

 M 2015

**U. PORTO**  
FEUP FACULDADE DE ENGENHARIA  
UNIVERSIDADE DO PORTO

# NEUROKINECT - QUANTIFICAÇÃO 3D DE MOVIMENTOS EM DOENÇAS NEUROLÓGICAS

**DANIELA MARIA FRAZÃO COSTA**  
DISSERTAÇÃO DE MESTRADO APRESENTADA  
À FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO EM  
ENGENHARIA BIOMÉDICA

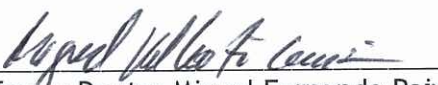



A Dissertação intitulada

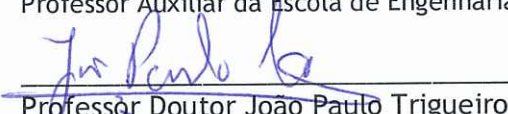
“NeuroKinect - Quantificação 3D de movimentos em doenças neurológicas”

foi aprovada em provas realizadas em 16-10-2015

o júri

  
Presidente Professor Doutor Miguel Fernando Paiva Velhote Correia  
Professor Auxiliar do Departamento de Engenharia Eletrotécnica e de Computadores  
da Faculdade de Engenharia da U. Porto

  
Professor Doutor André Paulo Almeida Whiteman Catarino  
Professor Auxiliar da Escola de Engenharia da U. Minho

  
Professor Doutor João Paulo Trigueiros da Silva Cunha  
Professor Associado do Departamento de Engenharia Eletrotécnica e de  
Computadores da Faculdade de Engenharia da U. Porto

O autor declara que a presente dissertação (ou relatório de projeto) é da sua exclusiva autoria e foi escrita sem qualquer apoio externo não explicitamente autorizado. Os resultados, ideias, parágrafos, ou outros extratos tomados de ou inspirados em trabalhos de outros autores, e demais referências bibliográficas usadas, são corretamente citados.

  
Autor - Daniela Maria Frazão Costa

Faculdade de Engenharia da Universidade do Porto





FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO



# **Neurokinect - Quantificação 3D de Movimentos em Doenças Neurológicas**

**Daniela Maria Frazão Costa**

VERSÃO DE TRABALHO

Mestrado em Engenharia Biomédica

Orientador: PhD João Paulo Cunha

30 de Outubro de 2015



# **Neurokinect - Quantificação 3D de Movimentos em Doenças Neurológicas**

**Daniela Maria Frazão Costa**

Mestrado em Engenharia Biomédica

Faculdade de Engenharia da Universidade do Porto

30 de Outubro de 2015



# Resumo

O sucesso das avaliações de epilepsia reside essencialmente na avaliação da técnica de Vídeo-Eletroencefalograma (Vídeo-EEG). Neste caso, os movimentos feitos pelos pacientes podem conseguir caracterizar o foco epilético e delinear com precisão as áreas a remover. Para os pacientes submetidos a cirurgia, a delineação da área alvo antes da cirurgia é fundamental, pelo que quanto maior rigor se conseguir caracterizar o tipo de epilepsia maior facilidade terá o médico de estabelecer a conexão entre os movimentos e a área cerebral afetada.

Para este estudo, desenvolveu-se uma interface em que é possível visualizar o vídeo em que ocorre as crises. Neste vídeo, é possibilitado ao utilizador marcar 8 pontos de interesse, neste caso serão a cabeça, torso, ombros, cotovelos e pulsos. Para além da visualização do vídeo tem também gráficos ajustáveis à preferência do utilizador, relativos aos movimentos dos membros. Para posterior análise, a interface também apresenta uma visualização 3D dos pontos previamente selecionados.

Para o seguimento destes pontos selecionados, é inicialmente calculada sua distância Euclidiana, em coordenadas mundo (metros), pelo que foi usada a profundidade adquirida pelo sensor (Kinect). O algoritmo de seguimento pressupõe dois blocos principais: estimativa e correção. A estimativa do novo ponto a cada *frame* é calculada através do fluxo ótico de densidade, pelo algoritmo de Farneback. Este algoritmo tem em conta a sua vizinhança e verifica o fluxo integral da região de interesse ao aproximar cada vizinho a um polinómio quadrático, após a diferença entre *frames* ser suavizada com um filtro Gaussiano. A correção é feita de modo a corrigir os pequenos erros relativos ao método anterior, em que verifica a similaridade de distâncias dos membros, através da profundidade e converte para coordenadas mundo para obter a distância em metros.

Este tipo de avaliação de movimento permite estabelecer relações entre os tipos de crises e posteriormente ajudar na decisão do médico.

**Palavras-chave:** Epilepsia, Vídeo-EEG, Kinect, seguimento, Farneback, similaridade, 3D.



# Abstract

The success of the Epilepsy evaluation relies specially on evaluating Video-Electroencephalogram (Video-EEG). In this case, uncoordinated movements can relate and characterize the epileptic focus and its involving areas. For patients that are submitted to surgery, the delineation of the target area before the surgery is fundamental, so as to achieve greater accuracy characterize the type of epilepsy best medical facility will have to establish the connection between the movement and the affected brain area.

For this study, it was developed a software in which is possible to visualize the video recording where it occurs the seizures. In the video it is possible to mark in the first frame 8 points of interest, joints of the body, which are head, torso, shoulders, elbows and wrists. Besides recording the video it also contains graphics of velocity, acceleration and positions of the points previously selected. For after analysis, the software also presents a 3D visualization of these same points.

It was proposed a tracking method for these points. After the selection of these points, the distances of the limbs are calculated, in world coordinates, using the depth acquired from the sensor, Kinect v2. The algorithm has 2 main blocks: estimation and correction. The estimation of the new point is achieved by using the Farneback algorithm, which is a dense optical flow method. This algorithm evaluates its neighbors and verifies the flow approximating each neighbor of both frames by quadratic polynomials, after the frames are smooth with a Gaussian filter. The correction is made when there is errors in the previous method, which is observed when the similarity of the distance of the limbs are less than a threshold.

This kind of movement analysis can establish relations between the type of seizures and after help in the doctor's decision.

**Key words:** Epilepsy, Video-EEG, Kinect, tracking, Farneback, similarity, 3D.





# Agradecimentos

Será difícil colocar em palavras todo o agradecimento que tenho pelas pessoas que direta ou indiretamente me ajudaram a alcançar os meus objetivos durante esta longa viagem de 5 anos.

Em primeiro lugar, gostava de agradecer profundamente a todo o amor, apoio, paciência e disponibilidade que os meus pais e irmãos me ofereceram ao longo do tempo. Sem eles, nada disto podia ter sido alcançado.

Gostava também de demonstrar o meu agradecimento ao grupo BRAIN, a todas as *brainMeetings* e *brainStorms* que foram proporcionadas ao longo do tempo e sem dúvida me ajudaram a evoluir. Em particular, gostava de mostrar a minha gratidão ao Prof. João Paulo Cunha que mostrou sempre paciência e vontade para melhorar, assim como ao Hugo Choupina e Pedro Costa por mostrarem sempre disponibilidade para ajudar.

À todos os meus amigos, agradeço pelo apoio e amizade incondicional. Aos que surgiram da vida académica, um especial obrigado por me proporcionarem das melhores memórias que poderia adquirir.

Por fim, um especial obrigada ao meu namorado, Mário, que tornou esta viagem tão mais fácil. Agradeço especialmente pela paciência, amor incondicional, carinho e sobretudo pela vontade própria de me ajudar e participar em tudo o que faço. Estes últimos anos não teriam sido tão especiais sem ti.

Daniela Costa



*“The best way to predict the future is to invent it.”*

Alan Kay



# Conteúdo

<b>Lista de Figuras</b>	<b>xi</b>
<b>Lista de Tabelas</b>	<b>xv</b>
<b>Abreviaturas e Símbolos</b>	<b>xvii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Enquadramento . . . . .	1
1.2 Contribuição . . . . .	2
1.3 Estrutura da Dissertação . . . . .	2
<b>2 Revisão Bibliográfica</b>	<b>3</b>
2.1 Epilepsia . . . . .	3
2.1.1 Tipos de Crise . . . . .	3
2.1.2 Diagnóstico . . . . .	5
2.1.3 Vídeo-EEG . . . . .	5
2.1.4 Mortalidade . . . . .	7
2.1.5 Tratamento . . . . .	8
2.2 Técnicas de Análise Motora . . . . .	8
2.3 Captura de Movimento . . . . .	10
2.4 Sensores RGB-D . . . . .	13
2.4.1 Sensores <i>Time-of-Flight</i> . . . . .	14
2.4.2 Microsoft Kinect . . . . .	15
2.5 Quantificação de Movimento em Epilepsia . . . . .	16

2.6	Proposta de trabalho . . . . .	19
<b>3</b>	<b>Seguimento de Membros Semi-Automático</b>	<b>21</b>
3.1	Pré-processamento . . . . .	21
3.2	Seguimento dos membros . . . . .	22
3.2.1	Estimativa: Algoritmo Farneback . . . . .	22
3.2.2	Correção: Similaridade entre pontos . . . . .	25
3.2.3	Aquisição das coordenadas mundo . . . . .	25
3.3	Resultados do Seguimento de Membros . . . . .	26
3.3.1	Caracterização dos Dados . . . . .	36
3.4	Validação do Algoritmo . . . . .	43
<b>4</b>	<b>Interface: Sistema de Análise de Movimento 3D</b>	<b>49</b>
4.1	NeuroKinect, KiT e KIMA . . . . .	49
4.2	Requisitos . . . . .	49
4.3	Aplicação . . . . .	51
<b>5</b>	<b>Conclusões e Trabalho Futuro</b>	<b>55</b>
5.1	Revisão dos Objetivos . . . . .	55
5.2	Contribuições da Dissertação . . . . .	55
5.3	Trabalho Futuro . . . . .	57
	<b>Referências</b>	<b>59</b>

# Lista de Figuras

2.1	Tipos de crises e respectivas características. Adaptado de [1] . . . . .	6
2.2	Esquema ilustrativo da Unidade de Monitorização. Adaptado de [2] . . . . .	7
2.3	Modelo Vareta. Adaptado de [3] . . . . .	9
2.4	Esquema de análise motora. Adaptado de [3] . . . . .	11
2.5	Esquemas de diversos tipos de técnicas de captura de movimento. Adaptado de [4]	12
2.6	Conceito de <i>Time-of-Flight</i> . Adaptado de [5] . . . . .	13
2.7	Conceito de <i>Structured – Light</i> . Adaptado de [5] . . . . .	14
2.8	Diferentes tecnologias de medição de profundidade. Adaptado de [6] . . . . .	15
2.9	Microsoft Kinect KW4. Adaptado de [7] . . . . .	16
2.10	Juntas relativas ao sensor v1 e v2. Adaptado de [7] . . . . .	16
2.11	Características da Kinect v2, K4W . . . . .	17
3.1	Resultado da Equalização de Histograma (antes e depois) . . . . .	22
3.2	Exemplo do resultado da diferença entre <i>frames</i> com filtro Gaussiano. As regiões de maior contraste indicam as regiões onde ocorreu movimento . . . . .	24
3.3	Articulações Corpo Humano. Adaptado de [8] . . . . .	28
3.4	Movimento do Pescoço. Adaptado de [8] . . . . .	29
3.5	Movimento da Lombar. Adaptado de [8] . . . . .	30
3.6	Movimento dos ombros. Adaptado de [8] . . . . .	31
3.7	Movimento do cotovelo. Adaptado de [8] . . . . .	32
3.8	Movimento do cotovelo. Adaptado de [8] . . . . .	32
3.9	Movimento do pulso. Adaptado de [8] . . . . .	33
3.10	Diagrama de Atividade UML do algoritmo proposto . . . . .	34

3.11	exatidão relativa à mudança da largura da janela de vizinhança . . . . .	35
3.12	Tempo despendido relativamente à mudança da largura da janela de vizinhança . . . . .	35
3.13	Configuração de Aquisição . . . . .	36
3.14	Representação do movimento ao longo do tempo durante a crise quantificada . . . . .	38
3.15	Traçado da velocidade e aceleração 3D na cabeça ao longo do tempo (em segundos) . . . . .	39
3.16	Traçado da velocidade e aceleração 3D no torso ao longo do tempo (em segundos) . . . . .	39
3.17	Traçado da velocidade e aceleração 3D do ombro direito ao longo do tempo (em segundos) . . . . .	40
3.18	Traçado da velocidade e aceleração 3D do ombro esquerdo ao longo do tempo (em segundos) . . . . .	40
3.19	Traçado da velocidade e aceleração 3D do cotovelo direito ao longo do tempo (em segundos) . . . . .	41
3.20	Traçado da velocidade e aceleração 3D do cotovelo esquerdo ao longo do tempo (em segundos) . . . . .	41
3.21	Traçado da velocidade e aceleração 3D do pulso direito ao longo do tempo (em segundos) . . . . .	42
3.22	Traçado da velocidade e aceleração 3D do pulso esquerdo ao longo do tempo (em segundos) . . . . .	42
3.23	Traçado dos centróides dos movimentos das juntas do corpo na coordenada X (em pixels por <i>frame</i> ) . . . . .	43
3.24	Traçado dos centróides dos movimentos das juntas do corpo na coordenada Y (em pixels por <i>frame</i> ) . . . . .	44
3.25	Traçado dos centróides dos movimentos das juntas do corpo na coordenada Z (em metros por <i>frame</i> ) . . . . .	44
3.26	Traçado da média do erro dos centróides (X e Y) das juntas do corpo ao longo do tempo (em <i>frames</i> ) . . . . .	44
3.27	Sequência de <i>frames</i> relativamente ao método usado sem correção (SC) e com correção (CC) . . . . .	46
3.28	Exatidão em X e Y para cada crise avaliada . . . . .	48
4.1	Diagrama de atividades UML da aplicação . . . . .	50



4.2	Diagrama de casos de uso da aplicação . . . . .	51
4.3	Janela de edição de gráficos a visualizar . . . . .	52
4.4	Janela principal da aplicação. Vermelho - menu que permite personalização. Azul - Manipulação do Vídeo. Verde - Visualização dos gráficos. Laranja - Visualização 3D do esqueleto. Amarelo - Comentários . . . . .	53
4.5	Exemplo do ficheiro guardado em <i>Excel</i> . Cada <i>sheet</i> apresenta os resultados ad- quiridos para cada articulação . . . . .	53



# Lista de Tabelas

3.1	Parâmetros Intrínsecos Estimados da Kinect v2, adaptado de [9] . . . . .	26
3.2	Caracterização das crises avaliadas . . . . .	37
3.3	Percentagem de erro em cada articulação . . . . .	47



# Abreviaturas e Símbolos

FEUP	Faculdade de Engenharia da Universidade do Porto
EEG	Eletroencefalograma
BRAIN	Biomedical Research And INnovation
WHO	World Health Organization
RGB-D	Red Green Blue - Depth
RGB	Red Green Blue
ILAE	League Against Epilepsy
AED	Anti Epileptic Drugs
3D	Three Dimensions
2D	Two Dimensions
TLE	Temporal Lobe Epilepsy
FLE	Frontal Lobe Epilepsy
ToF	Time-of-Flight
CCD	Charge-Coupled Device
LED	Light Emitting Diode
IR	Infrared
SDK	Software Development Kit
API	Application Programming Interface
MATLAB	MATrix LABoratory
ROI	Region Of Interest
UML	Unified Modeling Language

C	Cabeça
T	Torso
OE	Ombro Esquerdo
OD	Ombro Direito
CE	Cotovelo Esquerdo
CD	Cotovelo Direito
PE	Pulso Esquerdo
PD	Pulso Direito
SC	Sem Correção
CC	Com Correção
KiMA	Kinect Motion Analyser
KiT	Kinect Tracker

# Capítulo 1

## Introdução

### 1.1 Enquadramento

Movimentos descoordenados levam geralmente à falta de independência por parte dos pacientes e conseqüentemente a custos associados [10].

Epilepsia é considerada como um conjunto de sintomas neurológicos que afetam consideravelmente a parte motora, como movimentos incontroláveis, crises e outros fenómenos psíquicos. Os pacientes epiléticos podem mostrar resistência a fármacos, obrigando a uma intervenção cirúrgica. Numa fase inicial, um Vídeo-EEG é feito para avaliar movimentos durante as crises e associar ao foco epilético principal. Este método é crucial para capturar momentos de interesse que ajudam os profissionais de saúde a avaliar a(s) zona(s) a serem removidas, levando à diminuição das crises ou até mesmo parar com as mesmas. De acordo com a Organização Mundial de Saúde (WHO), doenças neurológicas têm-se tornado um assunto significativo e crescente, afetando, em 2010 mais de 450 milhões de pessoas mundialmente [11]. A Epilepsia é uma das doenças neurológicas mais comuns. A Epilepsia afeta 1 em 140 pessoas mundialmente, pelo que sugere que cerca de 75 mil pessoas em Portugal são portadoras desta condição e 30% de epilepsia refratária, em que requer intervenção cirúrgica [12]. Assim sendo, é necessário prestar mais cuidado a este assunto e desenvolver procedimentos clínicos que conseqüentemente melhorem a qualidade de vida, assim como a redução dos custos inerentes aos cuidados e fármacos necessárias para controlar esta condição. Um dos exames diagnósticos de epilepsia é a monitorização por Vídeo-EEG, que consiste na aquisição de sinais elétricos das atividades cerebrais de forma não-invasiva, concomitante com as imagens do paciente em tempo real. Este registo permite a distinção entre momentos de crises epiléticas dos eventos paroxísticos não-epiléticos. Esta documentação exige várias horas de observação do paciente de forma a qualificar o tipo de crise associada ao evento. Na maioria das vezes, este procedimento é analisado qualitativamente e depende muito da experiência do profissional de saúde. A análise de movimento foi frequentemente feita através de marcadores, que exige despender tempo na aprendizagem da colocação e na própria colocação em cada paciente, pelo que não é adequado para um ambiente agitado hospitalar.

De forma a assegurar o conforto do paciente, assim como a precisão dos dados, surge a ideia

de usar sensores RGB-D. Estes permitem a quantificação sem marcação dos movimentos, melhorando significativamente tanto a precisão de aquisição de dados, como a redução de custos elevados neste tipo de monitorização.

Os sensores RGB-D adquirem informação de profundidade que pode ser aproveitada na aquisição dos pontos no mundo em 3D de qualquer cenário, sem interromper com a rotina hospitalar e sem comprometer o conforto do paciente, pelo que apenas é necessário ser posicionado em frente do paciente.

Nesta dissertação será usado um sensor RGB-D que permite a aquisição do cenário de monitorização de epilepsia e dos eventos inerentes ao mesmo. De forma a quantificar os movimentos será criado um modelo 3D semi-automático de seguimento dos membros e uma interface que facilite a sua utilização.

## 1.2 Contribuição

Nesta dissertação, a principal contribuição é desenvolver um método capaz de, com auxílio do sensor RGB-D Microsoft Kinect v2, seguir sem marcação física os membros do corpo do paciente de epilepsia aquando sujeitos a monitorização de Vídeo-EEG. Este método é capaz de seguir os movimentos do paciente, tendo em conta os artefactos inerentes ao ambiente hospitalar. Estes artefactos incluem a oclusão frequente de membros, as posições pouco comuns aquando das crises e os movimentos bruscos provocados pelas mesmas. Este método permitirá então a estimação correta da posição dos membros.

De forma a facilitar o uso deste mesmo método é também desenvolvida uma interface Microsoft Visual Studio, em linguagem C# capaz de guardar todos os valores por ele adquiridos, como a velocidade 3D, aceleração 3D e posição de cada membro 3D.

## 1.3 Estrutura da Dissertação

Em adição à Introdução, esta dissertação contém mais 5 capítulos.

No capítulo 2 é feita uma revisão bibliográfica das contribuições científicas mais relevantes nos últimos anos, assim como uma explicação breve do cenário a ser estudado.

No capítulo 3 é apresentado o método utilizado para fazer o seguimento semi-automático dos membros. É também apresentado os resultados relativos ao método escolhido.

No capítulo 4 é feita uma descrição detalhada da interface desenvolvida para analisar os movimentos.

No capítulo 5 são feitas as conclusões finais da dissertação.



## Capítulo 2

# Revisão Bibliográfica

### 2.1 Epilepsia

Não foi assim há tanto tempo atrás que a Epilepsia não era considerada uma doença do cérebro. Era acreditado ser um estado induzido pelos deuses e que provocava crises severas de inconsciência nas pessoas. O nome Epilepsia é remanescente do verbo grego *epilambaneim*, que tem como significado predisposição para ter crises. Esta é considerada uma das doenças conhecidas mais antigas da humanidade e é estimado que 1% da população mundial sofra desta doença. Apesar de no mundo moderno ser considerado já uma doença, ainda existem certas culturas que creem outras origens da doença e tentam ao máximo esconder os sintomas da mesma [13]. A epilepsia é considerada uma doença cerebral que é caracterizada pela predisposição de ter crises imprevisíveis do normal funcionamento do cérebro, no qual resultam crises recorrentes. Estas crises não pressupõem uma causa identificável. Quando uma crise sucede, existe um distúrbio nas células nervosas que controlam o movimento. Durante a crise, a pessoa apresenta perda de consciência e apresenta um comportamento anormal. Os sintomas variam consoante a pessoa e o tipo de crise associada. Enquanto algumas podem simplesmente só fixar no vazio, outras apresentam movimentos repetidos num ou mais membros. Na maioria dos casos a origem de epilepsia é desconhecida. Nos países menos desenvolvidos, a epilepsia provém essencialmente de infeções provocadas. Relativamente a países economicamente mais desenvolvidos, causas vasculares dominam na origem de epilepsia. Outros fatores não menos importantes segue-se trauma, neoplástica, neuro degenerativo, fatores metabólicos ou tóxicos, esclerose, más formações, inflamações, entre outros fatores externos. Quando a epilepsia é de causa congénita a origem poderá ser genética, idiopática, gene único, cromossoma ou insultos pré-natais [14], [1].

