1}}{2}$
7	θ_{y10}	0	$+90^\circ$	0
8	θ_{x11}	0	$\theta_{z11} - 90^\circ$	$-\frac{R_{pz1}}{2} - \frac{R_{pz2}}{2}$
9	θ_{y11}	0	$+90^\circ$	0
10	θ_{x12}	0	$\theta_{z12} - 90^\circ$	$-\frac{R_{pz2}}{2} - \frac{R_{pz3}}{2}$
11	θ_{y12}	0	0	0

Matrizes de Translação e Rotação

$$T_0^1 = \begin{bmatrix} \cos(\theta_z + 90) & -\sin(\theta_z + 90) & 0 & 0 \\ \cos(\theta_x) \sin(\theta_z + 90) & \cos(\theta_x) \cos(\theta_z + 90) & -\sin(\theta_x) & 0 \\ \sin(\theta_x) \sin(\theta_z + 90) & \sin(\theta_x) \cos(\theta_z + 90) & \cos(\theta_x) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_1^2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta_y) & -\sin(\theta_y) & 0 \\ 0 & \sin(\theta_y) & \cos(\theta_y) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_2^3 = \begin{bmatrix} 1 & 0 & 0 & \frac{l_{y1}}{2} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -\frac{l_{z1}}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_3^4 = \begin{bmatrix} \cos(\theta_{z9} - 90) & -\sin(\theta_{z9} - 90) & 0 & 0 \\ \cos(\theta_{x9}) \sin(\theta_{z9} - 90) & \cos(\theta_{x9}) \cos(\theta_{z9} - 90) & -\sin(\theta_{x9}) & 0 \\ \sin(\theta_{x9}) \sin(\theta_{z9} - 90) & \sin(\theta_{x9}) \cos(\theta_{z9} - 90) & \cos(\theta_{x9}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_4^5 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ \cos(\theta_{y9}) & 0 & -\sin(\theta_{y9}) & 0 \\ \sin(\theta_{y9}) & 0 & \cos(\theta_{y9}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_5^6 = \begin{bmatrix} \cos(\theta_{z10} - 90) & -\sin(\theta_{z10} - 90) & 0 & \frac{l_{y1}}{2} - R_{px} \\ \cos(\theta_{x10}) \sin(\theta_{z10} - 90) & \cos(\theta_{x10}) \cos(\theta_{z10} - 90) & -\sin(\theta_{x10}) & \sin(\theta_{x10}) * \frac{R_{pz1}}{2} \\ \sin(\theta_{x10}) \sin(\theta_{z10} - 90) & \sin(\theta_{x10}) \cos(\theta_{z10} - 90) & \cos(\theta_{x10}) & -\cos(\theta_{x10}) * \frac{R_{pz1}}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_6^7 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ \cos(\theta_{y10}) & 0 & -\sin(\theta_{y10}) & 0 \\ \sin(\theta_{y10}) & 0 & \cos(\theta_{y10}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_7^8 = \begin{bmatrix} \cos(\theta_{z11} - 90) & -\sin(\theta_{z11} - 90) & 0 & 0 \\ \cos(\theta_{x11}) \sin(\theta_{z11} - 90) & \cos(\theta_{x11}) \cos(\theta_{z11} - 90) & -\sin(\theta_{x11}) & -\sin(\theta_{x11}) * \frac{R_{pz1}}{2} - \frac{R_{pz2}}{2} \\ \sin(\theta_{x11}) \sin(\theta_{z11} - 90) & \sin(\theta_{x11}) \cos(\theta_{z11} - 90) & \cos(\theta_{x11}) & \cos(\theta_{x11}) * \frac{R_{pz1}}{2} - \frac{R_{pz2}}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_8^9 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ \cos(\theta_{y11}) & 0 & -\sin(\theta_{y11}) & 0 \\ \sin(\theta_{y11}) & 0 & \cos(\theta_{y11}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_9^{10} = \begin{bmatrix} \cos(\theta_{z12} - 90) & -\sin(\theta_{z12} - 90) & 0 & 0 \\ \cos(\theta_{x12}) \sin(\theta_{z12} - 90) & \cos(\theta_{x12}) \cos(\theta_{z12} - 90) & -\sin(\theta_{x12}) & -\sin(\theta_{x12}) * -\frac{R_{pz2}}{2} - \frac{R_{pz3}}{2} \\ \sin(\theta_{x12}) \sin(\theta_{z12} - 90) & \sin(\theta_{x12}) \cos(\theta_{z12} - 90) & \cos(\theta_{x12}) & \cos(\theta_{x12}) * -\frac{R_{pz2}}{2} - \frac{R_{pz3}}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_{10}^{11} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta_{y12}) & -\sin(\theta_{y12}) & 0 \\ 0 & \sin(\theta_{y12}) & \cos(\theta_{y12}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

