

Leandro Weige Dias

**Análise Comparativa entre a Primeira e a Segunda Versão do Kinect na Biometria
do Caminhar**

Dissertação apresentada ao Programa de Pós-Graduação em Computação da Universidade Federal de Pelotas, como requisito parcial à obtenção do título de Mestre em Ciência da Computação

Orientador: Prof. Dr. Ricardo Matsumura de Araújo

Pelotas, 2017

Universidade Federal de Pelotas / Sistema de Bibliotecas
Catalogação na Publicação

D541a Dias, Leandro Weige

Análise comparativa entre a primeira e a segunda versão do kinect na biometria do caminhar / Leandro Weige Dias ; Ricardo Matsumura de Araújo, orientador. — Pelotas, 2017.

114 f. : il.

Dissertação (Mestrado) — Programa de Pós-Graduação em Computação, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas, 2017.

1. Aprendizado de máquina. 2. Biometria. 3. Reconhecimento do caminhar. 4. Kinect. I. Araújo, Ricardo Matsumura de, orient. II. Título.

CDD : 005

AGRADECIMENTOS

Agradeço primeiramente a Deus, pois tudo o que eu tenho eu devo a Ele, meu Criador e Salvador, Àquele que me deu forças e coragem durante todos os momentos de minha vida. Também não posso deixar de agradecer ao meu orientador, Prof. Ricardo Matsumura de Araújo, pela sua atenção, dedicação, disponibilidade de estar sempre me auxiliando em minhas dúvidas, me motivando a prosseguir e me ensinando com zelo o que necessitei aprender em todo o processo de desenvolvimento deste trabalho. Também agradeço aos professores do PPGC, pois com sua excelência no ensino transmitiram-me conhecimentos essenciais para realizar esta dissertação e me desenvolver academicamente.

Quero agradecer especialmente a minha amada esposa, Luiza Fagundes Dias, pois ela de forma especial e carinhosa me motivou, deu forças e me apoiou em todos os momentos. Ela é parte essencial desta conquista, acreditou em mim e sempre esteve ao meu lado me auxiliando a todo o tempo. Te amo de todo o meu coração!

Agradeço também aos meus pais, pois sem a educação e o apoio que me deram nunca teria chegado até aqui. Eles também são parte indispensável desta conquista, o que eu sou é resultado da dedicação deles em formar e educar a minha vida.

Também desejo agradecer aos meus filhos, Fabrício e Lara, que me foram de motivação para seguir adiante.

Por fim, agradeço a UFPel, pois através dela pude me tornar servidor e aluno do mestrado no PPGC.

*Grandes coisas fez o Senhor por nós,
por isso estamos alegres.*
— SALMOS 126.3

RESUMO

DIAS, Leandro Weige. **Análise Comparativa entre a Primeira e a Segunda Versão do Kinect na Biometria do Caminhar**. 2017. 114 f. Dissertação (Mestrado em Ciência da Computação) – Programa de Pós-Graduação em Computação, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas, Pelotas, 2017.

Este trabalho tem por objetivo central trazer uma comparação entre as duas versões do sensor Kinect em relação a identificação biométrica através do caminhar humano. Para realizar a comparação entre os sensores, foram implementadas quatro metodologias diferentes de extração de parâmetros do caminhar humano, sendo analisados tanto atributos antropométricos, cinemáticos como espaço-temporais. Uma análise individual de cada atributo das propostas de caracterização do caminhar foi realizada para identificar os predicados que traziam maior contribuição na identificação dos indivíduos. Com a finalidade de se criar uma base de exemplos para a extração dos parâmetros do caminhar, foram capturados 50 indivíduos através dos sensores Kinect e aplicados algoritmos de aprendizado de máquina para a classificação das pessoas. Os resultados obtidos mostraram uma acurácia superior do Kinect One na maioria das metodologias, justificando a sua utilização em relação ao Kinect 360. Neste trabalho também foram realizados como experimentos adicionais a comparação entre os sensores com relação a identificação do gênero do indivíduo e a análise da dependência de pose.

Palavras-Chave: aprendizado de máquina; biometria; reconhecimento do caminhar; kinect

ABSTRACT

DIAS, Leandro Weige. **Comparative Analysys of Kinect's First and Second Version on Biometrics of Gait**. 2017. 114 f. Dissertação (Mestrado em Ciência da Computação) – Programa de Pós-Graduação em Computação, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas, Pelotas, 2017.