#### 2.1.1 Tipos de Crise

Uma crise é um evento clínico, com manifestações clínicas objetivas e subjetivas. Assim, manifestações puramente eletrográficas não podem ser consideradas como crises. Existem duas categorias primordiais, caracterizados pela Liga Contra Epilepsia (League Against Epilepsy – ILAE), que são crises parciais ou focais e generalizadas. As crises parciais caracterizam-se por

iniciar numa área específica do cérebro e podem ser subdivididas em parciais simples, em que não existe qualquer alteração na consciência e crise complexa parcial, em que há comprometimento da mesma. As crises generalizadas podem ser repartidas como parciais com generalização secundária se uma descrição clínica de um sintoma antecedente existir (aura), ou um eletroencefalográfico claro em que existe uma assinatura marcada. Crises individuais generalizadas incluem ausência, mioclónica, tónico-clónica, atónicos, tónicos e sintomas clónicos. Os tipos de crises podem coexistir, mas quando é diagnosticado um tipo, este é assumido como o predominante. Apesar dos especialistas tentarem separar as crises por tipos, o diagnóstico não é linear e existe alguma dificuldade em atribuir um nome aos sintomas observados, pelo que podem permanecer inclassificáveis [14], [1]. Numa visão mais detalhada, a crise do tipo generalizada tónica-clónica (convulsão) não é vulgarmente precedente de um aviso, pois existe uma perda de consciência. O paciente fica rígido e, se em pé, cai. Nesta etapa existe um espasmo nos músculos da mandíbula o que leva ao paciente a morder a língua. Visto que existe uma contração nos músculos e nos membros, os músculos respiratórios são também afetados e o ar pode ser forçado a sair dos pulmões, o que provoca o paciente a lacrimejar. A fase tónica geralmente demora 20 segundos, seguindo uma fase de movimentos descoordenados nos membros, a fase clónica. A descoordenação pode aumentar em amplitude, enquanto a frequência diminui. Nesta fase também poderá criar salivação e a respiração poderá ser estertorante, assim como incontinência. Esta fase pode demorar entre 2 a 3 minutos, seguida de uma fase pós-ictal, em que o paciente recupera a consciência e queixa-se de dor de cabeça [14], [1]. O tipo de crise em que há ausência ocorre essencialmente na infância ou, menos frequente, em adolescentes. Esta crise envolve primeiramente uma reação em que a criança fixa no vazio. Caso estejam numa conversa, deixam de responder e podem continuar a caminhar. Ocorre geralmente um automatismo rápido, que ocupa poucos segundos. Se os automatismos forem prolongados pode sugerir que seja uma crise do tipo complexa parcial. Estas crises não são recordadas pelo paciente e podem ser confundidas como devaneios. A crise mioclónica caracteriza-se por um breve tremor descoordenado do músculo e pode ocorrer numa parte do corpo, como no membro ou na cabeça ou então em diversas regiões. Estes movimentos podem ser repetidos e subtis. Crises tónicas apresentam uma contração abrupta nos membros e na musculatura axial, o que normalmente causa o levantamento dos braços e extensão das pernas. Estes tipos de crise apenas duram 30 segundos e seguem de uma recuperação breve. Este tipo de crise pode provocar queda. Crises atónicas são caracterizadas pela abrupta perda do tônus postural, o que pode levar a lesões na cara ou dentes, visto que o paciente pode eventualmente cair. Crises clónicas envolvem movimentos rítmicos generalizados. Estes movimentos são mais frequentes em neonatais infantes. Em crises parciais simples, os pacientes tendem a experienciar sintomas visuais como luzes coloridas fortes ou manchas e perda de visão, o que pode ser relacionado com alucinações. Crises do lobo parietal podem gerar um distúrbio sensorial, pelo que pode ser expresso em ardor na pele, parestesia, assim como pode ser expressa como dor ou vertigens. Se as crises tiverem origem na vertente motora estas irão afetar os movimentos e a atividade motora. Quando é originalmente da área motora o paciente pode adotar um postura de ataque, chamado de “postura de esgrima”. Crises do lobo frontal podem também manifestar-se em crises motoras

como uma postura frenética em que podem iniciar movimentos de pontapear. É comum ocorrer crises frontais durante a noite que são breves e repetidos. Crises parciais complexas têm origem no lobo temporal. Os sintomas descritos envolvem perda de consciência. Contudo, as primeiras manifestações são imobilidade, seguindo de automatismos. Estes automatismos podem envolver a boca, com movimentos de deglutição. Os pacientes apresentam confusão após o período pós-ictal [14], [1]. Existe uma maior dificuldade em distinguir as crises secundárias generalizadas das crises focais, tónicas-clónicas, quando é generalizada desde o início. Torna-se especialmente desafiador quando a crise ocorre durante a noite e não existem ainda sintomas experienciados. As principais crises estão sumariadas na tabela 2.1.

### 2.1.2 Diagnóstico

Diagnosticar de forma errada em epilepsia é comum (20-30%). Geralmente crises não epiléticas psicogénicas são as que são pior diagnosticadas como epilepsia. Depois desta, segue-se síncope como a segunda pior diagnosticada. Outras patologias como hipoglicemia, ataques de pânico, movimentos descoordenados paroxismal, desordens a nível do sono paroxismal são também facilmente confundidas com epilepsia. TIAs, enxaquecas a TGA são as patologias pior diagnosticadas, mas menos frequentes. Estas más interpretações que levam ao mau diagnóstico ocorrem especialmente quando as avaliações de EEG são sobre interpretadas.

Vídeo-EEG melhorou imenso na precisão de diagnóstico e diversas investigações posteriores apresentam resultados promissores para a diminuição de maus diagnósticos.

O diagnóstico de epilepsia envolve a classificação em diferentes categorias. Estas categorias podem ser classificadas por manifestações clínicas e pelas características da EEG da crise (generalizadas ou focais), por causa (idiopática, criptogénica e epilepsias sintomáticas), por síndrome (incorporando observações clínicas, etiológico e EEG, entre outros fatores de diagnóstico) ou por fatores de alarme.

Este método de classificação foi aprovado pela ILAE em 1981 e foi revisto em 1989 [15].

### 2.1.3 Vídeo-EEG

O Vídeo-EEG consiste numa das abordagens de diagnóstico de Epilepsia. Esta abordagem consiste na gravação em sincronia de duas componentes, o vídeo onde é possível verificar que movimentos ocorreram durante a crise e o EEG onde ligam elétrodos ao couro cabeludo em locais específicos e permite a aquisição dos sinais elétricos do cérebro. Esta abordagem ocorre na Unidade de Monitorização de Epilepsia e pode levar o internamento do paciente entre 3 a 5 dias.

O diagnóstico da epilepsia pode trazer ambiguidade em alguns casos, pelo que é necessário uma análise mais detalhada do cenário aquando a avaliação do paciente. O Vídeo-EEG trouxe uma enorme vantagem e surgiu da necessidade de classificar e confirmar o tipo de crise. Esta metodologia é crucial quando os pacientes são considerados para resseção cortical focal. É também usado para confirmação da presença de episódios clínicos não epiléticos. Além disso, esta técnica é também usada para aceder a frequência das crises e para fatores precipitantes.

<b>Terminology</b>	
Aura	Subjective experience as a result of seizure discharge. It can occur alone (a simple partial seizure) or be followed by other seizure manifestations. An aura is a seizure, not a prodrome to a seizure.
Automatism	Complex automatic behaviour associated with loss of awareness, such as lip smacking or hand wringing occurring as part of a seizure.
Convulsion*	Term used for a seizure of any type with involuntary, irregular myoclonic, clonic or tonic-clonic movements of one or more limbs. It is most applicable to the form of generalized seizure with tonic-clonic movements ( <i>grand mal</i> ), emphasizing the generalized 'convulsive' type, as opposed to generalized 'non-convulsive' (e.g. childhood absence) epilepsy.
Cortical malformations	These are also referred to as <i>cortical dysgenesis</i> , <i>cortical dysplasia</i> or <i>neuronal migration disorders</i> and are developmental abnormalities of the cerebral cortex.
Cryptogenic*	Cause unknown, presumed symptomatic.
Déjà vu	Literally 'already seen', this refers to a false impression that a present experience is familiar. It is used to refer to something heard, experienced or seen. It can be the aura of a temporal lobe seizure, although it also happens in other settings.
Epilepsy	Two or more seizures.
Epilepsy in remission	No seizures for 2-5 years.
Fit	A popular lay term to denote a seizure of any type but also an outburst of temper or aggression.
Ictus	Literally a 'blow'. Usually used for an epileptic seizure (but can refer to other paroxysmal events, such as migraine, transient neurological events and stroke).
Idiopathic (primary)*	In relation to epilepsy, this implies that there is an underlying genetic cause.
Incidente	The rate of an event in a population over a certain period of time.
Jamais vu	Literally 'never seen', this refers to a false impression that a present experience is unfamiliar. It is used for something seen, heard or experienced. It can be the aura of a temporal lobe seizure, although it also happens in other settings.
Learning disability	Difficulties with daily living associated with intelligence quota less than 70 and age of onset less than 18 years.
Major seizure	A loose term that refers to a generalized convulsive seizure.
Minor seizure	A loose term that refers to seizures that are not generalized convulsive.
Partial (focal, localization related) seizure	A seizure originating in a specific cortical location.
People with epilepsy (PWE)	Term used instead of 'epileptics' to emphasize the whole person and that epilepsy is a condition someone has, and not their identity.
Postictal	The period of time after a seizure when behaviour or mood can be directly affected by the seizure.
Prevalence	The rate of an event in a population at a single point in time.
Primary generalized seizure	A seizure involving both hemispheres from onset.
Refractory epilepsy	Not responding to appropriate medical treatment over at least 2 years.
Seizure	Clinical manifestation of a sudden abnormal cerebral electrical discharge.
Standardized mortality ratio	The ratio of the observed deaths in a study population to the expected number experienced by the age- and sex-matched control group.
Sudden unexplained death in epilepsy (SUDEP)	Sudden, unexpected death in a person with epilepsy not due to drowning or other accident; without or with a seizure. Deaths during status epilepticus or those with an anatomical or toxicological explanation are excluded.
Symptomatic epilepsy	Having an identified underlying cause.
Syndrome	A constellation of clinical and investigational features for which there may be a number of aetiologies.
Typical absence (petit mal)	A generalized seizure characterized by loss of awareness, usually brief and associated with characteristic 3-4/second spike and wave discharges on an electroencephalogram. This must be distinguished from atypical absence and complex partial seizure.

\*These terms have acquired a specific meaning in epilepsy different to their use in other fields of medicine.

Figura 2.1: Tipos de crises e respetivas características. Adaptado de [1]

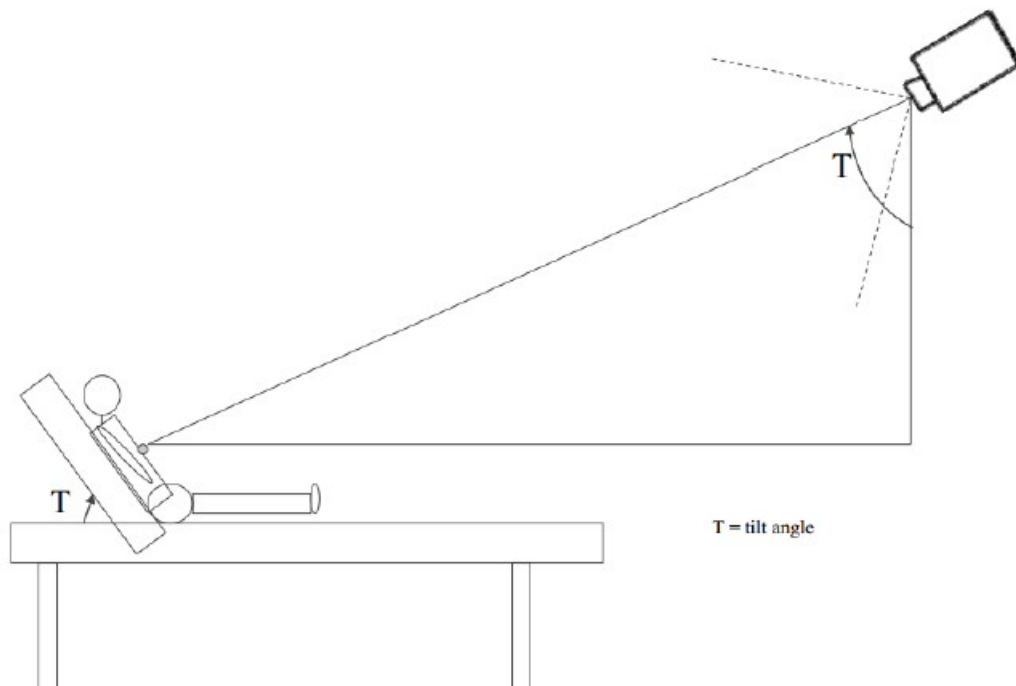


Figura 2.2: Esquema ilustrativo da Unidade de Monitorização. Adaptado de [2]

Esta metodologia permite uma análise qualitativa através da observação do vídeo, assim como uma análise mais detalhada e quantitativa através dos padrões relativos ao sinal elétrico (EEG) onde permite a observação dos padrões de movimento. Ao analisar o padrão de movimento durante as crises, é possível ao especialista fazer um diagnóstico mais preciso. O padrão de movimento poderá revelar a crise nesse preciso momento, assim como confirmar se existiu realmente uma crise ou o tipo inerente. O Vídeo-EEG é montado numa unidade especial para fazer o acompanhamento do paciente com epilepsia ou quando em dúvida se epilepsia [16]. É possível ver um desenho ilustrativo na figura 2.2 de como geralmente estas unidades estão montadas para poder monitorizar o paciente.

As desvantagens desta técnica vincula o custo inerente do estudo e da hospitalização que requer recursos especiais e pessoal destacado [17].

#### 2.1.4 Mortalidade

A epilepsia tem uma taxa de mortalidade bastante alta, entre 2 a 3 vezes mais alta que na população geral. A maioria dos óbitos ocorre uns anos depois do diagnóstico. Esta taxa também inclui as taxas de suicídio, que também é considerado mais elevado do que na população geral. Algumas das mortes são aquelas que ocorrem a pessoas que apresentavam ser saudáveis e ficaram epiléticos de forma inesperada, levando a acreditar que as mortes destes ocorrem durante crises. Os casos que são monitorizados nas unidades respetivas sugerem que a morte resulta de apneia [1].



### 2.1.5 Tratamento

Avaliar corretamente o tipo de crise é crucial para a administração do medicamento adequado. A terapia inicia com drogas anti-epiléticas (AEDs). Apesar dos AEDs serem recorridos mais frequentemente para o tratamento, os pacientes podem adotar um diferente estilo de vida para minimizar a ocorrência de crises. Relaxamento, bio feedback, privação de álcool e privação de sono podem mostrar resultados significativos em alguns casos, especialmente naqueles que contêm uma aura fidedigna que precede a crise complexa parcial ou uma secundária generalizada. Mudanças na dieta podem também ser significativas.

Apesar da medicação e das mudanças de estilo de vida poderem ser bastante eficazes, em alguns casos é necessária intervenção cirúrgica (20-30%). Este tratamento é indicado se as crises surgirem de uma área em que a remoção não causa deficiências neurológicas inaceitáveis.

Os candidatos para a cirurgia são monitorizados via Vídeo-EEG, neuro imagem e estudos neuro psicométricos. Esta cirurgia envolve uma lobectomia temporal, que é o procedimento mais vulgar em epilepsia. O putativo benéfico da cirurgia é para uma melhor qualidade de vida e para existir participação do paciente como membro da sociedade [18].

## 2.2 Técnicas de Análise Motora

O movimento humano tem sido ao longo dos anos bastante investigado. Esta investigação massiva deve-se essencialmente à procura do jogo mais realista, em que o jogador participa ativamente e de forma real num sistema virtual. A percepção e o comportamento motor têm, portanto, de ser consideravelmente conhecidos e reconhecidos para este objetivo ser conseguido.

Devido à exigência competitiva no mundo dos jogos, foi possível ao longo dos anos quantificar o movimento humano e este ser explorado noutros campos para além do mundo virtual. Existe então um grande compromisso para encontrar a melhor técnica de análise da cinemática humana, a qual terá de ter como base uma estrutura humanoide A priori comprimida por um conjunto de juntas. Este tópico tem sido alvo de diversas investigações, desde performance atlético, assim como um alicerce para o diagnóstico médico. O foco consiste na segmentação de partes do corpo na imagem, seguimento do movimento das juntas sobre a sequência de imagem e a recuperação tridimensional do subjacente da estrutura corporal. Seguidamente da aplicação destas técnicas de visão de computador, é possível adquirir informação relevante para uma análise detalhada do movimento. Para isso ser possível é necessário primeiramente uma segmentação do corpo, deteção e identificação de juntas, e recuperação da estrutura 3D das projeções em 2D na sequência de imagem. O movimento do corpo humano é vulgarmente associado ao movimento de membros e mãos, como a velocidade da mão ou dos segmentos dos membros, ou o ângulo com as diferentes partes do corpo. A análise do movimento baseia-se em duas abordagens: modelos que não têm uma forma associada e outro modelo em que tem uma forma associada. A primeira abordagem é feita através da estimação de juntas dos segmentos do corpo [3]. O modelo de varetas é um dos

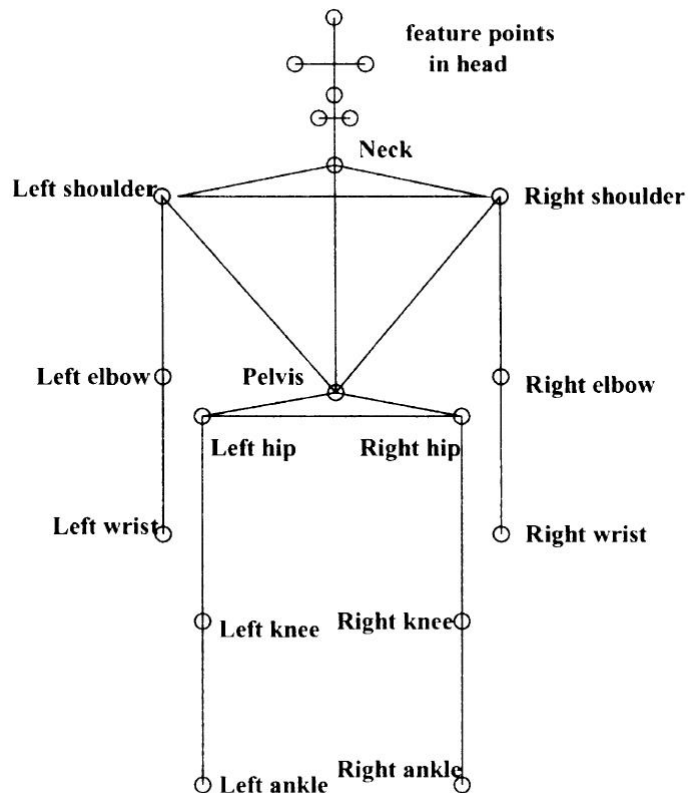


Figura 2.3: Modelo Vareta. Adaptado de [3]

exemplos desta primeira abordagem, que consiste em segmentos de linhas ligados pelas juntas, que pode ser visto na figura 2.3.

O movimento das juntas fornece a chave para a estimação do movimento e reconhecimento da figura completa. Outro método é usar contornos em 2D. Quando este método é usado, os investigadores usam técnicas de fluxo ótico para calcular a velocidade [19]. A maior dificuldade para este método é a segmentação do corpo. O segundo método, que requer um modelo *A priori* com forma é mais desafiante, pois é necessário corresponder as características do modelo às *frames* sucessivas. É geralmente usado já segmentos e juntas previamente definidas para facilitar a correspondência de características. É representado no modelo a cabeça, o torso, ancas, braços e pernas. A cabeça é usada para inicializar o seguimento e as restantes juntas são com base em estimativas. O maior inconveniente em modelos 2D é conseguir visualizar em diferentes ângulos da câmara. Este problema pode ser resolvido pela aplicação de um modelo volumétrico em 3D, mas apresenta como inconveniente ser computacionalmente oneroso durante o processo de correspondência, pois requer alguns parâmetros [3].

Alguns estudos de análise de movimento foram aplicados a patologias que têm desordens a nível motor, pelo que pode ajudar na quantificação do movimento e a reconhecer padrões associados a essas mesmas patologias. Estudos que envolvem a caracterização automática têm de incorporar um reconhecimento da atividade humana. Para esse fim, foram aplicadas características geométricas,

como aglomerados em 2D, projeções de perfil, textura e informação da cor. Este processo envolve uma aquisição de padrões dessas mesmas características que posteriormente pode ser reconhecido quando comparado a esses padrões previamente guardados. Para adquirir o padrão é normalmente usado um modelo correspondente, que é de baixo custo computacional, mas como inconveniente pode ser demasiado sensível à variância do tempo de movimento. Outro método para adquirir padrões é definido como estado-espço. Este método consiste em definir uma determinada postura corporal como um estado. As diferentes posições são então consideradas diferentes estados e a mudança de posição é considerada como um circuito de transição de estado estático para outro estado estático. Este método não é linear e tem um custo exigente computacional, pois requer um processo de treino [20].

Para uma boa análise de movimento é primordial que exista um bom seguimento do movimento do corpo humano. O seguimento do corpo tem como base estabelecer pontos do corpo previamente definidos, juntas, e seguir essas juntas ao longo do tempo e ao longo da sequência de imagens. O seguimento exige uma extração de características para fazer a correspondência às juntas. As características poderão ser a posição relativa, a velocidade, forma, textura e cor. O seguimento é geralmente um processo que envolve corresponder as imagens em que usa os pixels, pontos, linhas, *blobs*, que tem por base a sua forma e outras informações visuais. A maioria dos métodos usa apenas uma câmara para fazer o seguimento. Para melhorar o seguimento é então sugerido usar mais câmeras, pelo que aumenta a área monitorizada. Esta solução é um reforço para a limitação que surgia quando algo ficava obstruído num certo ângulo de visão. A maior dificuldade é fazer a correspondência de características, pois as características são gravadas de diferentes coordenadas. Para solucionar este problema terá de se usar o mesmo referencial para fazer uma correspondência precisa [3]. É possível reconhecer os inúmeros parâmetros que uma análise motora requer, os quais estão descritos no gráfico da figura 2.4.

Nos últimos anos o maior foco na análise de movimento tem sido desenvolver outros métodos de seguir numa perspetiva tridimensional. As regalias de uma análise tridimensional recaem essencialmente na diminuição de sensibilidade de variação de tamanho, relativamente às coordenadas e ângulos das juntas e podem fornecer informação mais proveitosa acerca do movimento. Para executar uma correspondência dos segmentos do corpo, os investigadores não atribuíram uma forma perfeitamente modelada ao corpo humano, mas como alternativa, uma aproximação à forma humana. Para isso, formas elementares foram usadas como esferas no *Bubbleman* [21], [22].