• Perna Esquerda

	X		Z	
	Rotação	Translação	Rotação	Translação
	α_i	a_i	θ_i	d_i
1	θ_x	0	$\theta_z + 90^\circ$	0
2	θ_y	0	0	0
3	0	$-\frac{l_{y1}}{2}$	0	$-\frac{l_{z1}}{2}$
4	θ_{x13}	0	$\theta_{z13} - 90^\circ$	0
5	θ_{y13}	0	$+90^\circ$	0
6	θ_{x14}	$-l_{px}$	$\theta_{z14} - 90^\circ$	$-\frac{l_{pz2}}{2}$
7	θ_{y14}	0	$+90^\circ$	0
8	θ_{x15}	0	$\theta_{z15} - 90^\circ$	$-\frac{l_{pz2}}{2} - \frac{l_{pz3}}{2}$
9	θ_{y15}	0	$+90^\circ$	0
10	θ_{x16}	0	$\theta_{z16} - 90^\circ$	$-\frac{l_{pz3}}{2} - \frac{l_{pz4}}{2}$
11	θ_{y16}	0	0	0

Matrizes de Translação e Rotação

$$T_0^1 = \begin{bmatrix} \cos(\theta_z + 90) & -\sin(\theta_z + 90) & 0 & 0 \\ \cos(\theta_x) \sin(\theta_z + 90) & \cos(\theta_x) \cos(\theta_z + 90) & -\sin(\theta_x) & 0 \\ \sin(\theta_x) \sin(\theta_z + 90) & \sin(\theta_x) \cos(\theta_z + 90) & \cos(\theta_x) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_1^2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta_y) & -\sin(\theta_y) & 0 \\ 0 & \sin(\theta_y) & \cos(\theta_y) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_2^3 = \begin{bmatrix} 1 & 0 & 0 & -\frac{l_{y1}}{2} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -\frac{l_{z1}}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_3^4 = \begin{bmatrix} \cos(\theta_{z13} - 90) & -\sin(\theta_{z13} - 90) & 0 & 0 \\ \cos(\theta_{x13}) \sin(\theta_{z13} - 90) & \cos(\theta_{x13}) \cos(\theta_{z13} - 90) & -\sin(\theta_{x13}) & 0 \\ \sin(\theta_{x13}) \sin(\theta_{z13} - 90) & \sin(\theta_{x13}) \cos(\theta_{z13} - 90) & \cos(\theta_{x13}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_4^5 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ \cos(\theta_{y13}) & 0 & -\sin(\theta_{y13}) & 0 \\ \sin(\theta_{y13}) & 0 & \cos(\theta_{y13}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_5^6 = \begin{bmatrix} \cos(\theta_{z14} - 90) & -\sin(\theta_{z14} - 90) & 0 & -l_{px} \\ \cos(\theta_{x14}) \sin(\theta_{z14} - 90) & \cos(\theta_{x14}) \cos(\theta_{z14} - 90) & -\sin(\theta_{x14}) & \sin(\theta_{x14}) * \frac{l_{pz2}}{2} \\ \sin(\theta_{x14}) \sin(\theta_{z14} - 90) & \sin(\theta_{x14}) \cos(\theta_{z14} - 90) & \cos(\theta_{x14}) & -\cos(\theta_{x14}) * -\frac{l_{pz2}}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_6^7 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ \cos(\theta_{y14}) & 0 & -\sin(\theta_{y14}) & 0 \\ \sin(\theta_{y14}) & 0 & \cos(\theta_{y14}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_7^8 = \begin{bmatrix} \cos(\theta_{z15} - 90) & -\sin(\theta_{z15} - 90) & 0 & 0 \\ \cos(\theta_{x15}) \sin(\theta_{z15} - 90) & \cos(\theta_{x15}) \cos(\theta_{z15} - 90) & -\sin(\theta_{x15}) & -\sin(\theta_{x15}) * \frac{l_{pZ2}}{2} - \frac{l_{pZ3}}{2} \\ \sin(\theta_{x15}) \sin(\theta_{z15} - 90) & \sin(\theta_{x15}) \cos(\theta_{z15} - 90) & \cos(\theta_{x15}) & \cos(\theta_{x15}) * \frac{l_{pZ2}}{2} - \frac{l_{pZ3}}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_8^9 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ \cos(\theta_{y15}) & 0 & -\sin(\theta_{y15}) & 0 \\ \sin(\theta_{y15}) & 0 & \cos(\theta_{y15}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_9^{10} = \begin{bmatrix} \cos(\theta_{z16} - 90) & -\sin(\theta_{z8} - 90) & 0 & 0 \\ \cos(\theta_{x16}) \sin(\theta_{z16} - 90) & \cos(\theta_{x16}) \cos(\theta_{z16} - 90) & -\sin(\theta_{x16}) & -\sin(\theta_{x16}) * \frac{l_{pZ3}}{2} - \frac{l_{pZ4}}{2} \\ \sin(\theta_{x16}) \sin(\theta_{z16} - 90) & \sin(\theta_{x16}) \cos(\theta_{z16} - 90) & \cos(\theta_{x16}) & \cos(\theta_{x16}) * \frac{l_{pZ3}}{2} - \frac{l_{pZ4}}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_{10}^{11} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta_{y16}) & -\sin(\theta_{y16}) & 0 \\ 0 & \sin(\theta_{y16}) & \cos(\theta_{y16}) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$