Machine Learning, Biometrics, Gait Recognition, Kinect

The main goal of this thesis is to bring a comparative between the two versions of the Kinect sensor in biometric identification through human gait. To compare the sensors, four different methods of human gait parameters extraction were implemented, being analyzed both anthropometric, kinematic and spatiotemporal attributes. An individual analysis of each attribute of the gait characterization methodologies was performed to identify the attributes that bring the greatest contribution in the identification of individuals. In order to create a base for extracting gait parameters, 50 individuals were captured through the Kinect sensors and machine learning algorithms were applied to classify the people. The results obtained show a superiority of Kinect One in most of the methodologies, justifying its use in relation to Kinect 360. In this work, a mapping between the sensors in relation to the identification of the individual's gender and an analysis of the pose dependency were also performed as additional experiments.

Keywords: machine learning; biometrics; gait recognition; kinect

LISTA DE FIGURAS

Figura 1	Ilustração do século 19 mostrando como realizar as medidas antropométricas de Alphonse. Fonte: YESHION (2014).	15
Figura 2	Sete componentes que formam a base para o caminhar humano. Fonte: VAUGHAN; DAVID; O'CONNOR (1999).	18
Figura 3	Ciclo completo. Fonte: VAUGHAN; DAVID; O'CONNOR (1999).	19
Figura 4	Ciclo completo do caminhar humano. Fonte: Adaptado de VAUGHAN; DAVID; O'CONNOR (1999).	19
Figura 5	Fases do ciclo do caminhar. Fonte: Adaptado por ANDERSSON (2014) de YAM; NIXON; CARTER (2004) e (YAM C. NIXON; CARTER, 2003). .	20
Figura 6	Rotação pélvica. Fonte: NETO (2015).	21
Figura 7	Inclinação da pélvis. Fonte: NETO (2015).	22
Figura 8	Deslocamento lateral da pélvis. Fonte: NETO (2015).	22
Figura 9	Duas teorias do caminhar. Fonte: SOUZA (2009).	23
Figura 10	Componentes de uma análise baseada em modelo. Fonte: YAM; NIXON (2015).	24
Figura 11	Figuras de Movimento. Fonte: Adaptado de YOO; NIXON (2011)	25
Figura 12	Figuras de Movimento. Fonte: Adaptado de YOO; NIXON (2011)	25
Figura 13	Ângulo de captura do método de identificação de Benabdelkader. Fonte: BENABDELKADER; CUTLER; DAVIS (2015).	27
Figura 14	Distâncias do corpo utilizadas como parâmetros estáticos. Fonte: BOBICK; JOHNSON (2001).	28
Figura 15	Trajectoria semi-circular percorrida pela pessoa. Fonte: Adaptado de ANDERSSON (2014).	29
Figura 16	Sequência de esqueletos 2D. Fonte: Adaptado de YOO JANG-HEE; MOON; HWANG DOOSUNG; NIXON (2008).	29
Figura 17	Usuário utilizando sensores de captura. Fonte: Adaptado de TANAWONG-SUW R.; BOBICK (2015).	30
Figura 18	Fluxograma do algoritmo de Wang. Fonte: Adaptado de WANG et al. (2005).	32
Figura 19	Fluxograma da metodologia de Balazia. Fonte: Adaptado de BALAZIA (2013).	32
Figura 20	Dimensões obtidas do indivíduo e a sua segmentação. Na esquerda são ilustradas a altura e largura da pessoa e na direita é apresentada a divisão em oito partes. Fonte: Adaptado de YEOH et al. (2011).	34
Figura 21	A silhueta típica de uma pessoa e suas sub-partes. Fonte: Adaptado de VENKAT; WILDE (2010).	35