## 2.3 Captura de Movimento

Devido ao largo foco que tem havido para a captura do movimento em que vai desde jogos ate análise biomédica, o interesse em sistemas sem marcadores tem aumentado na indústria. Existiu desde então um melhoramento dos sensores RGB-D e das câmeras RGB, o que levou a uma diminuição de custos para ser praticável nas diversas aplicações. Depois do lançamento destes sensores o objetivo tem sido incorpora-los em sistemas computacionais [4]. Pela análise da literatura, a captura do movimento pode ser agrupada em quatro categorias, que são mostradas na



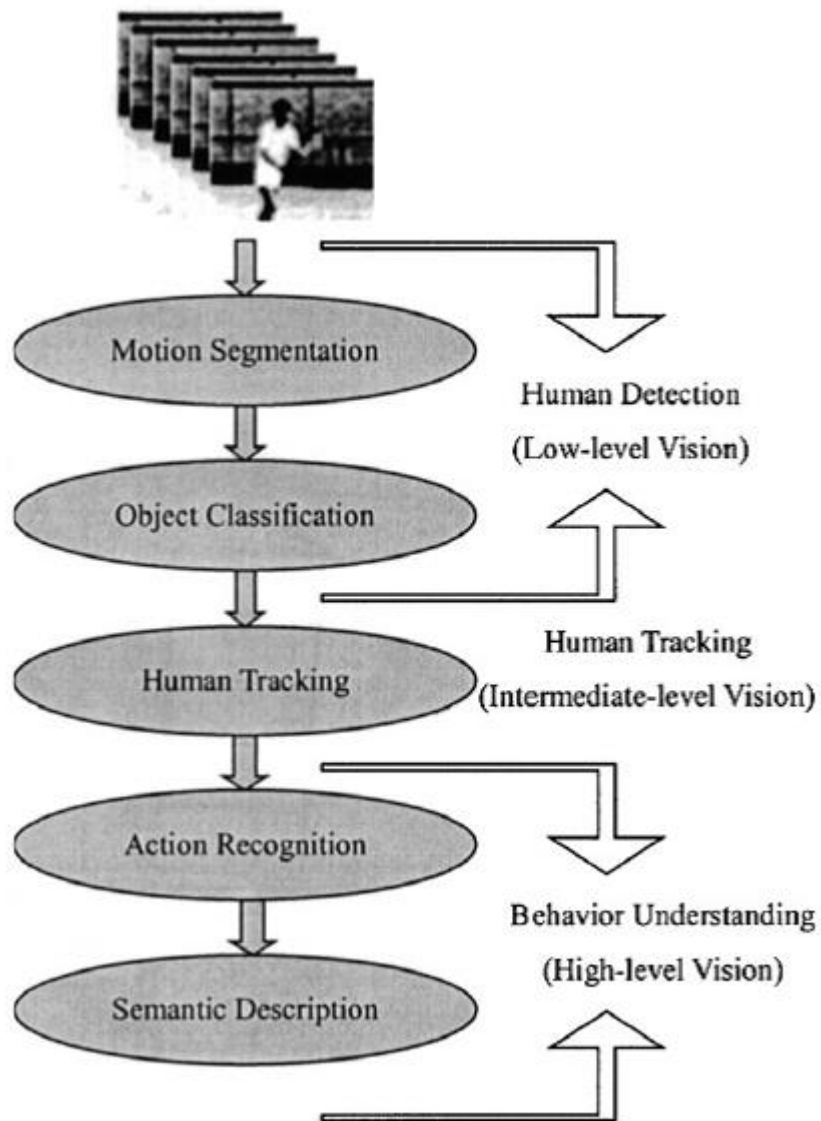


Figura 2.4: Esquema de análise motora. Adaptado de [3]

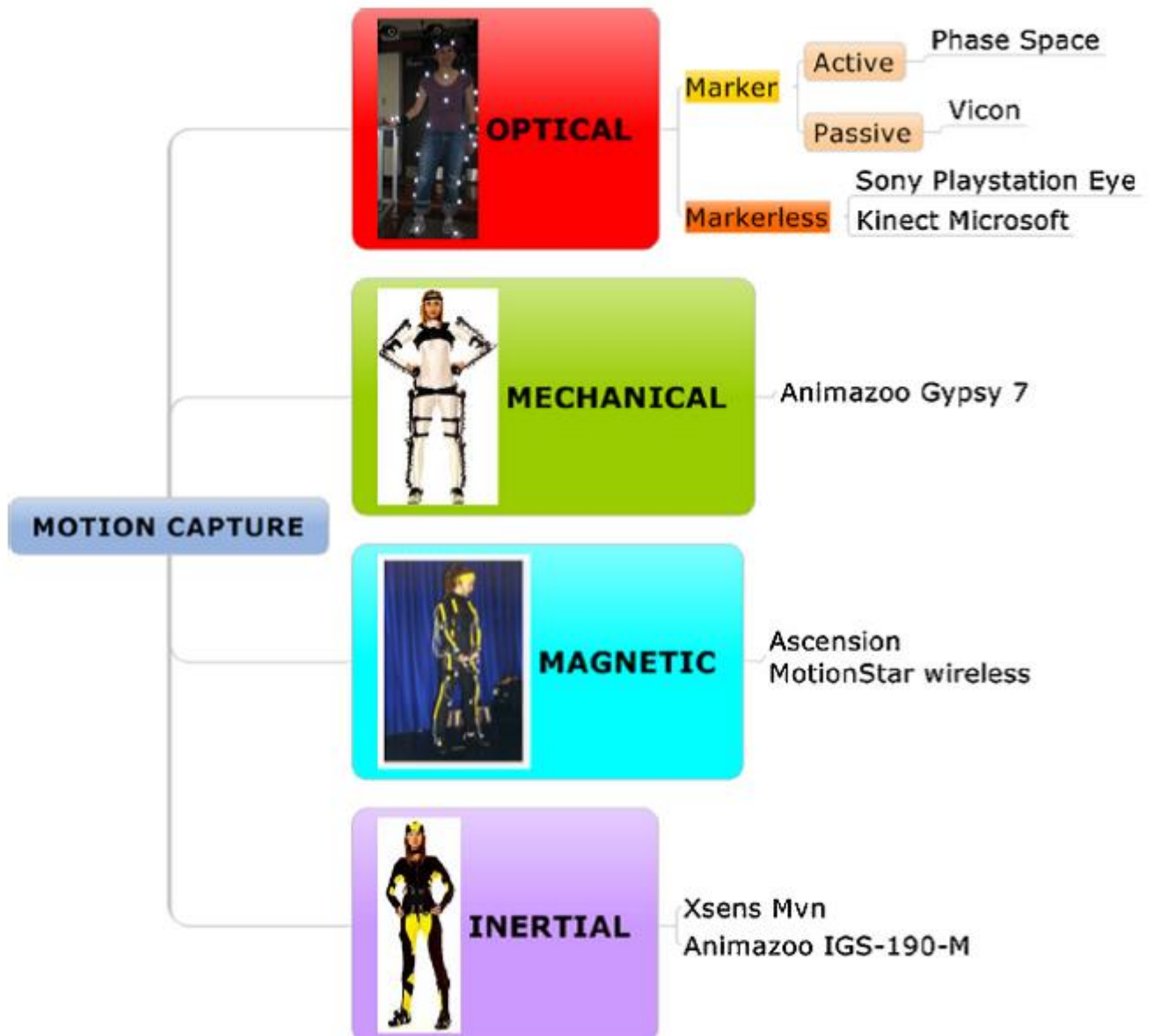


Figura 2.5: Esquemas de diversos tipos de técnicas de captura de movimento. Adaptado de [4]

figura 2.5 e descritas posteriormente. Sistemas óticos conseguem extrapolar a posição das juntas através da triangulação de imagens retiradas de diferentes câmeras, ou podem explorar os dados que são fornecidos da câmera de profundidade. Este método pode ser conseguido através de marcadores ou na ausência de marcadores. Pelo primeiro método, os marcadores são conectados ao corpo em pontos estratégicos. O segundo, é baseado em visão por computador para determinar a posição desses pontos [4].

O sistema mecânico é conseguido através de um exosqueleto que é vestido pela pessoa e consegue medir os movimentos executados. O exosqueleto é composto por segmentos conectados por goniómetros de metal ou plástico e potenciómetros de juntas. Os movimentos são adquiridos pelas

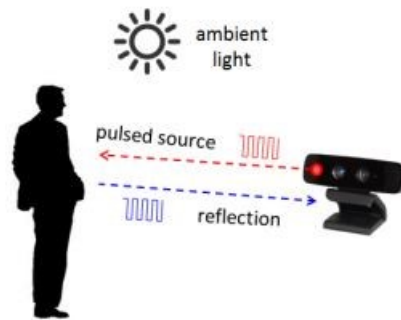


Figura 2.6: Conceito de *Time-of-Flight*. Adaptado de [5]

juntas e convertidos em sinais elétricos. Estes dados podem ser transferidos para o computador via *wireless* ou por cabos. Esta tecnologia é de baixo custo e apropriada para movimento complexo. A desvantagem deste sistema é o distúrbio que provoca ao normal comportamento do movimento na pessoa [4]. Sensores inerciais usam acelerómetros e giroscópios para adquirir movimentos. Esta técnica tem uma fácil portabilidade, mas os sensores podem sair do ponto fixado e afetar as leituras [4]. Os sensores magnéticos funcionam com base no fluxo magnético relativo de três bobinas ortogonais em que tanto os transmissores como os emissores são usados para determinar a posição e a orientação. Este sensor traz algumas desvantagens, como não ter uma resposta linear, especialmente entre arestas da área capturada, o que o condiciona, se comparado a outras tecnologias. Outra desvantagem é a rede de cabos dos sensores que tende a impossibilitar a execução dos movimentos, assim como materiais ferromagnéticos não poderem ser usados em cena [4]. Sistemas sem marcadores óticos estão a ser cada vez mais utilizados, parcialmente devido ao facto de não serem invasivos para a pessoa. Sensores RGB-D têm uma enorme vantagem pois apresentam um baixo custo, comparativamente a outros sistemas, conseguem fornecer excelentes resultados face a outras tecnologias mais dispendiosas, como o Vicon [4].

## 2.4 Sensores RGB-D

Os sensores RGB-D são geralmente compostos por dois grupos, dependendo da tecnologia usada para a medição de profundidade. Podem usar a tecnologia *time-of-flight* (ToF) ou *structured light*. Uma câmara 3D ToF funciona iluminando a cena com uma fonte de luz modulada, e observando a luz refletida. A distância é calculada pelo tempo despendido entre a iluminação e a reflexão [5]. A figura 2.6 mostra esse conceito.

A *structured-light* funciona através da projeção de padrões conhecidos para o sujeito e inspecionando a distorção desse padrão. Projeções sucessivas de padrões codificados ou de mudanças de fases são requeridos para extrair uma única *frame*, o que leva a uma menor taxa de *frames* por segundo. Isto implica que o sujeito terá de permanecer estático para que não ocorra manchas na imagem. O padrão refletido é sensível a interferências óticas. A *structured-light* funciona

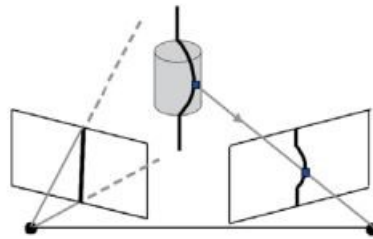


Figura 2.7: Conceito de *Structured – Light*. Adaptado de [5]

melhor como uma aplicação no interior. A figura 2.7 mostra exatamente este conceito. Por comparação, ToF é menos sensível a alinhamento mecânico e a luz de ambiente e é mecanicamente mais compacto. A atual tecnologia ToF tem uma menor resolução do que a *structured light*, mas é mais rapidamente melhorável. Estas câmeras têm a vantagem de diminuir os algoritmos para a aquisição de visão tridimensional. A tabela 2.8 compara os sistemas existentes.

#### 2.4.1 Sensores *Time-of-Flight*

Na literatura, devido ao interesse da questão, existe inúmeras tentativas de comparação deste tipo de sensores de baixo custo com outras tecnologias existentes. Em [6] são demonstradas as vantagens, limitações e a forma como estes sensores estão a ser aproveitados nas diversas aplicações. A profundidade é medida através deste sensor, pelo que o seu princípio é o “tempo-de-voo”. Pode ser medida através de dois meios, através de um sinal pulsado ou através de uma onda contínua modulada. Sem entrar em muito detalhe na tecnologia usada para conseguir medir o tempo de voo, este tipo de sensor consegue fornecer a quantidade de luz refletida pelo ambiente para cada pixel. Relativamente a outros sensores de profundidade, este apresenta propriedades interessantes, como o registo de profundidade assim como a intensidade das imagens, aquisição completa da imagem com uma taxa de *frames* alta, apresentam dimensões reduzidas, de baixo peso, um design compacto, não necessitam de partes móveis à parte e têm auto iluminação. Apesar das suas diversas vantagens, este sensor apresenta algumas limitações, nomeadamente erros sistemáticos e não sistemáticos. Contudo, estes problemas podem ser resolvidos através de calibração e por filtragem, respetivamente. Este tipo de sensores fornecem informação do mapa de profundidade sem comprometer a taxa de *frames*. Para além disso, a segmentação da informação de profundidade é bastante acessível, pelo que a maioria das aplicações requerem este tipo de sistema. Sensores ToF são também facilmente incorporados em aplicações de ambiente humano, pois são seguros e permitem evitar contacto físico e marcadores especiais ou hardware. Apesar de, atualmente, as tecnologias de *stereo vision* e *structured light* liderarem devido à sua melhor precisão e aquisição, a utilização de tecnologias ToF estão diariamente a aumentar devido às suas características distintas. Estes sensores demonstram uma facilidade em extrair dados rapidamente, em que o consumo de energia é baixo assim como o tempo de computação. Demonstram também um grande potencial em extrair características, reconstruir, planear e otimizar algoritmos. Estas são mais adequadas

Differences	ToF cameras	Stereo vision	Structured light
Correspondence problem	No	Yes	Yes
Extrinsic calibration	No, when used alone	Yes	Yes
Auto illumination	Yes	No	Yes
Untextured surfaces	Good performance	Bad performance	Good performance
Depth range	0.3 ÷ 7.5 m.	Base-line dependent	Light-power dependent
Image resolution	Up to 204x204	High resolution. Camera dependent	
Frame rate	Up to 25 fps.	Typically 25 fps. Camera dependent	

Figura 2.8: Diferentes tecnologias de medição de profundidade. Adaptado de [6]

para aplicações em tempo real e, em particular, para aquisição automática em 3D. Exemplos de sensores ToF RGB-D são as câmaras *DepthSense* (DS311 e DS325) [23], a *Panasonic D-Imager* [24] e, recentemente lançada, *Kinect for Windows* versão 2 (K4W) sensor [7]. A câmara DS325 é o sensor de profundidade mais preciso do mercado. É usado em aplicações de interação próxima (seguimento da mão e dos dedos). Tem uma resolução RGB mais alta quando comparada com a DS311 [23]. Outro sensor 3D ToF é *Panasonic D-Imager*, lançada pela *Panasonic Electric Works*. A *D-Imager* usa a propriedade de dispositivo de carga acoplada (CCD) juntamente com os LEDs infravermelhos próximos para conseguir a percepção dos gestos humanos e conseguir seguir todo o movimento do corpo, o que permite aos utilizadores experiência interativa. Este sensor possibilita captura precisa de movimento dos objetos com um campo de visão alto ao processar a informação com alta precisão (pixel a pixel). Funciona com uma taxa aproximadamente de 15-30 fps com resolução de 160x120. A maior vantagem deste sensor é que tem uma captura entre 1.2m a 9m [24]. *Asus Xtion Pro Live* [25] é outra alternativa para seguimento de movimento. Este sensor é feito pela mesma companhia (*PrimeSense*) como a *Kinect* e tem um *hardware* semelhante. As câmaras e os sensores quando perto seguem o movimento da pessoa de frente, extrapolado e até mesmo reconhecem pontos-chave, tudo em tempo real. Uma desvantagem do *Asus* é que devido à natureza do *hardware*, necessita de estar posicionado idealmente no meio do comprimento de exibição quando visto horizontalmente. Uma das diferenças destes dois dispositivos é a qualidade da câmara RGB, que é melhor na *Asus Xtion Pro Live* [25].

#### 2.4.2 Microsoft Kinect

A *Kinect v2* da Windows é um dispositivo com um sensor de profundidade, uma câmara de cor, um emissor infravermelho (IR), um *array* de microfones, pelo que permite adquirir a localização e os movimentos da pessoa assim como reconhecer as suas vozes. Este sensor também fornece um software num kit de desenvolvimento (SDK 2.0) e fornece aos utilizadores drivers,



Figura 2.9: Microsoft Kinect KW4. Adaptado de [7]

ferramentas, APIs, interfaces e amostras de código para facilitar o desenvolvimento da *Kinect* – habilitar aplicações de desenvolvimento comercial. Este SDK, comparado com os previamente lançados possibilitam seguir até 6 pessoas e apresenta 25 juntas contra as 20 do sensor anterior (v1), figura 2.10.

As características técnicas podem ser encontradas na tabela 2.11.

## 2.5 Quantificação de Movimento em Epilepsia

A epilepsia era considerada pelos anciões como uma doença entregue pelos deuses e era reportada como um demónio preso dentro da pessoa, era, então, tratado como um assunto religioso. Não foi assim há tanto tempo atrás que a epilepsia foi vista como uma doença do cérebro e começou a ser investigada por volta do século XIX. Atualmente é possível ter uma análise mais detalhada sobre a doença e investigar as suas possíveis causas, assim como avaliar as crises em tempo real. Durante vários anos, investigadores e profissionais de saúde tentaram encontrar a melhor maneira para aceder às crises epiléticas sem provocar desconforto nem interferir na quantificação dos movimentos para classificar as crises. Em 1875, Richard Caton elaborou um método de medir o campo elétrico do cérebro em tempo real [26]. No seu método ele foi capaz de gravar sinais do campo elétrico através de um eletroencefalógrafo. Apesar de ele ter usado macacos, foi inspiração para ser recriado em humanos. Levou alguns anos, pelo que foi apenas em 1920 que o psiquiatra Hans Berger, que criou o eletroencefalógrafo (EEG) que posteriormente foi usado como

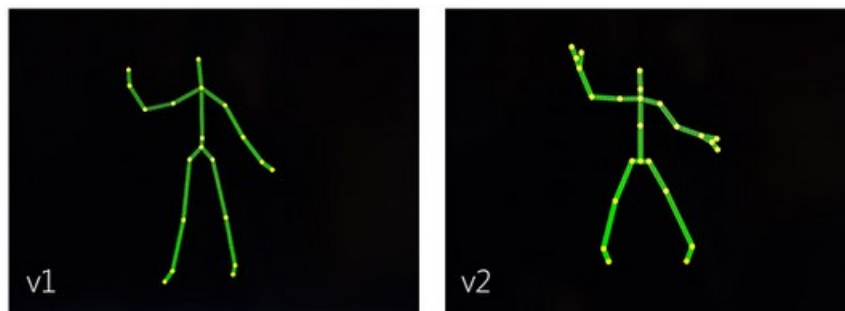


Figura 2.10: Juntas relativas ao sensor v1 e v2. Adaptado de [7]



Requer fonte de energia	Sim
Sensores	RGB e Depth Sensors, Microfone com tecnologia Array
Resolução	RGB: 1920 x 1080; Depth, Infrared and Body Index: 512 x 424 pixels @ 30 Hz
Tipo de plataforma	4GB de Memória (ou mais), i7 2.5Ghz (ou maior)
Interface	Built-in USB 3.0 host controller
Software	Microsoft Kinect SDK v2
Linguagens de programação	C#, C++, Visual Basic, Java, JavaScript
Dimensões	24.9cm x 6.6cm x 6.7cm

Figura 2.11: Características da Kinect v2, K4W

ferramenta para diagnóstico. É possível extrair e reconhecer padrões na atividade do cérebro através do EEG, que é uma representação das mudanças eletromagnéticas ao longo do tempo e são capturadas pelos sensores de EEG. Alguns dos padrões podem ser relacionados com epilepsia, as descargas de forma epilética, que podem ser durante a crise (ictal) ou entre as crises (interictal) [27]. Em contribuição à monitorização de EEG, em 2002, Gregory D. Cascino [17], melhorou a avaliação desta patologia ao introduzir a monitorização em Vídeo-EEG. Relativamente a esta técnica, sempre que em dúvida no diagnóstico de epilepsia, esta é crucial na distinção de evento paroxismal epilético ou não epilético. Contudo, como a epilepsia é associada à descoordenação de movimento, o tradicional Vídeo-EEG apresenta algumas limitações na inspeção visual, pois não fornece resultados objetivos e não pode por si só, analisar quantitativamente. Tendo em consideração essas limitações, investigadores promoveram outros métodos através de deteção por visão, análise e reconhecimento em epilepsia, em que conseguem aceder a movimentos de epilepsia por dois métodos distintos: através de marcadores ligados ao corpo ou através de métodos em que não seja necessário ligar marcadores ao corpo. O acesso ao movimento é baseado em quantificação, obtendo os seguintes parâmetros: amplitude, frequência, velocidade, ângulo, entre outros [27]. Até agora, os investigadores dividiram a avaliação de crises epiléticas em três grupos distintos: deteção de movimento, análise de movimento ou reconhecimento de movimento. O primeiro grupo atua como uma forma inteligente de monitorizar os pacientes e criar sistemas de alerta se detetar uma crise. Pode ser também usado para descartar secções de vídeo irrelevantes para a análise. O segundo grupo é focado na extração de informação objetiva que pode ser usado para diagnóstico e uma melhor proposta de tratamento. Por último, o terceiro grupo pode fornecer o tipo de crise ocorrida e a distinção entre eventos epiléticos e não epiléticos. Com o intuito de realizar uma análise motora mais detalhada, são usadas técnicas em que são aplicados marcadores ao corpo e outras que apenas são por visão computacional, como foi referido anteriormente. Através dos marcadores, estes são aplicados no corpo e seguidamente são usadas câmeras para poderem seguir os marcadores colocados. Estes marcadores são aplicados em posições estratégicas em que apresentam relevo para o estudo em questão, como juntas ou extremidades do corpo. É frequente ser avaliados componentes em 2D, em que pode ser fornecido as trajetórias, a amplitude relativa à

primeira *frame* e a distância de marcadores selecionados relativamente a outro marcador [27].

Em [28] foi usado, como marcadores, sensores refletivos infravermelhos. Estes marcadores foram aplicados em diferentes partes do corpo e seguidamente foram analisados e seguidos numa sequência de imagem com 25 *frames* por segundo durante a crise ocorrida. As ferramentas usadas para esse fim foi o MATLAB 7.4 e o MaxTraQ 2.1. Seguidamente à aquisição das coordenadas 2D, foi analisado a diferença dos automatismos dos membros superiores entre o lobo temporal e frontal. Foi possível concluir que os automatismos dos membros superiores apresentam uma maior extensão, cobrem maior distância e são mais rápidos no lobo frontal contra o lobo temporal. Neste estudo, a análise foi feita tendo em conta o comprimento do movimento, a extensão do movimento e a velocidade máxima e média.

Em [29] foi introduzido uma nova abordagem para quantificar o movimento em crises epiléticas. Foram usados *blobs* para aplicar no corpo e medir as respetivas características desse movimento, através de uma plataforma feita pelos autores, MovEpi3D, como a posição, distância, aceleração, volume e frequência. O cálculo do erro teve como base o cálculo da distância Euclidiana entre a posição do marcador em cada *frame* e a sua posição respetiva à primeira *frame*. O método 3D mostrou superioridade, especialmente para padrões de movimento perpendiculares ao plano 2D. Existem diversos tipos de crises e algumas das propostas recentes têm sido para distingui-las quantitativamente. A maioria dessas crises são caracterizadas por sintomas motores. Por conseguinte, existe uma possibilidade de aceder a essas crises e aprender um sistema capaz de reconhecer esses sintomas motores.

Em [30], foi analisada a trajetória do movimento em automatismos com origem no lobo temporal (TLE) e no lobo frontal (FLE), através de uma técnica baseada em aplicação de marcadores. Foram então aplicado 20 marcadores ao corpo dos pacientes e foi possível distinguir estes dois tipos de epilepsia (TLE de FLE). Este estudo foi baseado em [2], o qual foi o primeiro a analisar quantitativamente numa crise hipercinética. Diferentes semiologias das crises (hipercinética, postura tónica, postura de esgrima, giro de cabeça tónico) podem ser distinguidas pela média da amplitude, assim como taxas de amplitude entre os membros distais e proximais. Estas abordagens apenas podem fornecer informações sobre o movimento depois de uma análise *A posteriori*. Recentemente, um sensor foi criado para detetar crises enquanto o paciente se encontrava num ambiente clínico, pelo que pode monitorizar os movimentos do paciente através de um acelerómetro. Os sensores wireless são posicionados no corpo do paciente e sempre que um limite é atingido, o alarme é acionado através de um transmissor RF para ser possível notificar a pessoa encarregue. Este método foi bem sucedido pois apresentou uma taxa de precisão de 90% para detetar crises em tempo real [31].

Como uma alternativa de baixo custo, [32] usaram sensores magnéticos inerciais em que 3 foram colocados nos membros dos pacientes para obter movimentos em 3D. As crises foram gravadas e divididas em diferentes eventos, que correspondem a manifestações motoras e foram usadas como base para um processo de aprendizagem. As manifestações motoras em cada membro são chamadas de *motorogram* e são usadas para distinguir padrões em diferentes manifestações numa tentativa de ajudar os neuro fisiologistas a decidir o tratamento a aplicar. Os padrões foram cons-



truídos com base em redes neuronais artificiais. Este processo automático de classificação teve um erro de cerca de 20% entre a decisão do computador e a decisão do especialista que analisava as crises, demonstrando que existe potencial neste tipo de sistemas para melhorar o diagnóstico.

Em [33], apesar de intitulado de sem marcadores, o movimento foi calculado com base num pijama colorido vestido pelo paciente para ajudar no seguimento dos membros e recolher as características para a deteção e análise da crise. Um alarme é acionado quando uma crise é detetada, caso exista um movimento que exceda o limite definido. Apesar dos marcadores terem uma enorme importância na classificação e análise de crises em epilepsia, quando estes desaparecem da câmara trazem alguma dificuldade para estimar a sua posição, levando a incertezas.

Nas abordagens sem marcadores, a análise é feita através da automatização na aquisição do vídeo e no processamento de imagem. Para esta abordagem, o seguimento tem por base em marcar as juntas nas *frames* dos vídeos e seguir a junta correspondente. Apesar de não existirem marcadores físicos aplicados no corpo, pode ser assumido como uma abordagem sem marcadores, mas apenas quando as técnicas são completamente automáticas no processamento de análise dos vídeos é que podem ser consideradas técnicas sem marcadores. Recentemente, alguns autores apresentaram uma abordagem que não requer qualquer tipo de marcação nem alguém que marque as juntas em cada *frame* e apenas analisa automaticamente a imagem. Em pacientes pediátricos torna-se difícil fazer uma ligação com marcadores físicos, portanto em [34], a técnica utilizada consistiu em usar fluxo ótico pois os vetores de fluído conseguem indicar a força do movimento. É determinado um limite ótimo para distinguir épocas de movimentos de épocas de não-movimento, através duma validação cruzada de 3 elementos. Para a avaliação das crises, foi necessário um especialista marcar o início e o fim de cada época de movimento. Com base na intensidade de pixels, foi calculada a velocidade dos vetores em cada pixel em cada *frame* da sequência de vídeo. Para esta abordagem de fluxo ótico, foi possível distinguir, então, épocas de movimento e épocas sem movimento.

## 2.6 Proposta de trabalho

Após a análise dos objetivos discutidos, e os diversos requisitos ou considerações para encontrá-los, discutidos neste capítulo, foram determinados os diversos passos e instrumentos para os alcançar.

Nesta dissertação, o objetivo será então contribuir com um sistema de vídeo, já incorporado na Unidade de Monitorização de Epilepsia (em Munique, Alemanha), com base num sensor RGB-D, *Kinect* v2. Este sistema será feito com o objetivo de adaptar o seguimento dos membros, usando o sensor, no cenário em questão. O sistema será, portanto, não intrusivo e apenas será feito com base em visão por computador.

Nesta dissertação os maiores objetivos serão:

1. Desenvolver uma interface que tenha requisitos considerados cruciais para este tipo de cenário;

2. Desenvolver um algoritmo capaz de seguir os membros neste tipo de cenário;
3. Quantificar a três dimensões a doença de Epilepsia com base nos pontos previamente adquiridos pelo algoritmo.

## Capítulo 3

# Seguimento de Membros Semi-Automático

Neste capítulo será descrito em detalhe o método escolhido para o problema em questão, assim como a apresentação dos resultados e do erro associado ao algoritmo.

### 3.1 Pré-processamento

A imagem de infravermelhos adquirida pela *Kinect V2* foi selecionada para fazer o seguimento dos membros. No entanto, o contraste da imagem é pobre e foi necessário melhorar através da Equalização de Histograma. Esta transformação não-linear estende o histograma da imagem tal que os valores de cinza tenham a mesma probabilidade de ocorrência [35]. Considerando uma imagem em níveis de cinza  $\{x\}$ , em que  $n_i$  é o número de ocorrências do nível de cinza de  $i$ , a probabilidade de ocorrência do nível do pixel  $i$  na imagem é dada por (3.1), em que  $L$  é o número total de níveis de cinza (geralmente 256),  $n$  o número total de pixels na imagem, e  $p_x(i)$  a probabilidade de ocorrência do pixel  $i$ , normalizado de  $[0,1]$ . A nova imagem é transformada através de  $T$ , que é o produto do histograma cumulativo e do fator de escala. O fator de escala é necessário para ajustar o novo valor de intensidade atribuído, dentro da gama de valores de intensidade, sendo a nova imagem,  $x'$ , devolvida pela equação (3.2). Na imagem 3.1 pode-se ver a transformação relativa à equalização de histograma.

$$p_x(i) = p(x = i) = \frac{n_i}{n}, \quad 0 \leq i < L \quad (3.1)$$

$$x' = T(x) = \sum_{i=0}^x n_i \frac{\text{max.intensidade}}{n} \quad (3.2)$$



Figura 3.1: Resultado da Equalização de Histograma (antes e depois)

## 3.2 Seguimento dos membros

Ao longo desta dissertação foram testados vários métodos que inicialmente seriam promissores para este tipo de cenário, no entanto o método apresentado nesta dissertação foi o que obteve melhores resultados.

O seguimento dos membros consiste na construção de um esqueleto sobre as juntas ou pontos estratégicos que permitem a caracterização espacial do corpo humano ao longo do tempo. Após a seleção das juntas/pontos, feita pelo utilizador, são calculadas as distâncias entre si, assim como a definição dos ângulos possíveis relativos à(s) parte(s) anatómica(s). Este passo é importante para garantir que não hajam movimentos sem nexos ou impossíveis anatomicamente. Este passo foi dividido em dois: estimativa e correção. A estimativa prevê o próximo ponto consoante os movimentos anteriores e verifica a vizinhança de forma a reconhecer qual o ponto com maior similaridade. A correção verifica se o ponto selecionado está dentro das condições estipuladas, relativas ao membro e corrige, caso seja necessário, o novo ponto. Com o intuito de facilitar a concessão do algoritmo, foi feito em linguagem C#, onde se pode usufruir de bibliotecas como o *EmguCV* [36] e *AForge* [37], que oferecem uma vasta gama de métodos de visão computacional.

### 3.2.1 Estimativa: Algoritmo Farneback

O algoritmo de Farneback traz vantagens relativamente a outros algoritmos de fluxo ótico (*Lukas-Kanade* [38], *Horn-Schunck* [39]) pois não calcula o fluxo ótico exclusivamente para pontos característicos da região a seguir, mas sim para todos os pontos da imagem. Neste tipo de cenário, tanto os membros, como a cabeça ao rodarem podem modificar os pontos característicos ao longo do tempo, sendo que não se enquadraria. Para além disso, o algoritmo de Farneback verificou-se eficaz no seguimento de membros do corpo humano com o objetivo de reconhecimento de ações humanas [40].

O algoritmo de Farneback [41], [42], [43] propõe o fluxo ótico de densidade com base na representação polinomial. A representação polinomial da imagem é obtida pela aproximação dos vizinhos locais através dos polinómios quadráticos, pelo que é obtido de forma eficaz com o uso da transformada de expansão polinomial. O desfaseamento entre *frames* pode ser obtido equacionando os coeficientes da base polinomial. A ideia da expansão polinomial é aproximar a vizinhança de cada pixel a um polinómio. O interesse reside nos polinómios quadráticos, que fornece um modelo do sinal local. Este pode ser expresso num sistema de coordenadas dado por (3.3).

$$f(x) \approx x^T A x + b^T x + c, \quad (3.3)$$

onde  $A$  é uma matriz simétrica,  $b$  um vetor e  $c$  um escalar. Os coeficientes são estimados pelo peso mínimo dos quadrados para se ajustarem aos valores na vizinhança. O peso tem duas componentes denominadas de certeza e aplicabilidade. A certeza é acoplada ao sinal da vizinhança, pelo que os valores fora da imagem não têm qualquer impacto na estimativa de coeficientes para a vizinhança. A aplicabilidade determina o peso relativo de cada ponto da vizinhança, com base na sua posição. Geralmente, é atribuído um peso mais significativo ao ponto central da imagem, decrescendo radialmente. Assim como em todos os algoritmo de fluxo ótico, é assumido que o brilho na imagem é constante, ou seja o brilho em *frames* consecutivas é constante. Considerando um movimento de translação  $d$  encontrado no ponto  $(x, y)$  na imagem e construindo um novo sinal,  $f_2$ , por esse mesmo deslocamento,

$$\begin{aligned} f_2(x) &= f_1(x-d) = (x-d)^T A_1 (x-d) + b_1^T (x-d) + c_1 \\ &= x^T A_1 x + (b_1 - 2A_1 d)^T x + d^T A_1 d - b_1^T d + c_1 \\ &= x^T A_2 x + b_2^T x + c_2. \end{aligned} \quad (3.4)$$

Ao equacionar os coeficientes em dois polinómios supondo que o brilho é constante, temos:

$$\begin{aligned} A_2 &= A_1 \\ b_2 &= (b_1 - 2A_1 d) \\ c_2 &= d^T A d - b_1^T d + c_1. \end{aligned} \quad (3.5)$$

Assumindo que  $A$  é não-singular,

$$d = -\frac{1}{2} A^{-1} (b_2 - b_1), \quad (3.6)$$

Portanto, é possível obter o deslocamento em cada ponto da imagem, assumindo que existe uma sobreposição entre a região de interesse, isto é, a vizinhança nas *frames* adjacentes. Supondo que se tem o deslocamento  $\bar{d}$ , extrai-se a região de interesse (ROI) da vizinhança no ponto  $P(x, y)$  e no ponto  $P(x + d.x, y + d.y)$ . O polinómio é extraído e calcula-se o que foi explicado anteriormente.

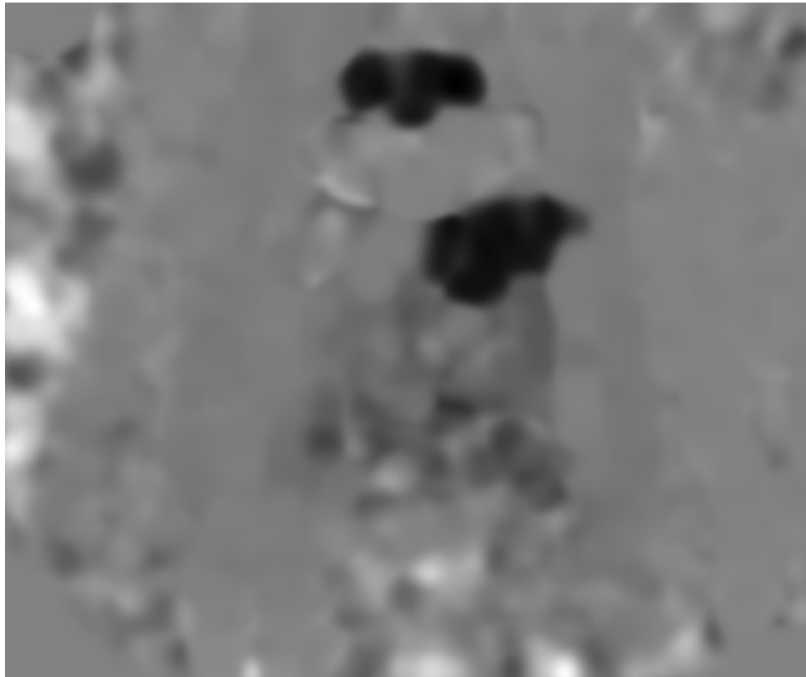


Figura 3.2: Exemplo do resultado da diferença entre *frames* com filtro Gaussiano. As regiões de maior contraste indicam as regiões onde ocorreu movimento

O deslocamento pode então ser estimado por

$$\begin{aligned} \bar{b}_2 &= b_1 - 2A_1(\bar{d}) \\ \text{sabendo } A_1, \bar{d} \text{ e } b_1 & \\ d &= 0.5 * A^{-1}(b_2 + \bar{b}_2 - b_1) \end{aligned} \quad (3.7)$$

Assim, um esquema iterativo pode ser usado a cada iteração sucessiva para obter uma melhor previsão do vetor de deslocamento. As iterações podem ser concluídas quando a mudança do vetor é inferior a um determinado limiar nas sucessivas iterações ou um número específico de iterações é definido. O vetor de deslocamento inicial é assumido  $(0,0)$ . Assim sendo, a ROI ou um fragmento da imagem na *frame* anterior e atual estão iguais.

Para que o fluxo ótico fosse mais robusto e não fosse influenciado por pequenas alterações na imagem, antes de medir o deslocamento entre *frames*, aplicou-se um filtro Gaussiano à diferença entre as 2 *frames*. O filtro Gaussiano, apesar de diminuir a velocidade de processamento, consegue uma maior robustez pois indica que o fluxo é mais fluído ignorando pixels que possam induzir em erro e não pertencem ao fluxo que se está a analisar, esta diferença está representada na imagem 3.2. Como todos os métodos de fluxo ótico, este algoritmo de Farneback respeita a expressão 3.8.

$$Frame_{anterior}(x,y) = Frame_{posterior}(x + flow_x, y + flow_y), \quad (3.8)$$

onde  $flow_x$  e  $flow_y$  são os deslocamentos em  $x$  e  $y$ , respetivamente.

### 3.2.2 Correção: Similaridade entre pontos

A correção ocorre quando a função de similaridade de distância entre os membros está abaixo de um limiar, indicando que o movimento foi brusco e a estimativa não foi suficiente para acompanhar a localização exata do membro, pelo que é necessário reajustar o valor do ponto a seguir. As propriedades de fotometria como a intensidade e a textura são úteis e necessárias para verificar a semelhança entre imagens, no entanto quando o que se está a seguir fica ocluso ou mudou de forma, torna-se impossível de verificar se o objeto está de facto naquele ponto. Desta forma, é imprescindível verificar propriedades geométricas dos objetos, como o seu comprimento (dos membros), normal e curvatura. A partir do momento que as propriedades geométricas estão definidas para cada ponto/membro, constrói-se, através da profundidade, um mapa 3D do ROI e verifica-se em que ponto desse mapa existe maior similaridade. Existem funções de propriedades  $f_1, f_2, \dots, f_n$  para um conjunto de pontos  $X$ , que mapeia qualquer ponto  $x \in X$  nos vetores de propriedades  $f_1(x), f_2(x), \dots, f_n(x)$ . Para imagens 2D, estas propriedades podem ser intensidades, cores, gradientes ou texturas. A nível 3D, as propriedades incluem distâncias, normais e curvaturas. Para cada função de propriedade  $f$  que tem uma definição entre dois pontos,  $x$  e  $y$ , pode-se induzir uma função de similaridade  $s(x, y, f)$ . Assim, a similaridade entre pontos final é definida pela combinação de funções de similaridades das propriedades. A distância dos membros é uma propriedade fixa ao longo do tempo, pelo que é então usada para ajustar o novo ponto. Para medir a similaridade das distâncias dos membros, calcula-se as distâncias entre pontos e a sua similaridade através da fórmula descrita em [44]. A similaridade é dada pela expressão 3.9.

$$s(x, y, f_{distancia}) = 1 - \frac{f_{distancia}(y) - f_{distancia}(x)}{f_{distancia}(x)}, \quad (3.9)$$

onde  $y$  é o novo ponto a ser avaliado e  $x$  o ponto definido na primeira *frame* como base de comparação. Para o novo ponto ser considerado, terá de satisfazer a condição  $f_{distancia}(y) \leq f_{distancia}(x)$ . Onde  $f_{distancia}$  é obtida através da expressão 3.10.

$$f_{distancia} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}, \quad (3.10)$$

onde  $x_1, y_1, z_1$  e  $x_2, y_2, z_2$  são as coordenadas dos pontos pertencentes a cada membro (cotovelo e pulso, por exemplo).

### 3.2.3 Aquisição das coordenadas mundo

O sensor *Kinect v2* consegue calcular todas as distâncias entre o ambiente e o sensor. Estas distâncias podem ser guardadas numa matriz, cujo tamanho é a resolução da imagem de profundidade, e podem ser posteriormente utilizadas. Para cada pixel da imagem de profundidade 512x424, o dispositivo estima em tempo real a distância até ao correspondente ponto do objeto. Este valor de saída é denominado de mapa de profundidade (*depthmap*), uma imagem 2D em que as coordenadas 3D podem ser calculadas. Após uns passos de processamento, que serão brevemente explicados, é possível obter indiretamente os pontos capturados da cena ou do objeto. Assumindo

que a *Kinect* usa lentes normalizadas, os parâmetros intrínsecos como o comprimento focal  $f$  e as coordenadas centrais  $(c_x, c_y)$  podem ser determinados. Assumindo também que a variação de sensor para sensor destes parâmetros é mínima e não influencia significativamente os resultados, são usados os valores listados em [9]. A coordenada  $Z$  já é conhecida, pois é dada pelo mapa de profundidade, conhecendo as coordenadas de pixel. Estes parâmetros são usados em conjunto com a perspectiva da relação da projeção de Mutto, em 2003, em que um ponto 3D pode ser estimado através das coordenadas de imagem de um pixel  $p = [u, v, 1]^T$ , através da equação 3.11. Cada pixel representa no mapa de profundidade um valor correspondente nas coordenadas mundo. O número total de coordenadas mundo é de 217088. Na tabela 3.1 estão os valores dos parâmetros intrínsecos usados para adquirir as coordenadas em 3D.

$$\sigma p = KX = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} = \begin{cases} X = \frac{u-c_x}{f_x} * Z \\ Y = \frac{v-c_y}{f_y} * Z \end{cases} \quad (3.11)$$

Tabela 3.1: Parâmetros Intrínsecos Estimados da Kinect v2, adaptado de [9]

Parametro	Valor (pixeis)	Desvio padrao
Comprimento focal ( $f_x$ )	364,7	1,6
Comprimento focal ( $f_y$ )	366,1	1,5
Ponto central ( $c_x$ )	255,8	2,4
Ponto central ( $c_y$ )	203,7	3,1

### 3.3 Resultados do Seguimento de Membros

De forma mais sucinta, o método proposto inicia então com a primeira *frame* e a marcação dos 8 pontos a seguir. Para poder corrigir sempre os valores de acordo com as novas posições, calcula-se a distância Euclidiana 3D entre os membros:

1. Cabeça (C.) - Torso (T.)
2. Ombro Esquerdo (O.E.) - Ombro Direito (O.D.)
3. O.E. - Cotovelo Esquerdo (C.E.)
4. O.D. - Cotovelo Direito (C.D.)
5. C. E. - Pulso Esquerdo (P.E.)



## 6. C. D. - Pulso Direito (P.D.).

Devido à anatomia especial do corpo humano, ajustou-se alguns parâmetros relativos à correção destas distâncias. Na figura 3.3 podemos verificar as articulações respetivas ao corpo humano. Vamos aproximar a cabeça ao pescoço ("Neck"), e o torso à lombar ("Lumbar Spine"). A cabeça tem um alcance espacial de movimento limitado, como podemos verificar na figura 3.4, a cabeça consegue rodar, fletir, estender e virar lateralmente (apenas até 35°), mas sem mover significativamente o local onde está. O torso também tem os seus movimentos limitados, mas que com ajuda de outras articulações permite uma maior liberdade de movimento, como podemos verificar na figura 3.5. Como também podemos verificar, como estas articulações estão interligadas, com o aumento do movimento da lombar, o movimento da cabeça pode aumentar ainda mais (em termos espaciais). Assim sendo, como o torso tem um limite espacial de movimento, considera-se que os seus movimentos, apesar de bruscos, não será necessário correção. Pelo que sempre que esta distância (C.- T.) estiver com similaridade inferior ao limiar estipulado, apenas a cabeça irá ser corrigida. Em caso de exceção, ou seja, aquando a cabeça permaneceu estática e o torso apresentou um grande deslocamento dado pelo fluxo ótico, será o ponto do torso a ser corrigido.

Relativamente aos ombros, estes apresentam 3 ângulos de liberdade [45], mas que em termos espaciais não apresenta grandes movimentos, como verificamos na figura 3.6. Pelo que a correção só ocorrerá entre os ombros caso um deles apresente um deslocamento superior ao outro e a similaridade for inferior ao limiar. Como o ombro consegue ser mais estável, mas dá grande movimento aos cotovelos, como mostram as figuras 3.7 e 3.8, sempre que é necessário a correção desta distância, será o cotovelo o valor reajustado. Após estes valores estarem mais robustos, e caso a similaridade entre o cotovelo e o pulso ser baixa, será o pulso a ser reajustado. O pulso, por ter grande movimento dado pelo ombro e cotovelo, pode mover-se de forma mais livre e incerta, como exemplifica a figura 3.9. Assim sendo, é o último valor a ser reajustado, caso necessário, após os outros terem sido avaliados e considerados como certos.

Após estas medições estarem guardadas avança-se para a próxima *frame*, e calcula-se o fluxo ótico de Farneback. O fluxo de Farneback devolve o deslocamento e assim se ajusta os novos valores. Caso a similaridade de alguma das distâncias entre os membros seja inferior a 80%, corrige-se esse valor com base no que foi descrito anteriormente, procurando o máximo da similaridade na sua vizinhança. Este processo é feito até atingir a última *frame*. O diagrama de atividade UML representado em 3.10 representa o algoritmo proposto. Para escolher a largura da vizinhança mais adequada tentou-se atingir um compromisso entre o tempo despendido versus a exatidão que o algoritmo oferecia. Nos gráficos 3.11 e 3.12 podemos verificar que existe um aumento de exatidão consoante o aumento da janela, no entanto, o aumento da janela implica um aumento bastante significativo do tempo despendido. Verifica-se também que a partir da janela de 20 pixeis, o valor de exatidão tende a manter-se, assim sendo, não será necessário ter uma janela de vizinhança superior a 20 pixeis. Assumindo que os valores de exatidão são suficientes para o objetivo final, optou-se por escolher um valor que não compromete-se significativamente o tempo de computação, sendo este valor entre 14 a 20 pixeis. Desta forma, optou-se por uma janela de vizinhança de 16 pixeis.

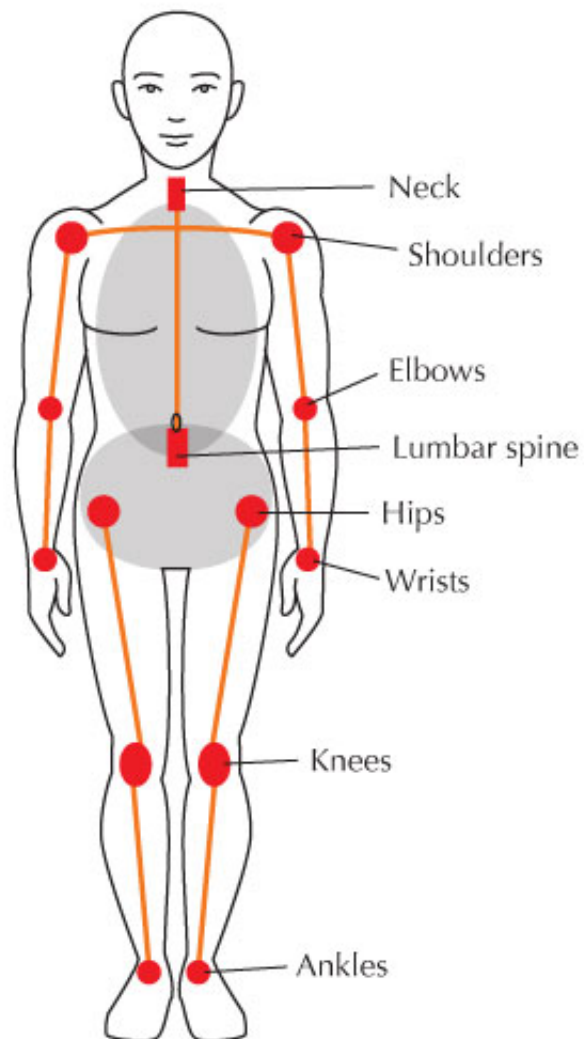


Figura 3.3: Articulações Corpo Humano. Adaptado de [8]

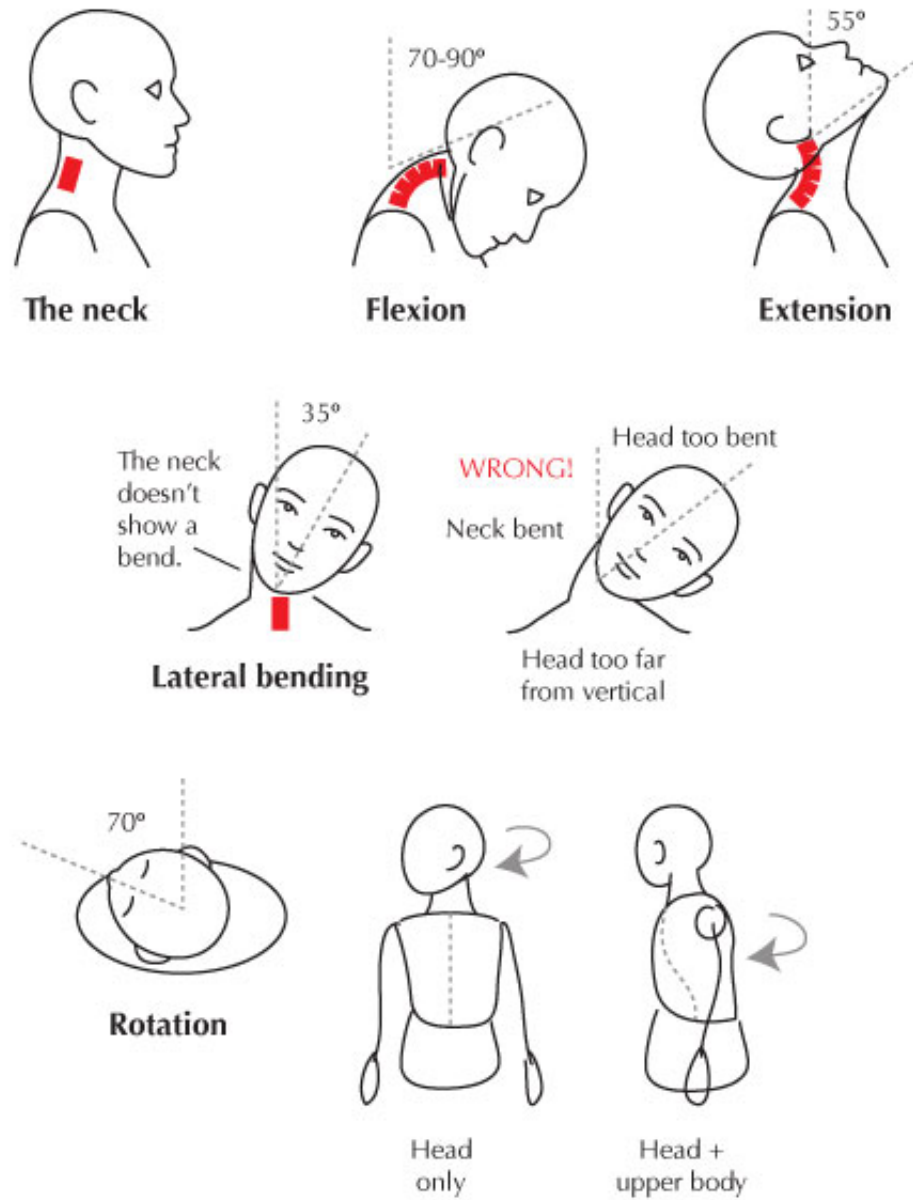


Figura 3.4: Movimento do Pescoço. Adaptado de [8]