Figura 22	GEIs extraídos de múltiplos ângulos. Pixels mais claros representam maior ocupação no ciclo do caminhar. Fonte: SIVAPALAN et al. (2011).	36
Figura 23	Na esquerda está um exemplo de um modelo de reconstrução de voxel, na imagem do meio é apresentado um GEV e na direita está uma representação cortada do GEV exibida para melhor ilustrar as densidades internas. Fonte: SIVAPALAN et al. (2011).	37
Figura 24	Demonstrativos dos ângulos de captura da base <i>CMU MoBo Database</i> . Fonte: SIVAPALAN et al. (2011).	37
Figura 25	Retângulo demarcando a silhueta da pessoa. Fonte: GUILLEN et al. (2009a).	38
Figura 26	Imagens gaussianas sendo transformadas em GHGI. Fonte: ARORA et al. (2015a).	39
Figura 27	Ângulos Euler utilizados para identificar o indivíduo. Fonte: (KASTANI-OTIS et al., 2015).	42
Figura 28	Ilustração do funcionamento do KNN. Fonte: (PETERSON, 2009).	51
Figura 29	Árvore de decisão para jogar tênis. Fonte: (GFBIOINFO, 2017).	53
Figura 30	Relações dos elementos da árvore de decisão para jogar tênis. Fonte: (GFBIOINFO, 2017).	53
Figura 31	Um neurônio artificial. Fonte: (GERSHENSON, 2015).	54
Figura 32	Aparência do Kinect 360. Fonte: (CNET, 2017).	55
Figura 33	Juntas mapeadas pelo Kinect 360. Fonte: (MICROSOFT, 2017b).	56
Figura 34	Aparência do Kinect One. Fonte: (XBOX, 2015).	57
Figura 35	Pontos mapeados pelos sensores Kinect. Fonte: (MASTER, 2017).	58
Figura 36	Fluxograma da metodologia para identificação dos indivíduos.	60
Figura 37	Posição dos sensores Kinect na captura dos dados.	60
Figura 38	Trajetos percorridos pelas pessoas nas capturas.	61
Figura 39	Trajetos das capturas.	62
Figura 40	Gênero das pessoas capturadas.	63
Figura 41	Altura das pessoas capturadas.	64
Figura 42	Peso das pessoas capturadas.	64
Figura 43	Idade das pessoas capturadas.	64
Figura 44	Em (a) é apresentado os segmentos que foram utilizados, em (b) é descrito os ângulos destes segmentos e em (c) é apresentado o ângulo de abertura do pé. (ANDERSSON, 2014).	67
Figura 45	Fases de extensão (vales) e flexão (picos) descritos pelos ângulos do joelho direito durante uma caminhada capturada (ANDERSSON, 2014).	68
Figura 46	Adaptação da metodologia de (ARORA et al., 2015a).	70
Figura 47	Área do retângulo (GUILLEN et al., 2009a).	70
Figura 48	Ângulo diagonal do retângulo (GUILLEN et al., 2009a).	71
Figura 49	Análise dos atributos antropométricos da metodologia de ANDERSSON (2014).	78
Figura 50	Comparativo dos sensores utilizando a metodologia de ANDERSSON (2014).	82
Figura 51	Resultados das Figuras de Movimento.	84
Figura 52	Resultados da Caixa Delimitadora.	86
Figura 53	Comparativo dos sensores na implementação da metodologia da Divisão em Segmentos.	88
Figura 54	Resultado da aplicação do conjunto de atributos formado pelas propostas de maior acurácia.	90

Figura 55 Comparativo do conjunto que seleciona atributos de todas as metodologias. 92

Figura 56 Comparativo entre os resultados. 93

Figura 57 Comparativo dos resultados da análise de dependência de pose. 97

LISTA DE TABELAS

Tabela 1	Resultado na identificação do mal de Parkinson	45
Tabela 2	Ângulos das capturas realizadas através do Kinect	47
Tabela 3	Algoritmos que obtiveram melhores resultados (* a última linha relaciona trabalhos que utilizaram algoritmos próprios de classificação.)	47
Tabela 4	Versão do Kinect utilizada nos trabalhos	48
Tabela 5	Resumo dos trabalhos na identificação do caminhar	49
Tabela 6	Resumo dos trabalhos na análise da saúde	50
Tabela 7	Exemplo de dados para árvore de decisão (GFBIOINFO, 2017).	52
Tabela 8	Implementação dos algoritmos de aprendizado de máquina no WEKA	72
Tabela 9	Resultado utilizando a metodologia de Andersson com Kinect 360	74
Tabela 10	Resultado utilizando a metodologia de Andersson com Kinect One	75
Tabela 11	Atributos antropométricos com Kinect 360	75
Tabela 12	Atributos antropométricos com Kinect One	75
Tabela 13	Impacto de cada atributo antropométrico no Kinect 360	76
Tabela 14	Impacto de cada atributo antropométrico no Kinect One	77
Tabela 15	Atributos cinemáticos com Kinect 360	78
Tabela 16	Atributos cinemáticos com Kinect One	79
Tabela 17	Impacto de cada atributo cinemático no Kinect 360	79
Tabela 18	Impacto de cada atributo cinemático no Kinect One	80
Tabela 19	Atributos espaço-temporais com Kinect One	80
Tabela 20	Atributos espaço-temporais com Kinect 360	80
Tabela 21	Impacto de cada atributo espaço-temporal com Kinect 360	81
Tabela 22	Impacto de cada atributo espaço-temporal com Kinect One	81
Tabela 23	Diferença entre os sensores na aplicação da proposta de ANDERSSON (2014)	82
Tabela 24	Resultado utilizando a metodologia das Figuras de Movimento com Kinect 360	83
Tabela 25	Resultado utilizando a metodologia das Figuras de Movimento com Kinect One	83
Tabela 26	Analisando o impacto de cada atributo das Figuras de Movimento no Kinect 360	83
Tabela 27	Analisando o impacto de cada atributo das Figuras de Movimento no Kinect One	84
Tabela 28	Resultado utilizando a metodologia da Caixa Delimitadora com Kinect 360	85
Tabela 29	Resultado utilizando a metodologia da Caixa Delimitadora com Kinect One	85
Tabela 30	Analisando o impacto de cada atributo da Caixa Delimitadora no Kinect 360	86

Tabela 31	Analisando o impacto de cada atributo da Caixa Delimitadora no Kinect One	86
Tabela 32	Resultado utilizando Sistema de Autenticação de Guillen	87
Tabela 33	Resultado utilizando a metodologia da Divisão em Segmentos com Kinect 360	87
Tabela 34	Resultado utilizando a metodologia da Divisão em Segmentos com Kinect One	88
Tabela 35	Resultado da combinação das propostas de melhor acurácia no Kinect 360	89
Tabela 36	Resultado da combinação das propostas de melhor acurácia no Kinect One	89
Tabela 37	Atributos removidos	91
Tabela 38	Resultado com Kinect 360	91
Tabela 39	Resultado com Kinect One	91
Tabela 40	Diferença nos resultados entre os sensores	92
Tabela 41	Análise da dependência de pose utilizando a proposta de ANDERSSON (2014) com Kinect 360	94
Tabela 42	Análise da dependência de pose utilizando a proposta de ANDERSSON (2014) com Kinect One	94
Tabela 43	Análise da dependência de pose utilizando a proposta de Figuras de Movimento com Kinect 360	95
Tabela 44	Análise da dependência de pose utilizando a proposta de Figuras de Movimento com Kinect One	95
Tabela 45	Análise da dependência de pose utilizando a proposta da Caixa Delimitadora com Kinect 360	95
Tabela 46	Análise da dependência de pose utilizando a proposta da Caixa Delimitadora com Kinect One	96
Tabela 47	Análise da dependência de pose utilizando a proposta de Divisão em segmentos com Kinect 360	96
Tabela 48	Análise da dependência de pose utilizando a proposta de Divisão em segmentos com Kinect One	96

LISTA DE ABREVIATURAS E SIGLAS

3D	Três Dimensões
CBIR	Content-Based Image Retrieval
CM	Centro de Massa
DNA	Deoxyribonucleic Acid
DTW	Dynamic Time Warping
EUA	Estados Unidos da América
FMR	False Match Rate
FPS	Frames per second
GAR	Genuine Acceptance Rate
GEI	Gait Energy Image
GEV	Gait Energy Volume
GHGI	Gradient Histogram Gaussian Image
HOG	Histogram of Oriented Gradients
ICC	Instâncias Corretamente Classificadas
JRA	Joint Relative Angle
JRD	Joint Relative Distance
KNN	K-nearest Neighbor
LS	Linguagem de Sinais
MLP	Multilayer Perceptron
PDV	Pose Depth Volume
RGB	Red, Green, Blue
SVM	Support Vector Machine
WEKA	Waikato Environment for Knowledge Analysis
WW	Wald-Wolfowitz
XBOX	Video-game Microsoft Xbox