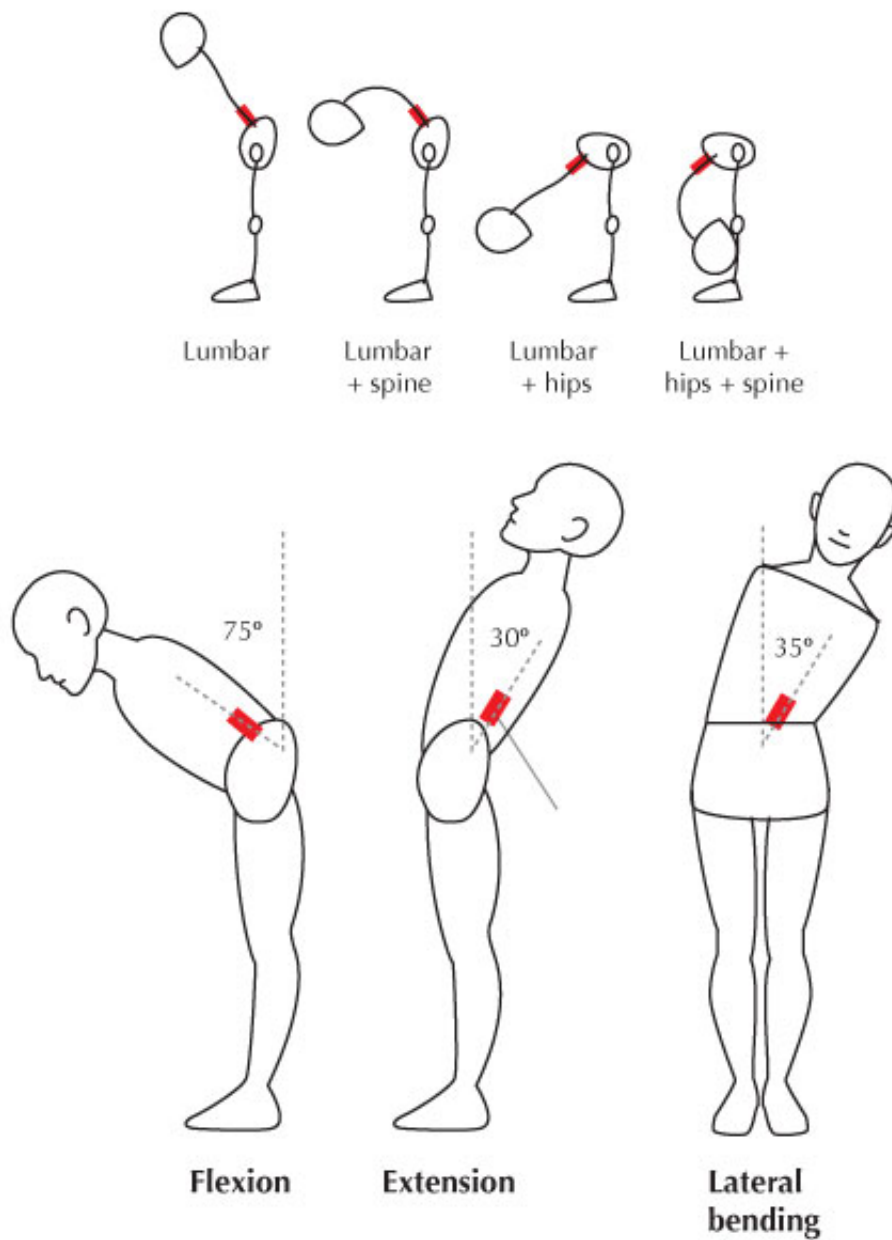


Figura 3.5: Movimento da Lombar. Adaptado de [8]

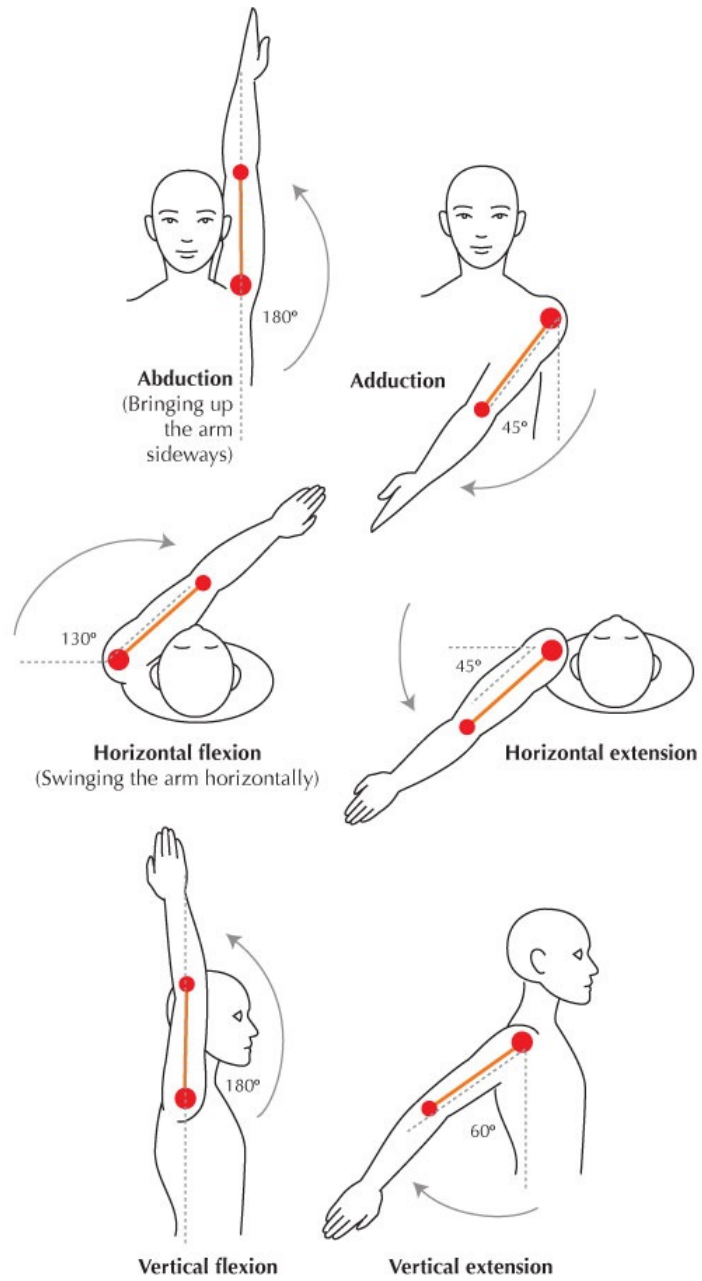


Figura 3.6: Movimento dos ombros. Adaptado de [8]

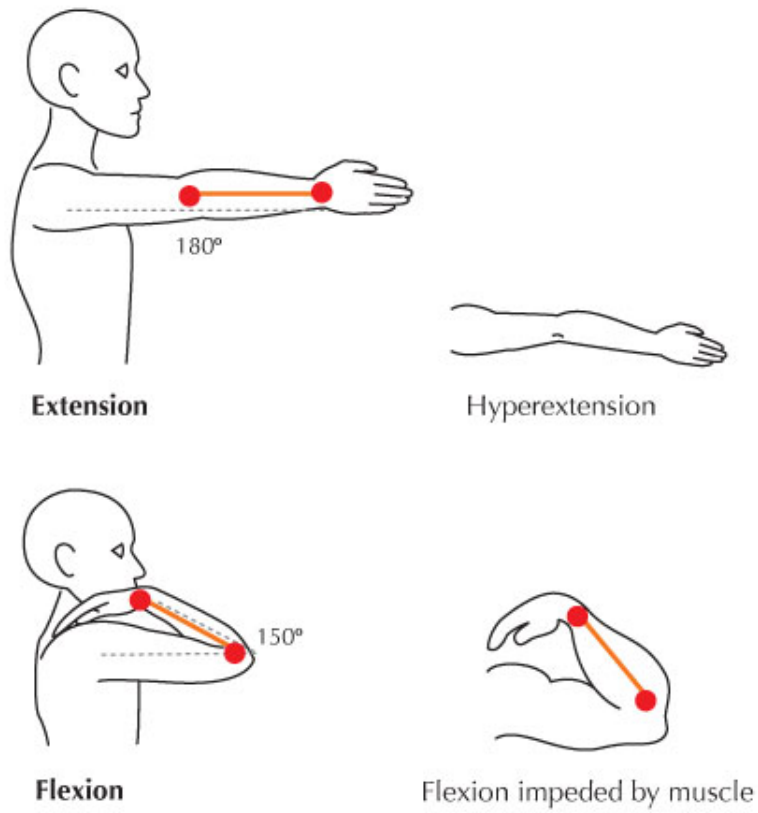


Figura 3.7: Movimento do cotovelo. Adaptado de [8]

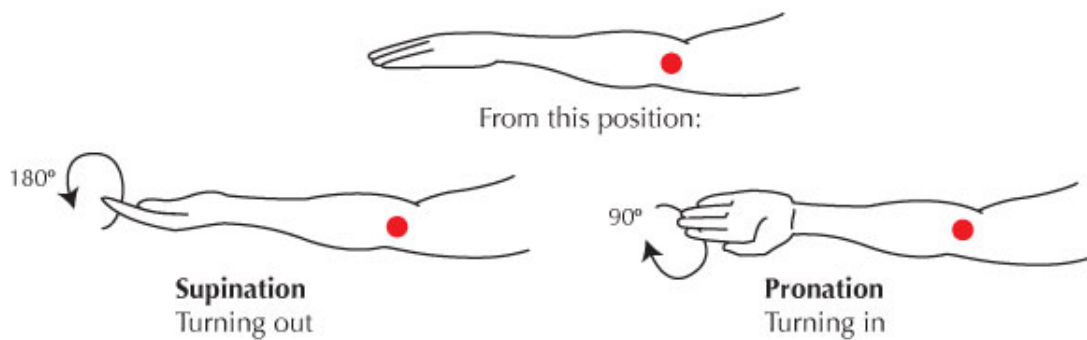


Figura 3.8: Movimento do cotovelo. Adaptado de [8]

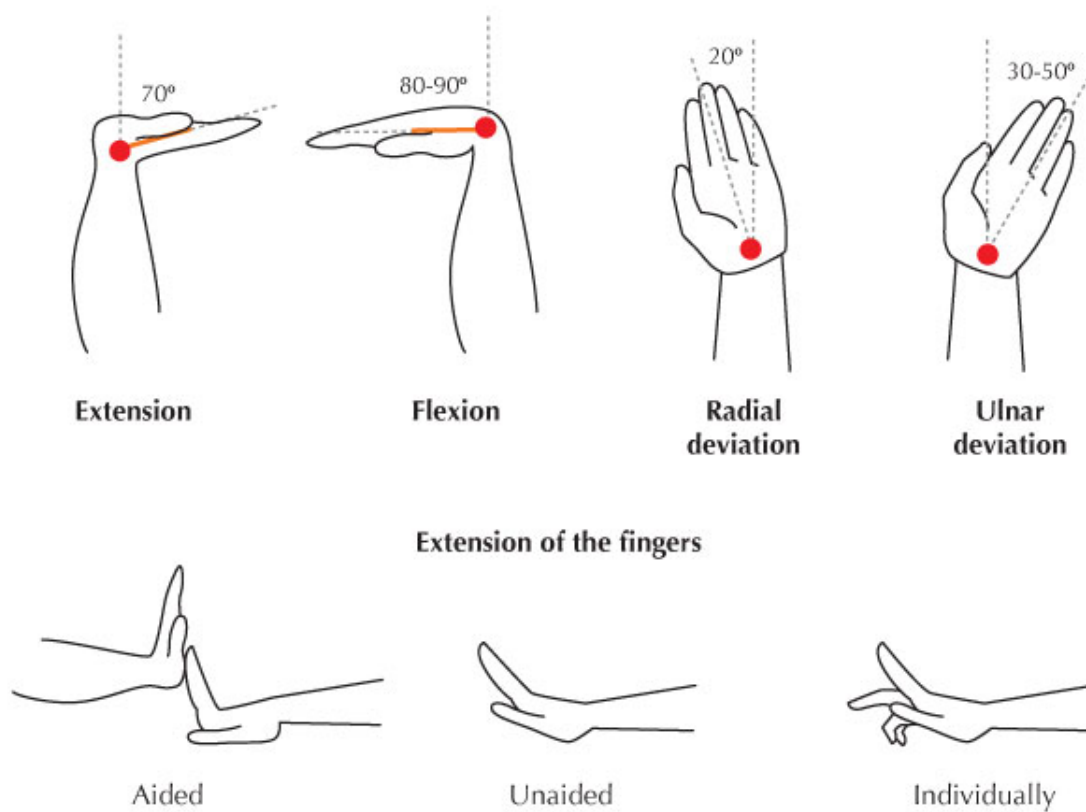


Figura 3.9: Movimento do pulso. Adaptado de [8]

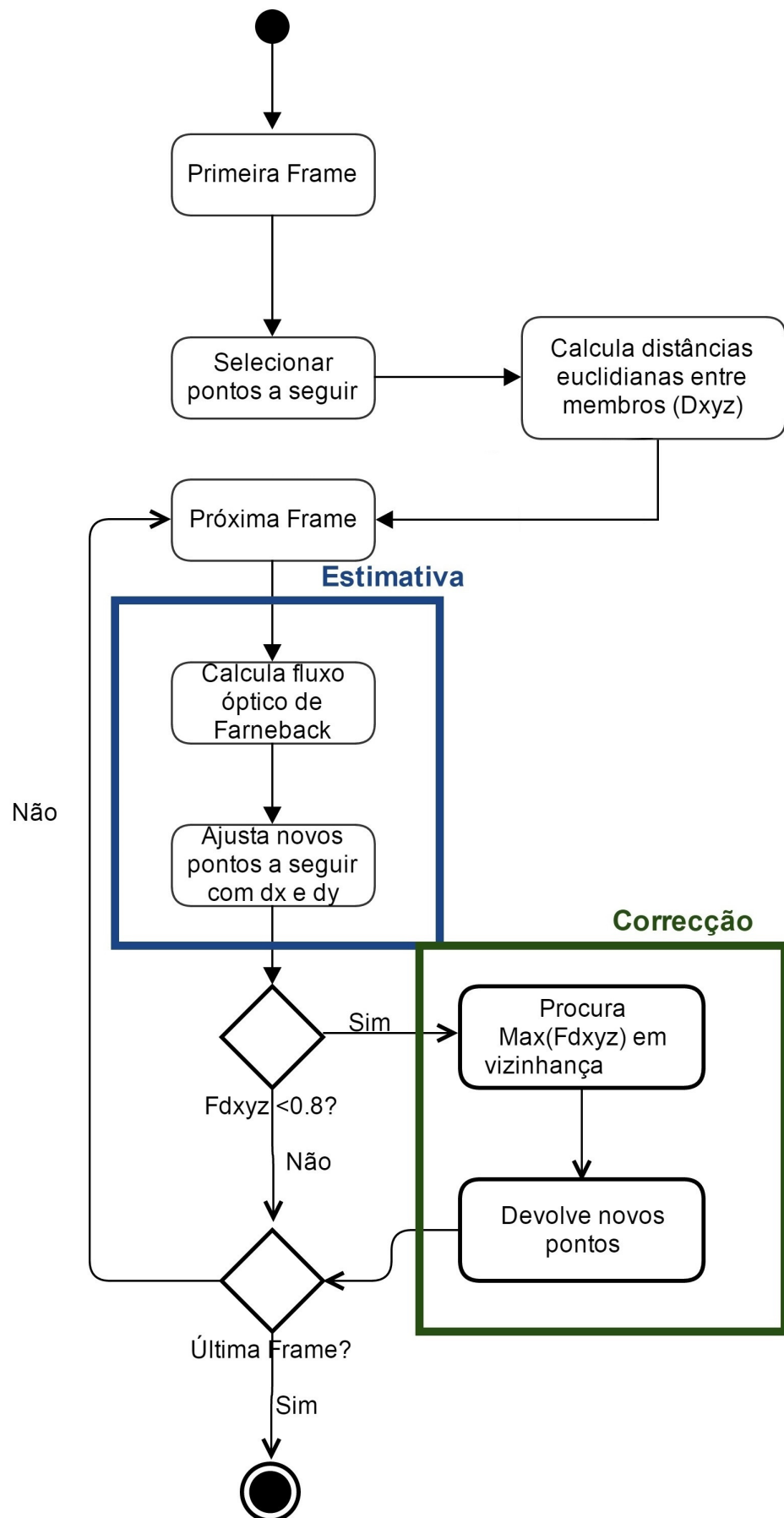


Figura 3.10: Diagrama de Atividade UML do algoritmo proposto



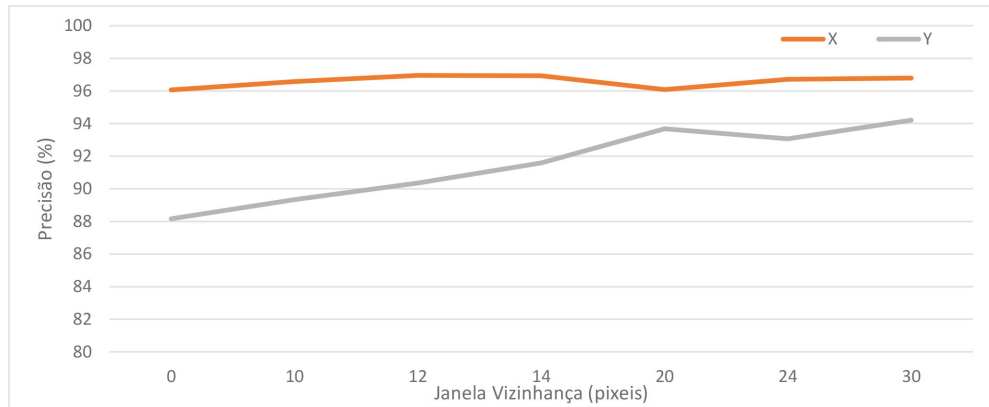


Figura 3.11: exatidão relativa à mudança da largura da janela de vizinhança

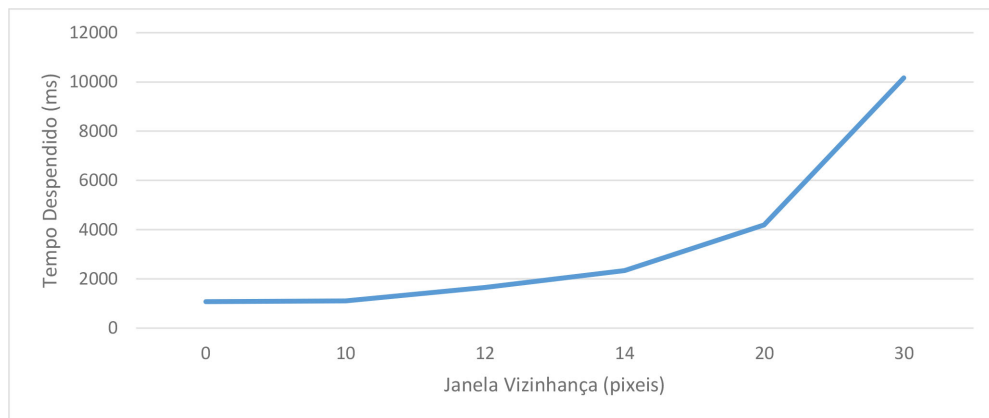


Figura 3.12: Tempo despendido relativamente à mudança da largura da janela de vizinhança

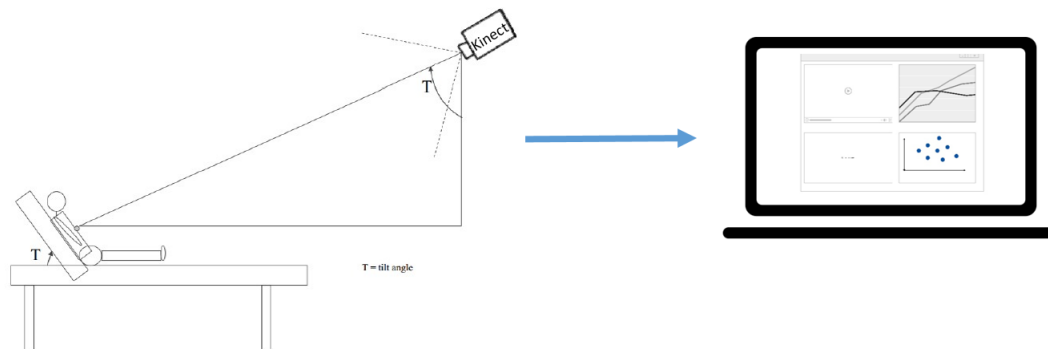


Figura 3.13: Configuração de Aquisição

### 3.3.1 Caracterização dos Dados

Todos os dados foram adquiridos na Unidade de Motorização de Epilepsia do Hospital de Munique, através do sensor RGB-D Microsoft Kinect K4W e a com a configuração de aquisição conforme é demonstrado na figura 3.13. Estes dados foram posteriormente analisados por profissionais de saúde onde marcaram as crises captadas pelo sensor. Esta análise foi processada na ferramenta KiMA ([46]), onde é possível selecionar e gravar eventos que sejam considerados importantes. No total foram compostos 24 crises (2 pacientes) selecionadas pelos profissionais de saúde. Destas 24 crises apenas foram analisadas 8, que obedeciam aos seguintes critérios:

1. Não seja interrompida por enfermeiros
2. Não sejam crises que apenas provoquem movimentos em membros inferiores
3. Não sejam crises com movimentos pouco significativos

Das 24 crises, 7 foram interrompidas por enfermeiros, 6 não apresentavam qualquer movimento significativo e 3 eram de crises relativamente aos membros inferiores. Dentro das 8 crises analisadas, estas decorriam com o paciente voltado frontalmente para o sensor ou na posição fetal, em algum momento da crise haviam membros que ficavam obstruídos pelo lençol ou até mesmo por outro membro (caso estivessem na posição fetal), todas estas características estão sintetizadas na tabela 3.2. No total foram analisadas 4344 *frames*.