SUMÁRIO

1	INTRODUÇÃO	14
2	FUNDAMENTAÇÃO TEÓRICA E ESTADO DA ARTE	18
2.1	O Caminhar Humano	18
2.2	O Ciclo do Caminhar	19
2.2.1	Abordagem em cinco eventos	19
2.2.2	Componentes cinemáticos	20
2.2.3	Teoria dos seis determinantes da marcha	21
2.2.4	Teoria do pêndulo invertido	22
2.3	Análise e Modelagem do Caminhar Humano	23
2.3.1	Análise baseada em modelo	23
2.3.2	Análise livre de modelo	34
2.4	Pesquisas com o Kinect na Análise Biométrica	40
2.4.1	Análise biométrica através do caminhar	40
2.5	Algoritmos de Aprendizado de Máquina	50
2.5.1	K-nearest Neighbor	50
2.5.2	Random Forest	51
2.5.3	Rede Bayesiana	53
2.5.4	Rede Neural Artificial	54
2.6	MICROSOFT KINECT	55
2.6.1	Kinect do Xbox 360	55
2.6.2	Kinect do Xbox One	57
3	OBJETIVOS E METODOLOGIA	59
3.1	Objetivos	59
3.2	Metodologia	59
3.2.1	Captura	60
3.2.2	Extração dos atributos	65
3.2.3	Classificação	71
3.2.4	Método de validação	72
3.2.5	Protocolo de captura	73
4	RESULTADOS	74
4.1	Proposta de Andersson	74
4.2	Proposta das Figuras de Movimento	82
4.3	Proposta da Caixa Delimitadora	85
4.3.1	Aplicação do sistema de autenticação de Guillen	87
4.4	Proposta da Divisão em Segmentos	87

4.5	Combinação de atributos	88
4.5.1	Combinação das propostas de maior acurácia	89
4.5.2	Combinação dos atributos que aumentam a acurácia	90
4.6	Análise dos Resultados	92
4.7	Análise da dependência de pose	94
4.7.1	Proposta de Andersson	94
4.7.2	Proposta de Figuras de Movimento	95
4.7.3	Proposta da Caixa Delimitadora	95
4.7.4	Proposta de Divisão em Segmentos	96
4.7.5	Análise dos resultados da dependência de pose	96
5	CONCLUSÃO	98
	REFERÊNCIAS	100
	APÊNDICE A ORGANIZAÇÃO DOS DADOS	109
A.1	Arquivo contendo a delimitação dos trajetos	109
A.2	Arquivo contendo os dados das capturas	110
	APÊNDICE B PROTOCOLO DE CAPTURA	112

1 INTRODUÇÃO

O reconhecimento de pessoas é importante no cotidiano do ser humano, pois através dele identificamos amigos, familiares, conhecidos, diversas pessoas que de alguma forma tiveram contato conosco e da qual nos lembramos.

Porém, embora sejamos capazes de reconhecer pessoas da qual tivemos contato, foi necessário criar-se métodos que assegurassem a identidade dos indivíduos. Os primeiros processos de reconhecimento de pessoas tinham objetivo civil, como identificar escravos, e posteriormente a sociedade sentiu a necessidade de criar meios de identificar criminosos, para atribuir responsabilidades e penalizações (PASQUALI; ARAÚJO, 2017).

Um dos meios utilizados para identificar pessoas é pelas suas medidas corporais, esta ação é conhecida como antropometria. Segundo PETROSKI (1995), a origem da antropometria está nas artes, inspirada na assimetria, harmonia e filosofia pitagórica, sendo que o seu princípio está demarcado na antiga civilização da Índia, Egito e Grécia, por meio do uso de dimensões de partes do corpo como padrão de medida. Esta metodologia foi utilizada na tentativa de estabelecer o perfil das proporções do corpo humano (DELGADO, 2015).

O inventor do termo antropometria, segundo BEUNEN (2009), foi um médico alemão, chamado Johann Sigismund Elsholtz (1623-1688). A tese de Johann em 1654 contém uma ilustração do instrumento que ele desenvolveu, o qual possuía uma haste transversal (régua) que movia para cima e para baixo uma haste vertical e que foi provavelmente um dos primeiros instrumentos construídos com o objetivo de medir as extremidades do corpo humano (MUEHLENBEIN, 2015).

MUEHLENBEIN (2015) também nos diz que os instrumentos de medidas antropométricas tiveram um avanço especial no campo da auxologia (estudo do crescimento humano), destacando o trabalho de Count Philibert Guéneau de Montbeillard como sendo o primeiro estudo longitudinal do crescimento humano. Em seu trabalho, Count Guéneau mediu seu filho a cada 6 meses, entre os anos de 1759 e 1777.

Em 1858, William Herschel, trabalhando no serviço civil da Índia, gravou a impressão da mão de cada trabalhador atrás de seus contratos para poder identificá-los. Esta foi a primeira captura sistemática da impressão de mãos para propósitos de identificação (UPDATE, 2015).