Foi analisado nas 8 crises o seguimento de 8 articulações do corpo, anteriormente descritas. Após o seguimento estar completo consegue-se quantificar a crise através do deslocamento de *frame* para *frame*, fornecido pelo Fluxo Ótico de Farneback, variação da profundidade, velocidade

Tabela 3.2: Caracterização das crises avaliadas

Crises	Articulações obstruídas (lençol ou outro membro)						Articulações com movimentos bruscos						Posição					
	C.	T.	O.E.	O.D.	C.E.	C.D.	P.E.	P.D.	C.	T.	O.E.	O.D.	C.E.	C.D.	P.E.	P.D.	Frontal	Fetal
Run1	X			X		X	X	X	X	X								X
Run10	X			X		X	X	X	X						X			X
Run11						X	X	X	X	X					X			X
Run12				X		X	X	X	X				X	X	X			X
Run13						X	X	X	X	X					X			X
Run16						X	X	X	X	X					X			X
Run20	X					X	X	X	X						X			X
Run21						X	X	X	X			X	X	X	X			X



Figura 3.14: Representação do movimento ao longo do tempo durante a crise quantificada

e aceleração 3D. Após o mapeamento das coordenadas 2D para coordenadas mundo 3D, é possível calcular a velocidade e a aceleração.

$$Velocidade_{3D} = \sqrt{v_x^2 + v_y^2 + v_z^2} \approx \sqrt{\frac{\Delta x^2 + \Delta y^2 + \Delta z^2}{\Delta t^2}} \quad (3.12)$$

$$Aceleracao_{3D} = \sqrt{a_x^2 + a_y^2 + a_z^2} \approx \sqrt{\frac{\Delta v_x^2 + \Delta v_y^2 + \Delta v_z^2}{\Delta t^2}} \quad (3.13)$$

Numa das crises avaliadas, figura 3.14, foi aplicado o algoritmo de forma a quantificar este movimento. O movimento implica principalmente movimentos sistemáticos da cabeça, no entanto também apresenta movimentos no torso e membros. Esta crise dura alguns segundos e após o seguimento dos membros estar concluída, foi possível obter a velocidade, aceleração e posições ( $X$ ,  $Y$  e  $Z$ ) das 8 articulações. Como se pode verificar nas várias *frames*, o paciente tem um movimento brusco na cabeça, enquanto que um dos pulsos se encontra obstruído pelo lençol, e o resto do corpo se move ligeiramente para o lado.

Na figura 3.15, comparativamente com as figuras 3.16, 3.17, 3.18, 3.19, 3.20, 3.21, 3.22, consegue-se perceber que a velocidade e aceleração da cabeça foram praticamente o dobro das restantes articulações.

As figuras 3.23, 3.24, 3.25 mostram os traçados de  $X$ ,  $Y$  e  $Z$ , respetivamente de cada membro seguido. A figura do eixo do  $Z$ , em particular, demonstra a importância que têm de se fazer

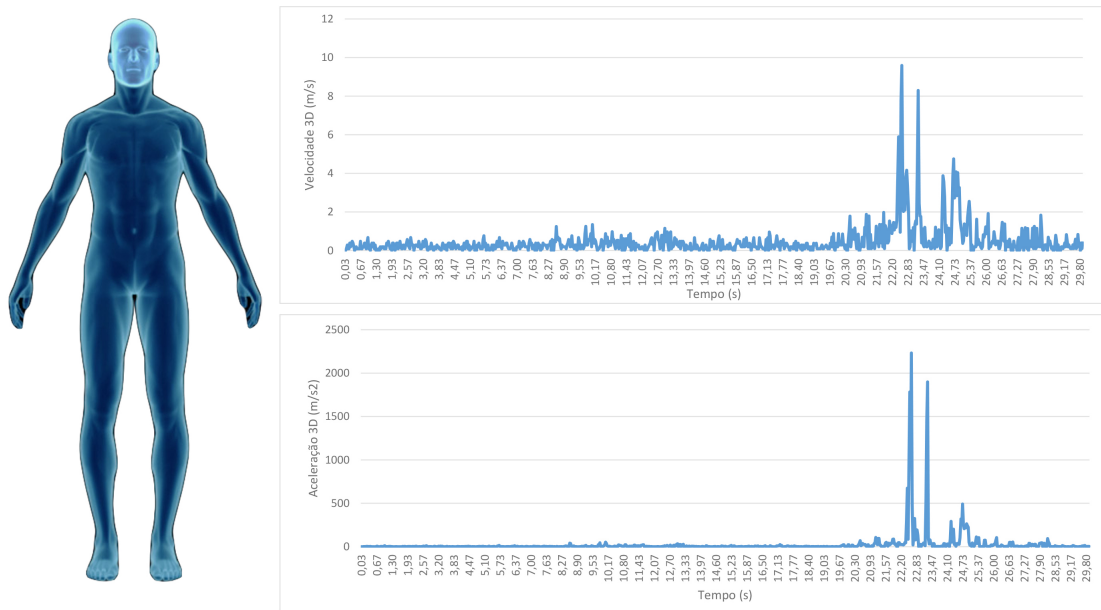


Figura 3.15: Traçado da velocidade e aceleração 3D na cabeça ao longo do tempo (em segundos)

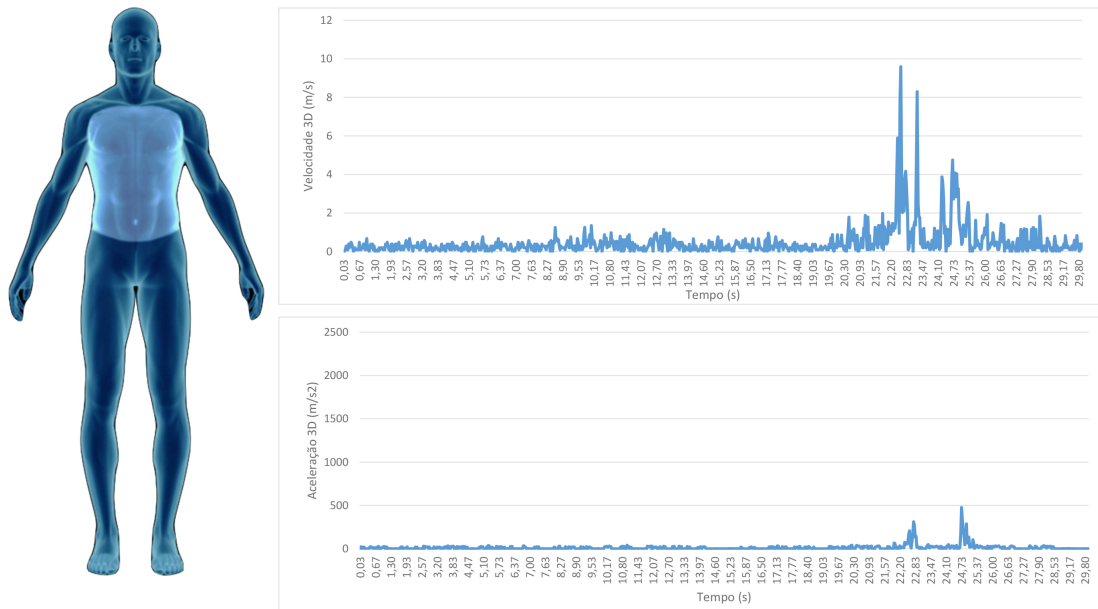


Figura 3.16: Traçado da velocidade e aceleração 3D no torso ao longo do tempo (em segundos)

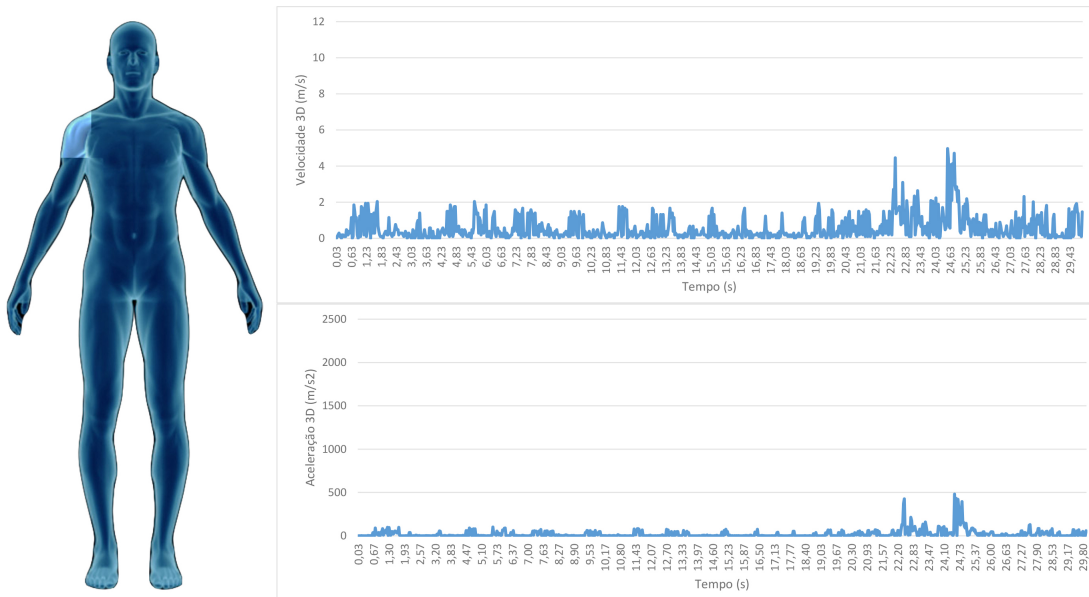


Figura 3.17: Traçado da velocidade e aceleração 3D do ombro direito ao longo do tempo (em segundos)

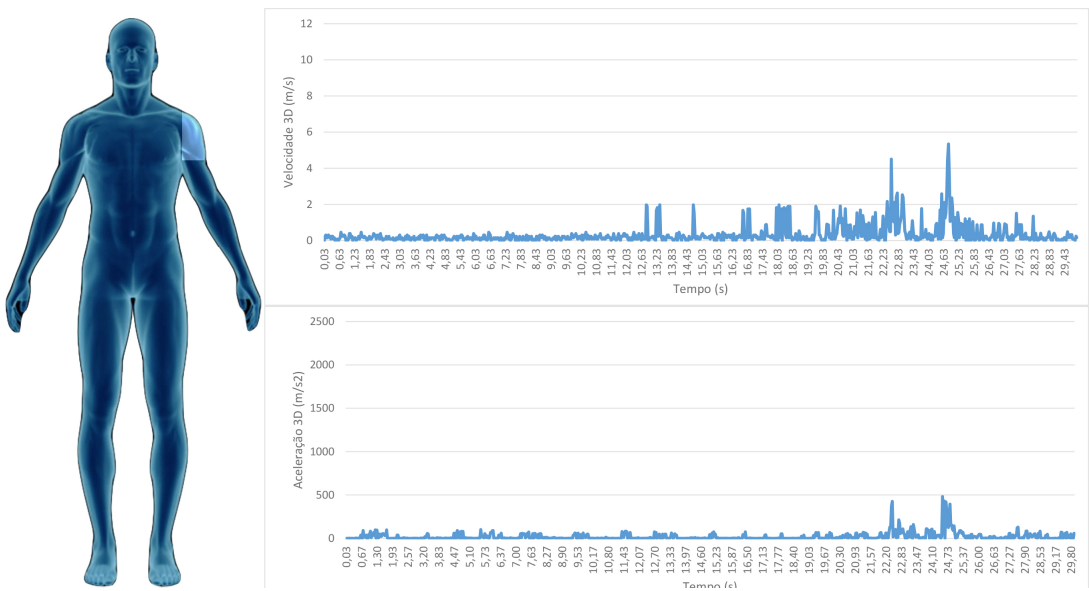


Figura 3.18: Traçado da velocidade e aceleração 3D do ombro esquerdo ao longo do tempo (em segundos)

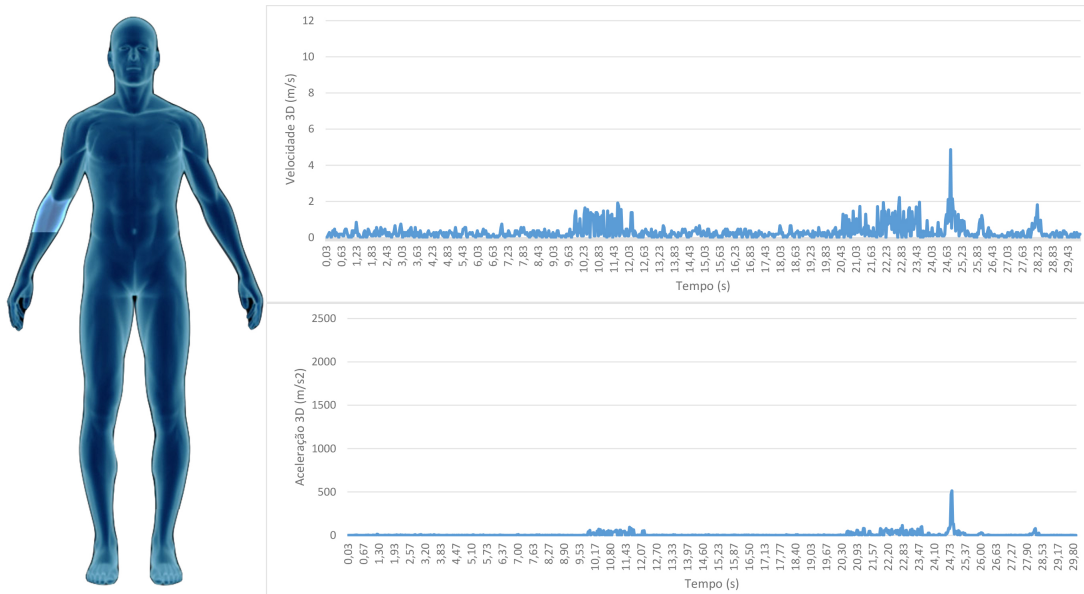


Figura 3.19: Traçado da velocidade e aceleração 3D do cotovelo direito ao longo do tempo (em segundos)

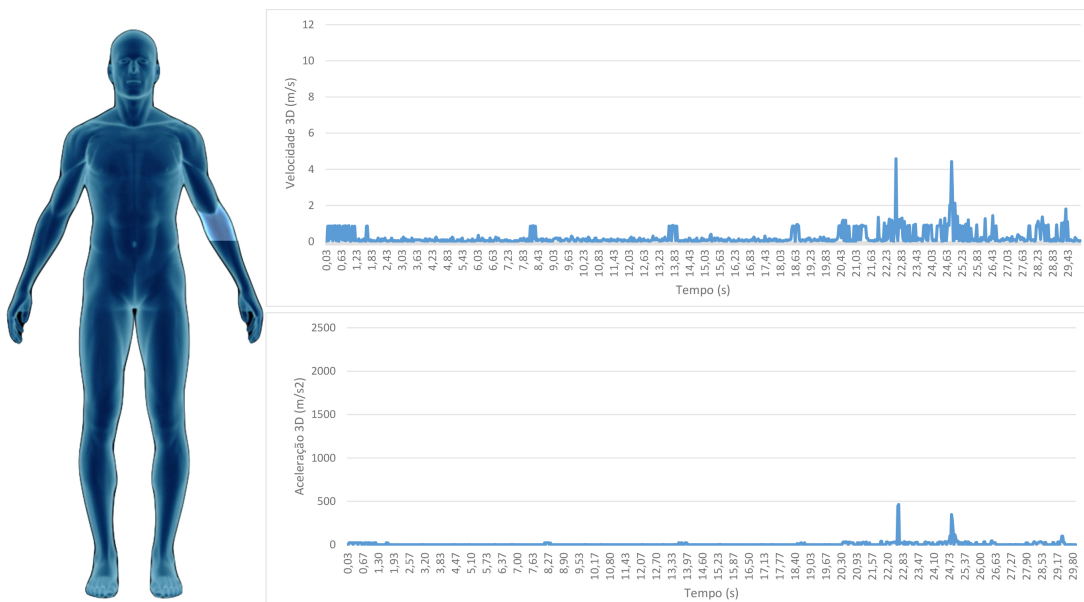


Figura 3.20: Traçado da velocidade e aceleração 3D do cotovelo esquerdo ao longo do tempo (em segundos)

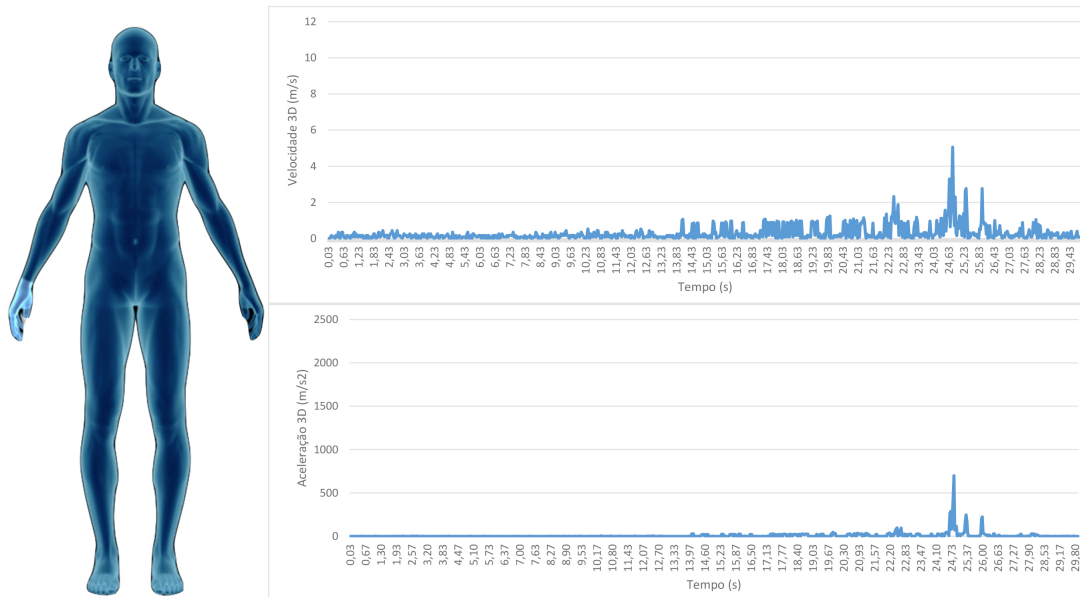


Figura 3.21: Traçado da velocidade e aceleração 3D do pulso direito ao longo do tempo (em segundos)

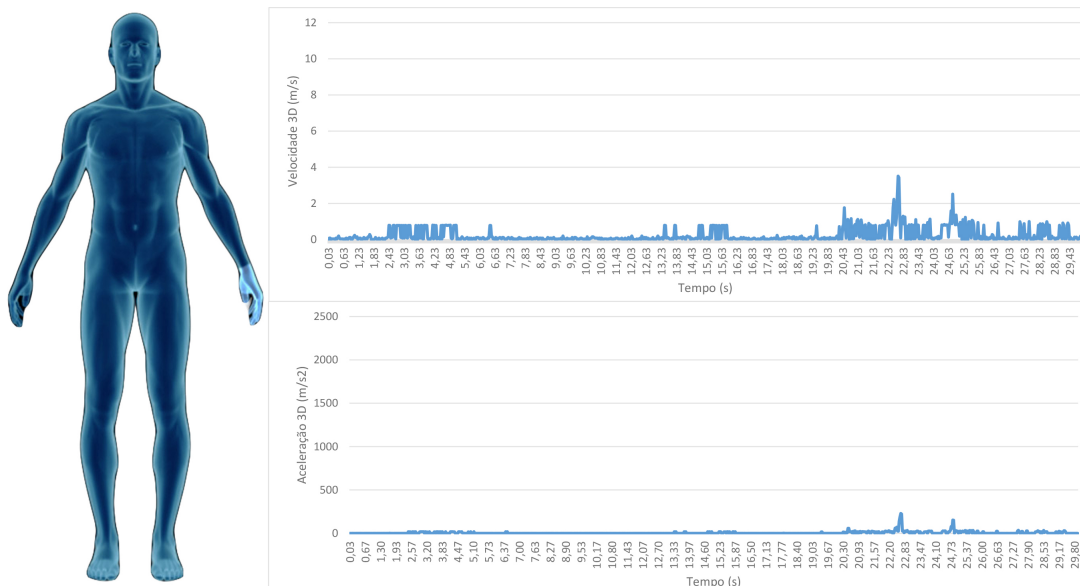


Figura 3.22: Traçado da velocidade e aceleração 3D do pulso esquerdo ao longo do tempo (em segundos)



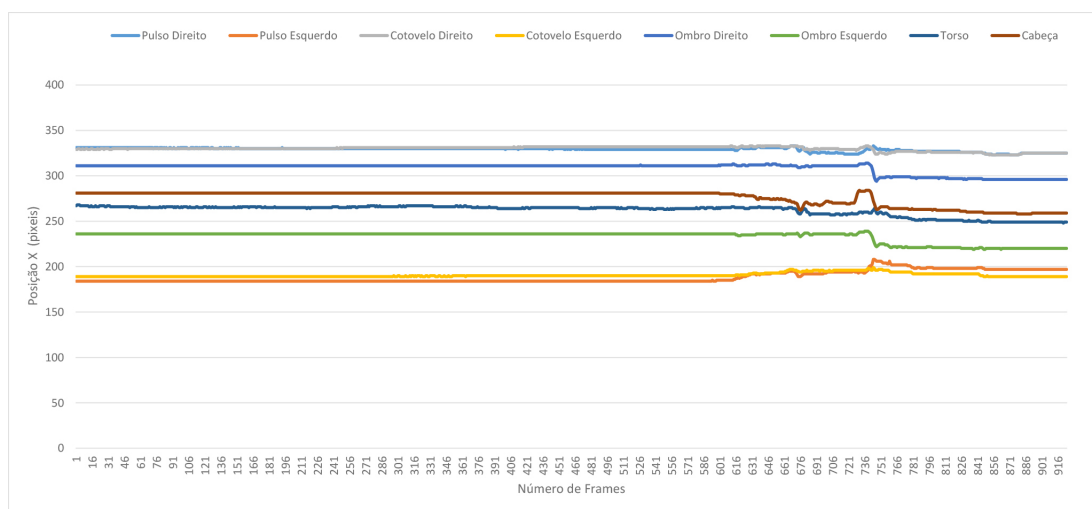


Figura 3.23: Traçado dos centróides dos movimentos das juntas do corpo na coordenada X (em pixels por *frame*)

uma quantificação em três dimensões. Visto que não seria possível obter esta informação, em profundidade, caso o seguimento fosse feito apenas em duas dimensões.

Para obter uma perspectiva mais detalhada do erro ao longo das *frames* é também demonstrado num gráfico da figura 3.26 a diferença entre os pontos devolvidos pelo algoritmo e as marcações feitas à mão no MaxTRAQ [47]. Pode-se concluir que o erro apresenta-se significativamente baixo, mas que aumenta quando os movimentos são mais bruscos.

### 3.4 Validação do Algoritmo

Como foi referido anteriormente, existiam alguns artefactos relativamente ao conjunto de crises avaliadas, nomeadamente a obstrução de articulações pelo lençol. A correção foi introduzida devido a falta de contraste devido à imagem ser em Infravermelhos, ou seja, as cores eram em níveis de cinza. Este artefacto trouxe muitas vezes dificuldade na distinção de membros entre outros objetos nas crises analisadas e consecutivamente à falha relativa do fluxo ótico, que considera primordialmente os valores dos pixels e seus vizinhos. Ao colmatar com a correção, este algoritmo demonstrou-se mais robusto e dificilmente perdia-se o seguimento da articulação. No gráfico 3.11 podemos verificar que sem correção existe alguma diferença do que com correção. É de notar, que apesar da exatidão do algoritmo sem correção ter valores consideravelmente altos, estes no entanto não correspondem totalmente ao sucedido, pois após 10 *frames* as articulações dos pulsos e cotovelos foram perdidas. Enquanto que com correção era mais difícil de se perder. É fácil de perceber, através do gráfico que quanto maior fosse a janela de procura, maior a exatidão e maior dificuldade tinha de perder o seguimento das articulações. Estes valores foram adquiridos ao implementar o algoritmo na crise que mais artefactos tinha do conjunto de dados, assim como movimentos mais violentos, pelo que o fluxo ótico nestas condições não obteve muito sucesso. Pode ver-se na sequência de *frames* a diferença entre não ter correção (SC) e ter correção (CC), na

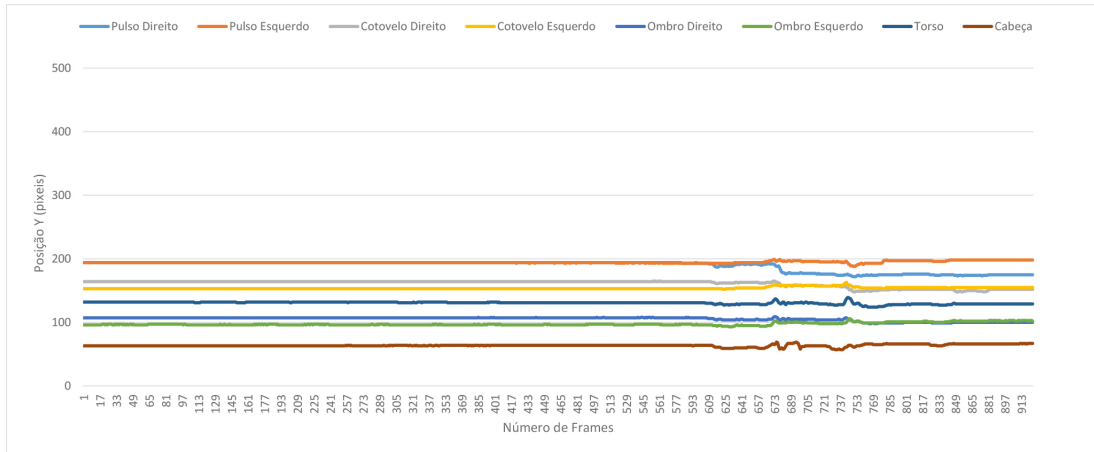


Figura 3.24: Traçado dos centróides dos movimentos das juntas do corpo na coordenada Y (em pixels por *frame*)

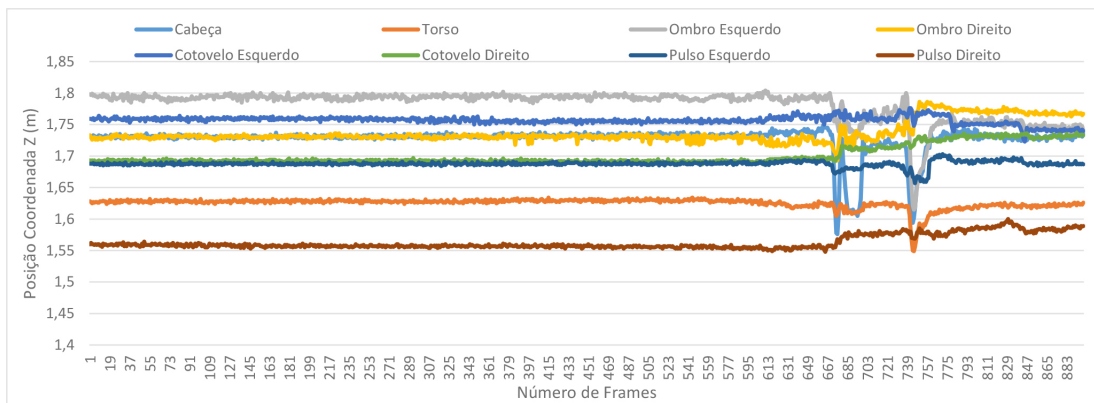


Figura 3.25: Traçado dos centróides dos movimentos das juntas do corpo na coordenada Z (em metros por *frame*)



Figura 3.26: Traçado da média do erro dos centróides (X e Y) das juntas do corpo ao longo do tempo (em *frames*)

figura 3.27.

No entanto, para além da obstrução provocada pelos lençóis, ocorria por vezes obstrução provocada por outra articulação. Nessas situações tornava-se impossível fazer a correção, visto que o sensor não conseguia obter a profundidade dos objetos obstruídos por outros.

De forma a validar o algoritmo, foi marcado manualmente no MaxTRAQ [47] todas as crises avaliadas. Com essa marcação foi possível verificar o erro obtido com este algoritmo, assim como o tempo despendido nas duas abordagens. O gráfico da figura 3.28 apresenta a exatidão relativa a cada crise para as coordenadas de  $X$  e  $Y$ . Estes valores estão significativamente altos, principalmente nas crises em que o paciente estava frontalmente para o sensor. Para as crises em que os pacientes se encontravam em posição fetal, os valores de exatidão descenderam, sendo que na maioria era impossível corrigir devido a obstrução contínua de certas articulações. No entanto, este método obteve uma exatidão à volta de 92%.

No MaxTraQ, para conseguir marcar 8 articulações em 248 *frames* levou cerca de 38 minutos, enquanto que com o algoritmo proposto seria apenas necessário à volta de 8/10 minutos (dependendo das correções a serem feitas). Houve portanto um ganho de 78% em termos de tempo despendido para avaliar uma crise com 8 articulações.

Em mais detalhe, é possível verificar na tabela 3.3 o erro relativo a cada membro seguido para cada vídeo testado.

Através do gráfico 3.28 e da tabela 3.3 é possível perceber que a crise *Run21* teve a pior exatidão. Esta crise encontrava-se em posição fetal e na maior parte do tempo teve a maioria dos membros obstruídos, sendo impossível fazer a correção. No entanto, este tipo de algoritmo não foi pensado para testar em crises quando os pacientes se encontravam em posição fetal, mas sim em posição frontal, onde seria possível analisar fielmente os movimentos. Nestes casos (posição fetal), seria necessário incorporar outro método que previsse esta posição e a constante obstrução de membros, como por exemplo, estimar que o membro permanecia no mesmo sítio até reaparecer e/ou acrescentar características relativas aos membros em que fosse possível ao algoritmo estimar que este se encontrava ocluído, por não conseguir encontrar essas mesmas características. Outra das falhas do algoritmo é quando existe coordenadas que podem apresentar uma maior similaridade e que estejam vizinhos à articulação em causa. Esta falha ocorreu na crise *Run10* no pulso esquerdo, perdendo esta articulação e assim sendo obteve um erro mais significativo (75,60%). Para além desta crise, apenas as que apresentavam uma posição fetal, *Run1* e *Run21*, é que obtiveram valores mais altos de erro, devido as razões mencionadas anteriormente.

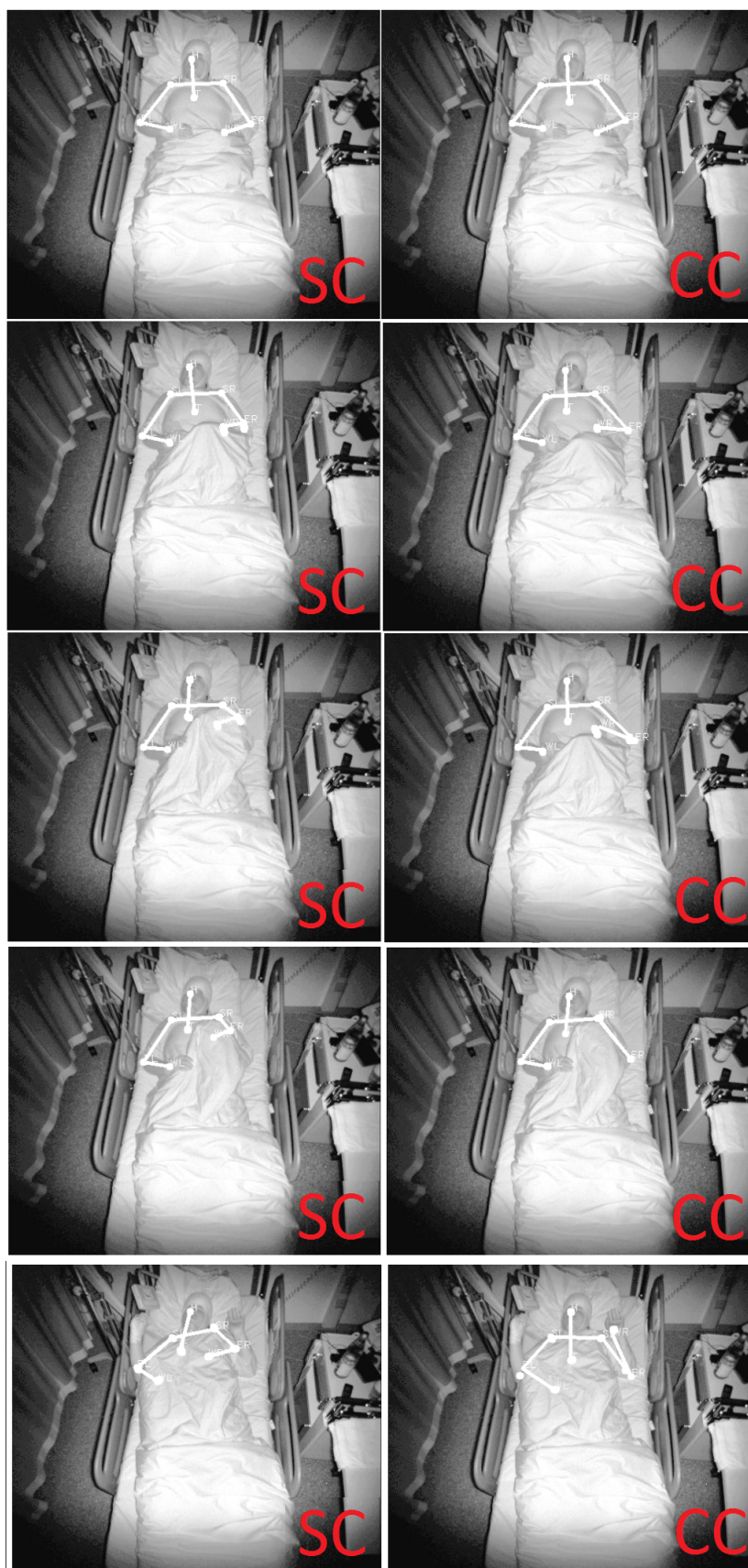


Figura 3.27: Sequência de *frames* relativamente ao método usado sem correção (SC) e com correção (CC)

Tabela 3.3: Percentagem de erro em cada articulação

Crise	C.		T.		O. E.		O. D.		C. E.		C. D.		P. E.		P. D.		Total	
	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y
run1	2,14	10,19	18,05	28,55	5,82	18,60	7,19	8,34	8,45	7,01	3,05	5,65	11,81	2,76	4,94	7,72	7,68	11,10
run10	1,08	3,13	0,37	1,97	2,40	5,60	0,33	3,31	1,86	14,76	0,86	3,67	4,99	75,60	0,30	1,52	13,71	13,69
run11	0,92	3,56	4,49	8,20	3,00	2,66	1,06	3,55	3,54	6,62	2,83	3,43	0,56	2,40	3,19	9,73	2,45	5,42
run12	1,64	0,43	3,95	4,37	1,93	2,37	3,67	2,74	3,45	4,88	4,15	6,30	1,20	7,61	5,63	8,62	4,15	4,66
run13	2,07	9,96	0,94	6,52	2,11	6,41	3,33	2,98	1,02	4,10	2,89	9,11	5,84	7,95	4,27	3,85	2,81	6,36
run16	2,80	3,01	1,46	4,72	2,29	6,81	0,54	1,34	3,33	4,73	6,70	2,13	4,92	1,23	8,68	4,61	3,84	3,57
run20	3,98	11,21	1,48	8,60	2,27	7,02	0,80	7,77	3,57	2,00	0,71	3,42	2,83	4,88	0,82	2,38	2,06	5,91
run21	4,22	10,49	9,45	17,11	16,32	16,74	3,41	11,68	10,24	34,19	16,30	58,30	4,64	41,41	8,42	33,16	9,13	27,88
<b>Total</b>	<b>2,35</b>	<b>6,50</b>	<b>5,02</b>	<b>10,01</b>	<b>4,52</b>	<b>8,27</b>	<b>2,54</b>	<b>5,21</b>	<b>4,43</b>	<b>9,78</b>	<b>4,69</b>	<b>11,50</b>	<b>4,60</b>	<b>17,98</b>	<b>4,53</b>	<b>8,95</b>	<b>5,73</b>	<b>9,82</b>

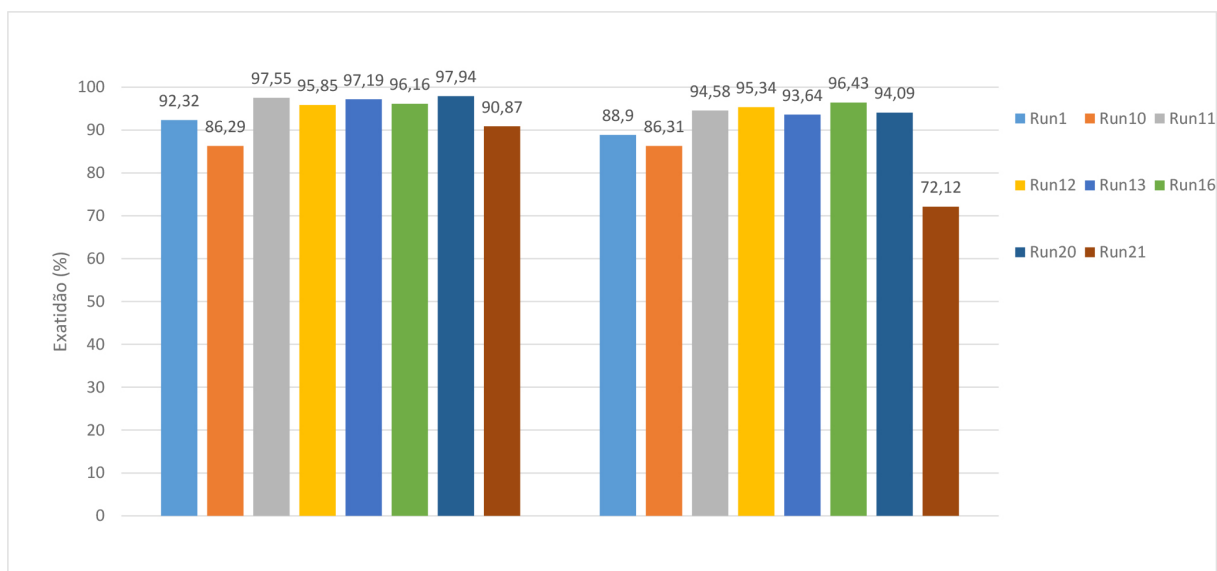


Figura 3.28: Exatidão em X e Y para cada crise avaliada

## Capítulo 4

# Interface: Sistema de Análise de Movimento 3D

Este capítulo irá endereçar as principais características da interface desenvolvida, assim como a informação relativa ao ambiente onde este foi desenvolvido (Microsoft Visual Studio C# - *Windows Presentation Foundation (WPF) Application*). Sendo este trabalho desenvolvido no seguimento de projetos anteriores, é também feita uma revisão dos mesmos.

### 4.1 NeuroKinect, KiT e KiMA

O conceito de usar a *Kinect* para avaliação de movimento em doenças neurológicas foi introduzido pelo Dr. João Paulo Cunha. O KiT (Kinect Tracker) e o KiMA (Kinect Motion Analyser) constituem o NeuroKinect, que avalia movimento com base na câmara RGB-D, nomeadamente *Microsoft Kinect*. Ambas as aplicações foram produzidas em C#. O KiT é usado para adquirir e guardar informação relativa à atividade humana, composto por cor, profundidade e esqueleto e o KiMA permite a manipulação dessa mesma informação. Esta interface desenvolvida vem complementar o KiMA, de forma a seguir corretamente os pontos do esqueleto, pois foi adequado ao ambiente clínico em questão, a adquirir e extrair dados relevantes como gráficos de velocidade e posições 3D dos pontos do esqueleto.

### 4.2 Requisitos

Esta interface permite a manipulação de vídeos de infravermelhos, anteriormente adquiridos pelo KiMA. Os principais requerimentos para esta interface são:

1. Visualização *offline* da informação previamente guardada pelo KiMA
2. Manipulação do vídeo (*play*, pausa, stop, escolha de *frame*)
3. Seleção das juntas a seguir

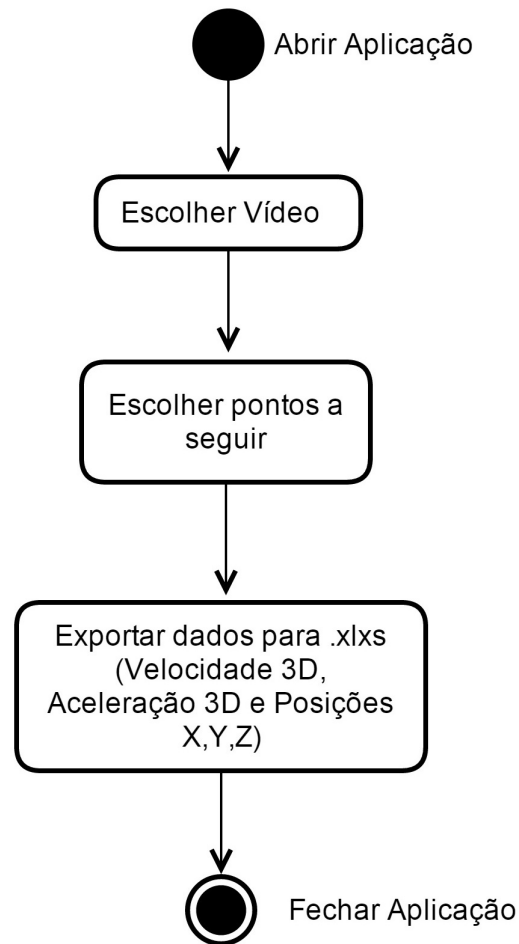


Figura 4.1: Diagrama de atividades UML da aplicação

4. Seleção e Visualização dos gráficos a mostrar
5. Visualização 3D dos pontos do esqueleto
6. Comentar vídeo
7. Guardar todas as informações mencionadas em ficheiro *Excel*

Na figura 4.2 pode ver-se as funcionalidades que a aplicação oferece, através de um diagrama de casos de uso, ao utilizador, que poderá ser um engenheiro ou até mesmo um profissional de saúde, dependendo da tarefa principal. As crises visualizadas terão de estar previamente extraídas através do KiMA, como foi referido anteriormente, e guardadas em formato de vídeo. Após o utilizador selecionar a crise, este poderá selecionar as articulações com a sequência definida (Cabeça, Torso, Ombro Esquerdo, Ombro Direito, Cotovelo Esquerdo, Cotovelo Direito, Pulso Esquerdo, Pulso Direito), sendo possível observar o esqueleto destas mesmas articulações ao longo do decorrer da crise. É permitido também assistir a visualização em 3D dessas mesmas articulações, assim



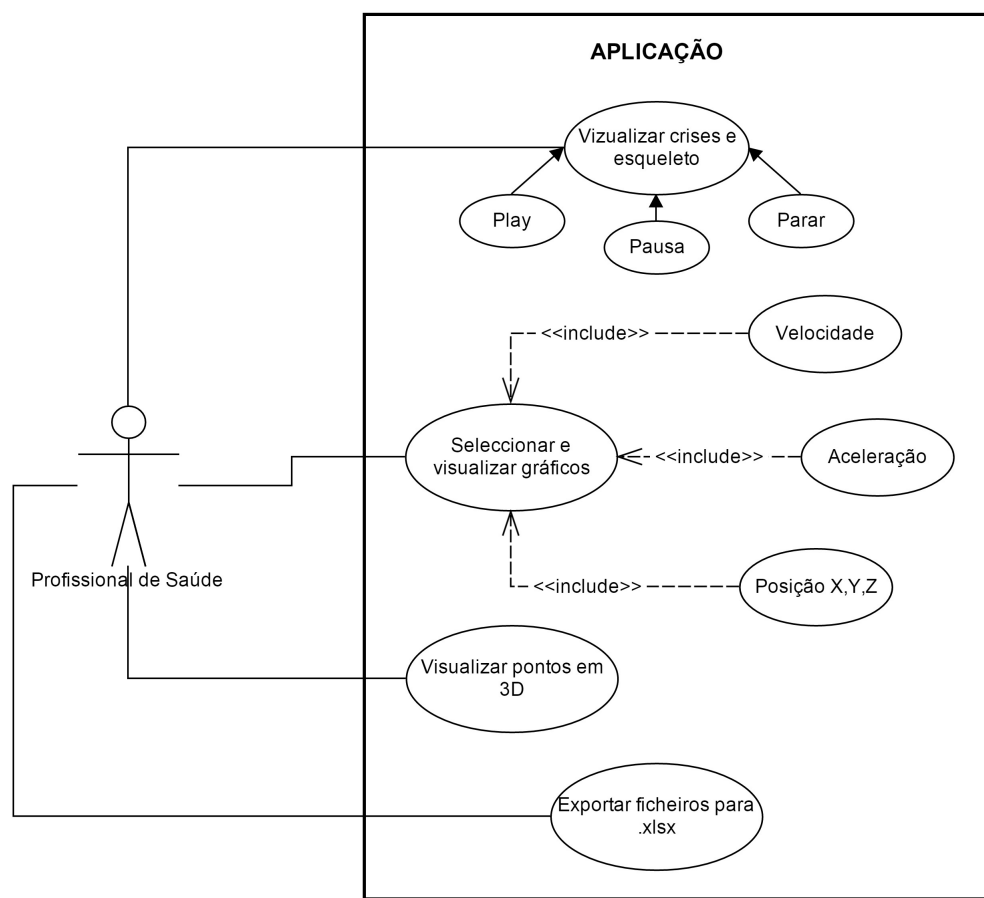


Figura 4.2: Diagrama de casos de uso da aplicação

como fazer a seleção dos gráficos das articulações e respetivas métricas (velocidade, aceleração ou posições X, Y, Z). Todas essas métricas serão gravadas num ficheiro *Excel* após o fim da crise ou antes, caso o utilizador assim o deseje.

### 4.3 Aplicação

A aplicação foi feita com o intuito de ser amigo do utilizador e de fácil utilização. Esta divide-se em 4 componentes:

1. Vídeo
2. Comentários
3. Gráficos
4. Pontos do esqueleto em três dimensões

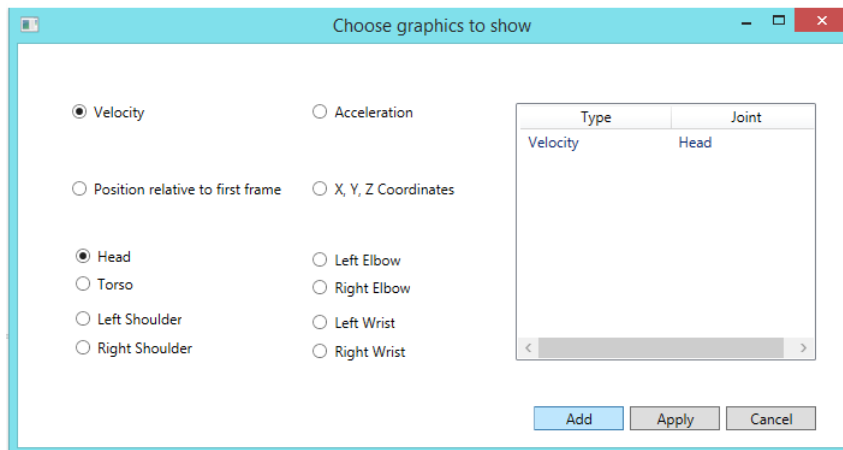


Figura 4.3: Janela de edição de gráficos a visualizar

como descrito na figura 4.4. É dado ao utilizador a opção de selecionar um novo vídeo, guardar todos os dados que foram adquiridos pela aplicação, bem como os comentários que fez, os pontos das juntas de interesse, velocidade e aceleração, através do menu, da figura 4.4 a vermelho. Pode também editar os gráficos que deseja ver na aplicação, como se pode verificar na figura 4.3. Adicionalmente, é também permitido ao utilizador parar, pausar, voltar a visualização do vídeo (figura 4.4 a azul) e selecionar o tempo em que deseja começar a visualizar o vídeo, através do controlo deslizante, que consecutivamente varia consoante o vídeo avança. Para uma melhor visualização dos pontos a 3D, (figura 4.4 a laranja), o utilizador pode rodar o eixo com o uso do rato e visualizar de diferentes ângulos e/ou aumentar ou diminuir a área de visualização ao premir + ou – no teclado, respetivamente.

É possível exportar, como foi referido anteriormente, todos estes dados para um ficheiro *Excel*, como no exemplo demonstrado na figura 4.5, para ser possível posterior análise e não ser necessário voltar a correr o algoritmo de seguimento dos membros. Para isso, basta selecionar "*File*" e de seguida "*Save*". Todos os valores são guardados, sem exceção. A estrutura deste ficheiro é dividida em Livros, em que cada junta representa um Livro no ficheiro de *Excel*. Para cada junta é então guardado os pontos  $X, Y$  em pixeis, os pontos  $X, Y, Z$  em coordenadas mundo, a velocidade e a aceleração. É também criado outro Livro para guardar os comentários efetuados.