No ano de 1870, Alphonse Bertillon criou o sistema para identificar indivíduos baseados

em medidas antropométricas detalhadas (UPDATE, 2015). Segundo RUIZ-GARZÓN (2013), o sistema biométrico de Alphonse foi o primeiro método científico utilizado pela polícia com o objetivo de identificar criminosos. Gabriel nos diz que este método é baseado em 11 medidas corporais, cor do olhos, cabelos e pés, utilizando também fotografias de frente e de perfil dos criminosos e que através deste método foi possível, no ano de 1884, identificar 241 criminosos reincidentes, sendo adotado pela polícia britânica, francesa e norte-americana.

Alphonse fazia a medição de cada indivíduo 3 vezes e registrava a média delas em cartões. Foi determinado que a probabilidade de duas pessoas terem as mesmas medidas era de 1:4, ou seja, as chances de duas pessoas terem as mesmas medidas das mãos e dos pés, por exemplo, era de 1:4x4. A probabilidade de duas pessoas terem as 11 medidas iguais era de 1:411 (YESHION, 2014).

Para cada prisioneiro cadastrado, eram criados dois cartões, um ficava na prisão e uma cópia dele ia para o escritório Albany para ser indexado. O cadastro dos cartões era dividido dentro de 243 categorias, e o sistema começou a enfrentar problemas devido ao crescimento da quantidade cartões.

Os indexadores do Bertillon de Nova Iorque começaram a ter dentro das 243 categorias mais de 200 cartões, fazendo com que o tempo de pesquisa passasse de minutos para horas. Outro problema enfrentado foi a imprecisão das medidas registradas, que entre outros fatores, foi causado pela inexperiência ou incompetência dos examinadores (DCJS, 2015). Este método foi utilizado durante aproximadamente 20 anos, sendo substituído pela identificação através da impressão digital, desenvolvida por meio de um estudo de Francis Galton no ano de 1892 (UPDATE, 2015).

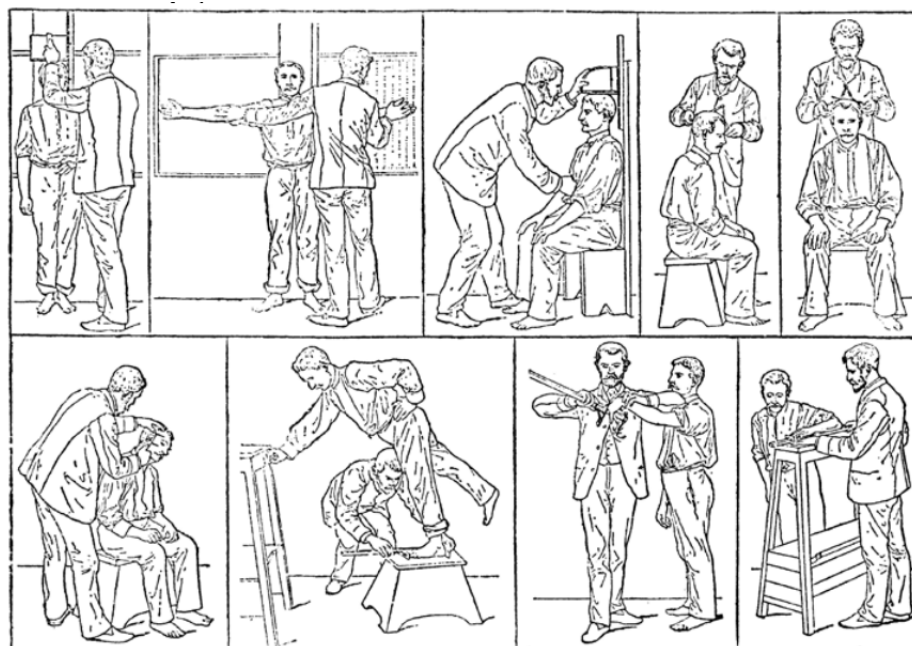


Figura 1 – Ilustração do século 19 mostrando como realizar as medidas antropométricas de Alphonse. Fonte: YESHION (2014).

Francis Galton, inclusive, é considerado um dos fundadores da biometria. As suas pesquisas em capacidades e disposições mentais, que incluíam estudo de gêmeos idênticos, foram pioneiras em mostrar que vários traços são genéticos. Devido a paixão de Galton pela medição, ele abriu o Laboratório de Antropometria na Exibição Internacional de Saúde em 1884, onde recolheu estatísticas de milhares de pessoas (SINFIC, 2015).