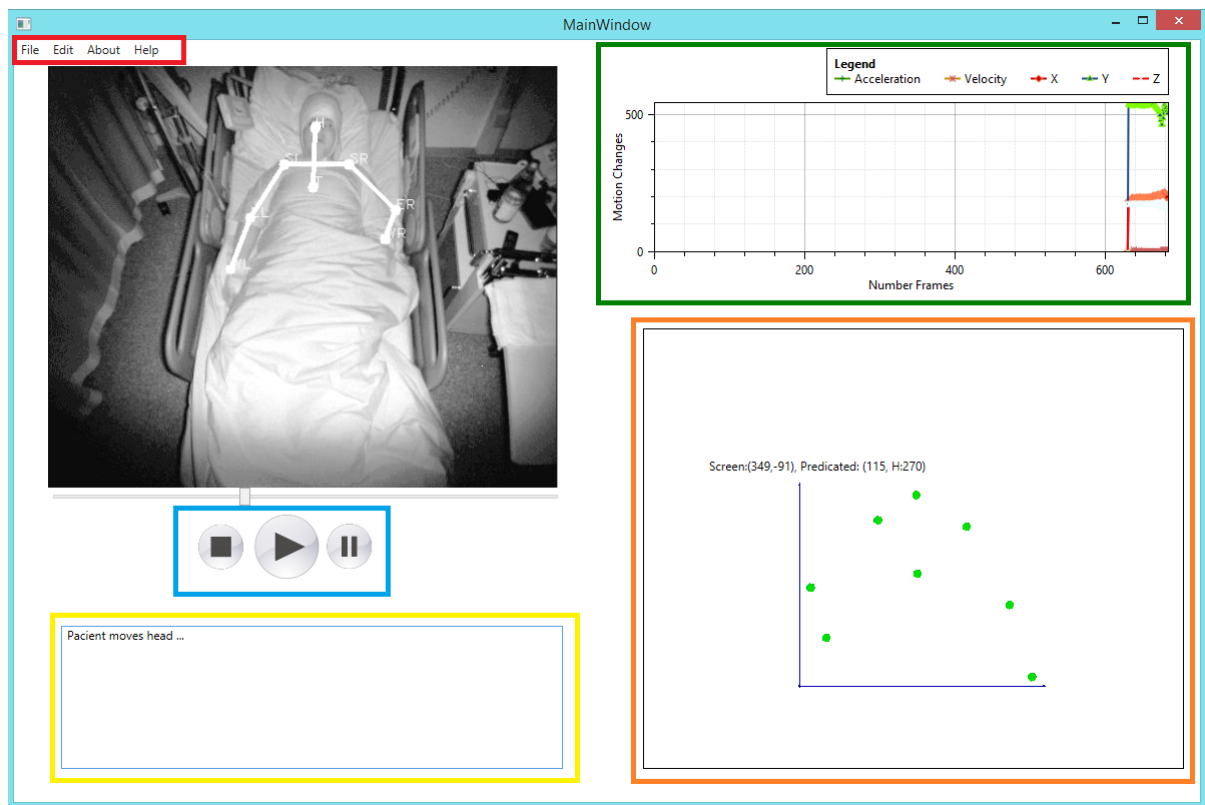


Figura 4.4: Janela principal da aplicação. Vermelho - menu que permite personalização. Azul - Manipulação do Vídeo. Verde - Visualização dos gráficos. Laranja - Visualização 3D do esqueleto. Amarelo - Comentários

Frames	X Cooridin	Y Cooridin	Z Cooridin	X Velocity	Y Velocity	Z Velocity	X Pixel	Y Pixel					
1	169,8559	526,408	17,31	0,196243	2,323914	0,02	281	63					
2	169,8559	526,408	17,31	0	0	0	281	63					
3	169,7578	526,1038	17,3	0	0	0	281	63					
4	170,0522	527,0162	17,33	0	0	0	281	63					
5	169,954	526,7121	17,32	0	0	0	281	63					
6	169,8559	526,408	17,31	0	0	0	281	63					
7	169,4634	525,1915	17,27	0	0	0	281	63					
8	169,6597	525,7998	17,29	0	0	0	281	63					
9	170,0522	527,0162	17,33	0	0	0	281	63					
10	169,5615	525,4957	17,28	0	0	0	281	63					
11	169,954	526,7121	17,32	0	0	0	281	63					
12	169,7578	526,1038	17,3	0	0	0	281	63					
13	169,7578	526,1038	17,3	0	0	0	281	63					
14	169,8559	526,408	17,31	0	0	0	281	63					
15	170,0522	527,0162	17,33	0	0	0	281	63					
16	169,954	526,7121	17,32	0	0	0	281	63					
17	169,8559	526,408	17,31	0	0	0	281	63					
18	169,7578	526,1038	17,3	0	0	0	281	63					
19	169,6597	525,7998	17,29	0	0	0	281	63					
20	170,1503	527,3203	17,34	0	0	0	281	63					
21	170,0522	527,0162	17,33	0	0	0	281	63					
22	169,954	526,7121	17,32	0	0	0	281	63					
23	169,7578	526,1038	17,3	0	0	0	281	63					
24	169,5615	525,4957	17,28	0	0	0	281	63					
25	169,954	526,7121	17,32	0	0	0	281	63					
26	169,7578	526,1038	17,3	0	0	0	281	63					
27	169,8559	526,408	17,31	0	0	0	281	63					
28	170,2484	527,6244	17,35	0	0	0	281	63					
29	169,5615	525,4957	17,28	0	0	0	281	63					

Figura 4.5: Exemplo do ficheiro guardado em Excel. Cada sheet apresenta os resultados adquiridos para cada articulação



## Capítulo 5

# Conclusões e Trabalho Futuro

Neste capítulo é apresentado uma revisão dos objetivos e o trabalho desenvolvido para atingir esses mesmo objetivos. Algumas sugestões de trabalho futuro também são fornecidas.

### 5.1 Revisão dos Objetivos

Esta dissertação teve como objetivo desenvolver um método capaz de quantificar movimentos de epilepsia através de um algoritmo semi-automático capaz de seguir membros e avaliá-los posteriormente com os resultados devolvidos, assim como o desenvolvimento de uma interface que facilitasse o uso deste mesmo método. Posteriormente, este serviria como avaliação pré-cirúrgica dos pacientes de epilepsia.

A interface consiste numa apresentação do vídeo na Unidade Motora de Epilepsia, gráficos relativos ao movimento efetuado pelos membros e a localização 3D desses mesmos membros. Este método serviria para dotar o uso do sensor Kinect v2 com um seguimento de membros mais adequado e preciso à situação em questão. Esta abordagem poderá fornecer outras visões ao profissional de saúde e conseqüentemente permite um maior um apoio ao diagnóstico.

### 5.2 Contribuições da Dissertação

Apesar do objetivo inicial da Microsoft Kinect v2 não ter sido aqui usado de forma convencional, esta foi de extrema importância neste projeto de forma a usufruir das capacidades e das funcionalidades que esta apresenta. Através da aquisição da imagem IV, que possibilita também a visualização à noite, em conjunto com o mapa de profundidade foi possível adequar o sensor a um cenário completamente irreverente e de forma precisa sem um grande custo computacional. Sem dúvida que a maior vantagem foi usar o sensor como uma forma de registrar movimento sem marcadores, apenas por visão computacional e a baixo custo.

Apesar da Kinect v2 dar um bom resultado para cenários mais comuns, esta não o consegue fazer com precisão nos cenários estudados, por diversas razões como o ângulo em que o sensor se encontra (encostado ao teto do quarto virado de frente para o paciente) ao contrário dos cenários

mais comuns onde o sensor está posicionado de frente para a pessoa. Também por este cenário ter diversos artefactos de imagem, como membros escondidos, debaixo dos lençóis ou até mesmo uma posição que não a frontal, como a posição fetal. Foi necessário criar um algoritmo capaz de mitigar estes problemas e adequar ao cenário específico e adquirir com precisão os pontos em 3D dos membros seguidos ao longo das *frames*. O método fornece uma simplificação do seguimento de membros até mesmo quando estes se encontram ocluídos (pelo lençol), uma das falhas que o sensor tinha. Espera-se que com este método se consiga fazer contribuições na Unidade Motora de Epilepsia e influenciar positivamente a forma como atualmente se processa os dados de Vídeo-EEG, sem criar mais tempo despendido aos profissionais de saúde.

Esta interface oferece a possibilidade de visualizar e exportar todos os valores adquiridos pelo sensor e consecutivamente pelo algoritmo. Apresenta opções nos dados a ser observados, como a seleção dos gráficos a visualizar, como as coordenadas da posição de algum membro, a velocidade e a aceleração e a exportação desses mesmos dados. O tempo despendido é também um fator importante neste processamento de dados, pois o processo manual é cansativo e requer diversas horas para as várias crises de todos os pacientes. Enquanto que ao marcar manualmente poderá demorar até horas, com o método semi-automático demora alguns minutos, por crise, sem requerer a vigilância do decorrer do processo e apenas visualizar os resultados posteriormente dados. A exportação dos dados reflete isso mesmo, a visualização dos dados *offline*, sem a necessidade de voltar a correr o algoritmo de seguimento de membros, diminuindo ainda mais o tempo despendido para avaliar uma crise. A interface foi desenvolvida em C# para facilitar o uso da mesma e ser facilmente portátil.

Numa perspectiva clínica é reconhecida a importância da avaliação dos padrões de movimento durante as crises epiléticas. Ao usar este método, este conceito apresenta maior significância pois apresenta resultados interessantes e significativos na análise de movimento. Apesar do algoritmo não ser em tempo-real, este consegue fornecer em poucos milissegundos o resultado da mudança do movimento e apenas apresenta um atraso maior quando necessita de correção, pois necessita de verificar diversos pontos do mapa de profundidade e convertê-los para coordenadas mundo para que corrijam efetivamente o ponto que possa ter ficado ocluído ou caso tenha ocorrido um movimento brusco em que a estimativa não conseguiu reconhecer com precisão. Sendo a distância euclidiana 3D entre membros uma característica invariante ao longo do tempo e com a precisão do sensor de adquirir a profundidade de cada pixel é possível a correção do membro para o ponto respetivo.

É possível afirmar que o método desenvolvido teve resultados positivos e adequou-se à problemática sem despende um custo muito elevado de computação. Apesar deste método não ter sido aplicado a muitas crises, apresentou grande potencialidade e poderá ser aplicado noutros para melhorar e fazer a adaptação necessária para casos mais desafiantes. A apresentação de resultados ajuda a perceber a importância que este método e interface podem ter e estimar o potencial noutras aplicações futuras.

### **5.3 Trabalho Futuro**

De forma a aumentar a confiança do método será necessário testar em mais crises, e até mesmo ser testado por um profissional de saúde para melhorar a interface e colocar à prova a utilidade do mesmo. Futuramente, mais opções e características serão acrescentadas à interface como:

- adaptar a interface às especificações do KiMA para ser automática a transição
- carregar um vídeo que foi anteriormente analisado
- acrescentar os membros inferiores ao seguimento
- adicionar outras métricas, como ângulos entre juntas.





# Referências

- [1] Yvonne M. Hart. Epidemiology, natural history and classification of epilepsy. *Medicine*, 40(9):471 – 476, 2012. Neurology: Part 2 of 3. URL: <http://www.sciencedirect.com/science/article/pii/S1357303912001296>, doi:<http://dx.doi.org/10.1016/j.mpmed.2012.06.002>.
- [2] Zhanjian Li, A.M. da Silva, e J.P.S. Cunha. Movement quantification in epileptic seizures: a new approach to video-eeeg analysis. *Biomedical Engineering, IEEE Transactions on*, 49(6):565–573, June 2002. doi:10.1109/TBME.2002.1001971.
- [3] J. K. Aggarwal e Q. Cai. Human motion analysis: A review. *Computer Vision and Image Understanding*, 73:428–440, 1999.
- [4] Daniele Regazzoni, Giordano de Vecchi, e Caterina Rizzi. {RGB} cams vs rgb-d sensors: Low cost motion capture technologies performances and limitations. *Journal of Manufacturing Systems*, 33(4):719 – 728, 2014. URL: <http://www.sciencedirect.com/science/article/pii/S0278612514000910>, doi:<http://dx.doi.org/10.1016/j.jmsy.2014.07.011>.
- [5] Laary Li. Texas Instruments, year = 2014, url = <http://www.ti.com/lit/wp/sloa190b/sloa190b.pdf>.
- [6] S. Foix, G. Alenya, e C. Torras. Lock-in time-of-flight (tof) cameras: A survey. *Sensors Journal, IEEE*, 11(9):1917–1926, Sept 2011. doi:10.1109/JSEN.2010.2101060.
- [7] Microsoft, 2014. URL: <http://www.microsoft.com/en-us/kinectforwindows/>.
- [8] Medlej Joumana. Human anatomy fundamentals flexibility and joint limitations, 2014. URL: <http://design.tutsplus.com/articles/human-anatomy-fundamentals-flexibility-and-joint-limitations--vector-25401/>.
- [9] E. Lachat, H. Macher, M.-A. Mittet, T. Landes, e P. Grussenmeyer. First Experiences with Kinect v2 Sensor for Close Range 3d Modelling. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, páginas 93–100, Fevereiro 2015. doi:10.5194/isprsarchives-XL-5-W4-93-2015.
- [10] Henning Stolze, Stephan Klebe, Christoph Baecker, Christiane Zechlin, Lars Friege, Sabine Pohle, e Günther Deuschl. Prevalence of gait disorders in hospitalized neurological patients. *Movement Disorders*, 20(1):89–94, 2005. URL: <http://dx.doi.org/10.1002/mds.20266>, doi:10.1002/mds.20266.
- [11] Common neurological disorders, year = 2013, url = <http://www.livestrong.com/article/98547-common-neurological-disorders/ixzz2RVGzYPX8>.

- [12] S. Hinde S. Palmer F. Beyer J. Minton-A. Marson U. Wieshmann N. Woolacott J. Burch e M. Soares. The clinical effectiveness and cost-effectiveness of technologies used to visualise the seizure focus in people with refractory epilepsy being considered for surgery: a systematic review and decision-analytical model. *Health Technol Assess*, 16(34):1 – 157, 2012.
- [13] Ivan Rektor, Steven C. Schachter, Shahar Arzy, Stavros J. Baloyannis, Carl Bazil, Milan Brázdil, Jerome Engel Jr., Gerhard Helmstaedter, Dale C. Hesdorffer, Marilyn Jones-Gotman, Ladislav Kesner, Vladimír Komárek, Günter Krämer, Ilo E. Leppik, Michael W. Mann, Marco Mula, Gail L. Risse, Guy W. Stoker, Dorothee G.A. Kasteleijn-Nolst Trenité, Michael Trimble, Ivana Tyrliková, e Amos D. Korczyn. Epilepsy, behavior, and art (epilepsy, brain, and mind, part 1). *Epilepsy Behavior*, 28(2):261 – 282, 2013. URL: <http://www.sciencedirect.com/science/article/pii/S1525505013001509>, doi:<http://dx.doi.org/10.1016/j.yebeh.2013.03.011>.
- [14] R. Rocamora, I. Mader, J. Zentner, e A. Schulze-Bonhage. Epilepsy surgery in patients with multiple cerebral cavernous malformations. *Seizure*, 18(4):241 – 245, 2009. URL: <http://www.sciencedirect.com/science/article/pii/S1059131108002318>, doi:<http://dx.doi.org/10.1016/j.seizure.2008.10.006>.
- [15] S. Benbadis. The differential diagnosis of epilepsy: A critical review. *Epilepsy Behavior*, 15(1):15 – 21, 2009. Management of Epilepsy: Hope and Hurdles Critical Reviews and Clinical Guidance. URL: <http://www.sciencedirect.com/science/article/pii/S152550500900081X>, doi:<http://dx.doi.org/10.1016/j.yebeh.2009.02.024>.
- [16] Nicola Specchio, Antonio Carotenuto, Marina Trivisano, Simona Cappelletti, Federico Vigeveno, e Lucia Fusco. Ictal yawning in a patient with drug-resistant focal epilepsy: Video/eeg documentation and review of literature reports. *Epilepsy Behavior*, 22(3):602 – 605, 2011. URL: <http://www.sciencedirect.com/science/article/pii/S152550501100477X>, doi:<http://dx.doi.org/10.1016/j.yebeh.2011.08.013>.
- [17] Gregory D. Cascino. Video-eeg monitoring in adults. *Epilepsia*, 43:80–93, 2002. URL: <http://dx.doi.org/10.1046/j.1528-1157.43.s.3.14.x>, doi:10.1046/j.1528-1157.43.s.3.14.x.
- [18] S Shorvon. The treatment of chronic epilepsy: a review of recent studies of clinical efficacy and side effects. *CURR OPIN NEUROL*, 20(2):159 – 163, June 2007.
- [19] B. K. P. Horn e B. G. Schunk. Determining optical flow. *Artificial Intelligence*, 17:185–204, 1981.
- [20] Martin Casdagli, Stephen Eubank, J. Doyne Farmer, e John Gibson. State space reconstruction in the presence of noise. *Phys. D*, 51(1-3):52–98, Setembro 1991. URL: [http://dx.doi.org/10.1016/0167-2789\(91\)90222-U](http://dx.doi.org/10.1016/0167-2789(91)90222-U), doi:10.1016/0167-2789(91)90222-U.
- [21] D.M Gavrilá. The visual analysis of human movement. *Comput. Vis. Image Underst.*, 73(1):82–98, Janeiro 1999. URL: <http://dx.doi.org/10.1006/cviu.1998.0716>, doi:10.1006/cviu.1998.0716.
- [22] J. O'Rourke e N.I. Badler. Model-based image analysis of human motion using constraint propagation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PAMI-2(6):522–536, Nov 1980. doi:10.1109/TPAMI.1980.6447699.

- [23] Softkinetic. Softkinetic, the interface is you. “ds311 datasheet, year = 2014, url = [http://www.softkinetic.com/Portals/0/Documents/PDF/WEB\\_20130527\\_SKD\\_S311\\_Datasheet\\_V3.0.pdf](http://www.softkinetic.com/Portals/0/Documents/PDF/WEB_20130527_SKD_S311_Datasheet_V3.0.pdf),.
- [24] Panasonic. Panasonic electric works. “d-imager 3d sensing technology, year = 2014, url = <http://www2.panasonic.biz/es/densetsu/device/3DImageSensor/en/product.html>,.
- [25] Asus. Asus “xtion pro live, year = 2014, url = <http://www.asus.com/Multimedia/Xtion>,.
- [26] Richard Caton. The Electric Currents of the Brain. *Br Med. J*, 1875.
- [27] Matthew Padiaditis, Manolis Tsiknakis, e Norbert Leitgeb. Vision-based motion detection, analysis and recognition of epileptic seizures—a systematic review. *Computer Methods and Programs in Biomedicine*, 108(3):1133 – 1148, 2012. URL: <http://www.sciencedirect.com/science/article/pii/S0169260712001873>, doi:<http://dx.doi.org/10.1016/j.cmpb.2012.08.005>.
- [28] Jan Rémi, João P. Silva Cunha, Christian Vollmar, Özgür Bilgin Topçuoğlu, Alexander Meier, Steffen Ulowetz, Pedro Beleza, e Soheyl Noachtar. Quantitative movement analysis differentiates focal seizures characterized by automatism. *Epilepsy Behavior*, 20(4):642 – 647, 2011. URL: <http://www.sciencedirect.com/science/article/pii/S1525505011000242>, doi:<http://dx.doi.org/10.1016/j.yebeh.2011.01.020>.
- [29] João Paulo Silva Cunha, Luís M. Paula, Virgílio F. Bento, Candas Bilgin, Eduardo Dias, e Soheyl Noachtar. Movement quantification in epileptic seizures: A feasibility study for a new 3d approach. *Medical Engineering Physics*, 34(7):938 – 945, 2012. URL: <http://www.sciencedirect.com/science/article/pii/S1350453311002785>, doi:<http://dx.doi.org/10.1016/j.medengphy.2011.10.013>.
- [30] Xuhong Yang, Lei Chen, Yonghong Liu, Dong Zeng, Yufeng Tang, Bo Yan, Xu Lin, Ling Liu, Hongru Xu, e Dong Zhou. Motor trajectories in automatism and their quantitative analysis. *Epilepsy Research*, 83(2–3):97 – 102, 2009. URL: <http://www.sciencedirect.com/science/article/pii/S0920121108002854>, doi:<http://dx.doi.org/10.1016/j.epilepsyres.2008.09.010>.
- [31] S. Bonnet, P. Jallon, A. Bourgerette, M. Antonakios, R. Guillemaud, Y. Caritu, G. Becq, P. Kahane, P. Chapat, B. Thomas-Vialettes, F. Thomas-Vialettes, D. Gerbi, e D. Ejnes. An Ethernet motion-sensor based alarm system for epilepsy monitoring. *Irbm*, 32:155–157, 2011. doi:10.1016/j.irbm.2011.01.021.
- [32] Guillaume Becq, Stéphane Bonnet, Lorella Minotti, Michel Antonakios, Régis Guillemaud, e Philippe Kahane. Classification of epileptic motor manifestations using inertial and magnetic sensors. *Comput. Biol. Med.*, 41(1):46–55, Janeiro 2011. URL: <http://dx.doi.org/10.1016/j.compbiomed.2010.11.005>, doi:10.1016/j.compbiomed.2010.11.005.
- [33] H.-L. E. B. M. D. W. S. C. a. Y.-L. N. Haiping Lu. Markerless video analysis for movement quantification in pediatric epilepsy monitoring. *33rd Annual International Conference of the IEEE EMBS*, 2011.
- [34] Kris Cuppens, Lieven Lagae, Berten Ceulemans, Sabine Van Huffel, e Bart Vanrumste. Automatic video detection of body movement during sleep based on optical flow in pediatric patients with epilepsy. *Medical Biological Engineering Computing*, 48(9):923–931, 2010. URL: <http://dx.doi.org/10.1007/s11517-010-0648-4>, doi:10.1007/s11517-010-0648-4.

- [35] Rafael C. Gonzalez e Richard E. Woods. *Digital Image Processing*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2nd edição, 2001.
- [36] EmguCV, 2015. URL: [http://www.emgu.com/wiki/index.php/Main\\_Page](http://www.emgu.com/wiki/index.php/Main_Page).
- [37] AForge.NET, 2015. URL: <http://www.aforge.net.com/>.
- [38] B. D. Lucas e T. Kanade. An iterative image registration technique with an application to stereo vision. *Proceedings of Imaging Understanding Workshop*, pages 121–130, 1981.
- [39] B.K.P. Horn e B.G. Schunck. Determining optical flow. *Artificial Intelligence*, vol 17, pp 185–203, 1981.
- [40] Karla Brkic, Srdan Rasic, Axel Pinz, Sinisa Segvic, e Zoran Kalafatic. Combining spatio-temporal appearance descriptors and optical flow for human action recognition in video data. *CoRR*, abs/1310.0308, 2013. URL: <http://dblp.uni-trier.de/db/journals/corr/corr1310.html#BrkicRPSK13>.
- [41] Gunnar Farneback. *Two-Frame Motion Estimation Based on Polynomial Expansion*. 2003. doi:10.1007/3-540-45103-X50.
- [42] Gunnar Farneback. *Polynomial Expansion for Orientation and Motion Estimation*. Tese de doutoramento, Linköping University, Sweden, SE-581 83 Linköping, Sweden, 2002. Dissertation No 790, ISBN 91-7373-475-6.
- [43] Gunnar Farneback. Motion-based segmentation of image sequences using orientation tensors, 1997.
- [44] Jing Huang e Suya You. Point cloud matching based on 3d self-similarity. Em *Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2012 IEEE Computer Society Conference on, páginas 41–48, June 2012. doi:10.1109/CVPRW.2012.6238913.
- [45] I.P. Herman. *Physics of the Human Body*. Biological and Medical Physics, Biomedical Engineering. Springer, 2007. URL: <https://books.google.pt/books?id=vtubxNaSAdAC>.
- [46] Hugo Choupina. Neurokinect: Kinect-based system for motion analysis and quantification in neurological diseases,. Tese de mestrado, FEUP, October 2014.
- [47] Maxtraq 2d, 2014. URL: <http://www.innovision-systems.com/Products/MaxTraq2D.html>.