Com o avanço da tecnologia, outros métodos de identificação biométrica foram sendo desenvolvidos, como por exemplo, o reconhecimento por voz ou a técnica de reconhecimento por digitais, desenvolvida por Francis Galton no ano de 1892. A análise de DNA, que é tão comum atualmente, foi utilizada pela primeira vez na identificação de pessoas em cena de crime no ano de 1986 pelo professor Jeffreys (LEICESTER, 2015; UPDATE, 2015).

Também foi possível desenvolver, através dos avanços tecnológicos, métodos de identificação através do caminhar. Os primeiros experimentos de análise do caminhar foram feitos por Giovanni Borelli (1608-1679) (BORELLI, 1989) e posteriormente diversos cientistas, como Isaac Newton (1642-1727) (NEWTON, 1999), contribuíram para o entendimento de como acontece o caminhar das pessoas (BAKER, 2006). Mais tarde, os pesquisadores STEVENAGE; NIXON; K. (1999) descobriram que praticamente todas as pessoas possuíam um modo único de caminhar e a partir deste momento iniciou-se pesquisas para identificar as pessoas pelo modo como elas caminham (CHEN, 2014).

A maioria das técnicas de reconhecimento através do caminhar procuram extrair a silhueta do corpo da pessoa para obter os parâmetros do caminhar (BALAZIA, 2013). Este método de identificação possui a vantagem de não ser invasivo, ou seja, a pessoa não precisa tocar em nada para ser identificado. Através deste tipo de identificação é possível reconhecer pessoas a distância, sem ela nem mesmo saber que está sendo identificada. Também possui a vantagem da dificuldade de se tentar enganar um sistema de reconhecimento por caminhar (GUILLEN et al., 2009a).

Um dos meios de capturar as informações de pessoas caminhando e assim poder trabalhar na identificação dos indivíduos é através do sensor Kinect. Este sensor, desenvolvido pela Microsoft, possui duas versões, a versão 1 desenvolvida para o vídeo-game Xbox 360 (MICROSOFT, 2015a) e a 2 desenvolvida para o Xbox One (MICROSOFT, 2015b). Neste sentido, o campo de pesquisa deste projeto é com relação a comparação entre as duas versões do Kinect na identificação de pessoas através da caracterização do caminhar delas e aplicando como meio para classificação dos dados algoritmos de aprendizado de máquina.

A utilização do sensor Kinect se justifica devido a sua capacidade de capturar dados de imagem e profundidade que permitem obter posições nas dimensões x , y , z de pontos do corpo da pessoa, possibilitando a caracterização do esqueleto do indivíduo com baixo custo. Outro ponto que justifica este trabalho é a escassez de pesquisas com a versão 2 do sensor Kinect e também o fato de que não foi encontrado nenhum trabalho comparativo entre as duas versões em relação a biometria do caminhar, de modo que existe uma lacuna em relação a mostrar qual é a diferença entre eles neste campo de pesquisa, e se há benefícios em utilizar

o novo sensor em detrimento do antigo.

2 FUNDAMENTAÇÃO TEÓRICA E ESTADO DA ARTE

2.1 O Caminhar Humano

O início do caminhar humano se dá a partir do cérebro, quando uma ideia é convertida em atividade dos músculos que são necessários ao caminhar. A execução deste comando possui dois componentes: ativação dos centros neurais inferiores que ativam as atividades musculares e o feedback sensorial de articulações, músculos, e outros receptores que modificam os movimentos (VAUGHAN; DAVID; O'CONNOR, 1999). A interação entre o sistema nervoso e a musculatura estão ilustrados na Figura 2.

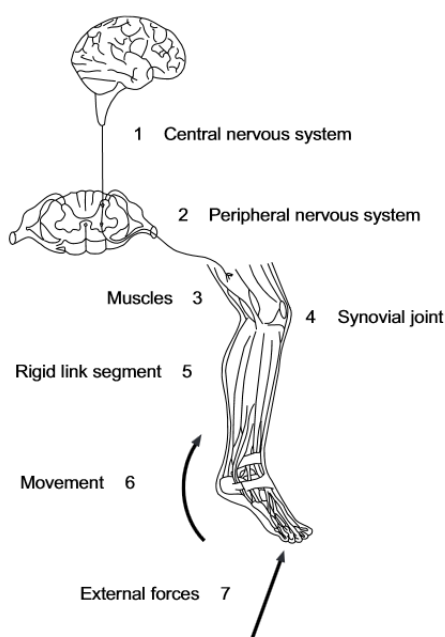


Figura 2 – Sete componentes que formam a base para o caminhar humano. Fonte: VAUGHAN; DAVID; O'CONNOR (1999).

Existem dois requisitos básicos para o caminhar: o movimento periódico de uma posição para outra de cada pé, sendo que através desse movimento um pé serve como suporte para o outro, e forças de reação suficientes, aplicado através dos pés, para suportar o corpo da pessoa (VAUGHAN; DAVID; O'CONNOR, 1999).

2.2 O Ciclo do Caminhar

Quando alguém está caminhando, um padrão cíclico de movimento, que se repete várias e várias vezes, está acontecendo. Descrições do caminhar geralmente são confinadas a somente um ciclo, assumindo que sucessivos ciclos são sempre os mesmos, embora isso não seja totalmente verdade, se aproxima o suficientemente do que acontece na realidade para a maioria das pessoas (VAUGHAN; DAVID; O'CONNOR, 1999). A Figura 3 ilustra, através de uma roda, a ideia de movimento cíclico:

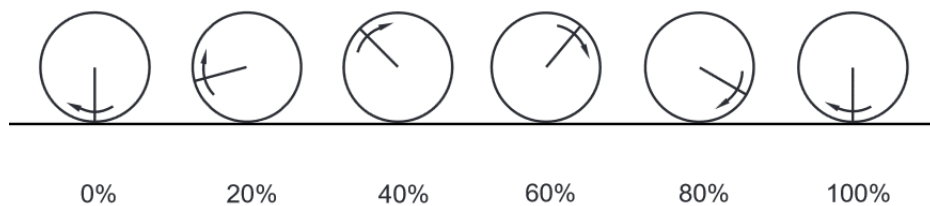


Figura 3 – Ciclo completo. Fonte: VAUGHAN; DAVID; O'CONNOR (1999).

A Figura 4 traz a ilustração do movimento cíclico do caminhar utilizando o desenho de uma pessoa como ilustração:

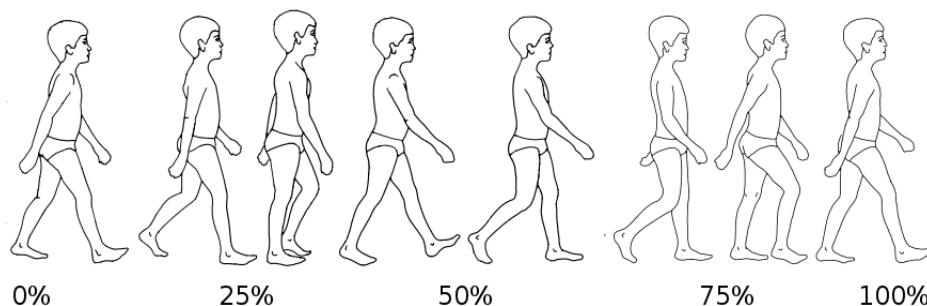


Figura 4 – Ciclo completo do caminhar humano. Fonte: Adaptado de VAUGHAN; DAVID; O'CONNOR (1999).

2.2.1 Abordagem em cinco eventos

Adelaar (1986) propôs uma divisão do ciclo do caminhar em 5 etapas, nestas etapas existem momentos em que apenas um dos pés está apoiado com o solo e existem momentos em que os dois pés estão apoiados com o solo. Com isso, podemos caracterizar o ciclo como sequência de simples e duplo apoio (ALEXANDER, 1984).

As fases propostas por Adelaar são (BRUNIEIRA, 2015):

- (a) Calcanhar do primeiro apoio em contato com o solo;
- (b) Contato total do apoio com o solo;
- (c) Primeiro apoio fora do solo;

- (d) Fase de oscilação;
- (e) Calcanhar do segundo apoio em contato com o solo.

Estas etapas propostas por Adellar podem ser resumidas ainda em 2 grandes categorias: A fase de apoio (a,b,c,e), quando o pé está no chão e a fase quando o pé não está mais em contato com o chão e a perna está se movendo em preparação para o próximo passo (d) (VAUGHAN; DAVID; O'CONNOR, 1999).

2.2.2 Componentes cinemáticos

Murray, Drought e Kory (1964) utilizaram uma técnica fotográfica para medir e temporizar componentes cinemáticos no plano sagital, frontal e transversão dos indivíduos (ANDERSSON, 2014), a Figura 5 mostra os termos utilizados neste estudo para descrever o ciclo do caminhar humano:

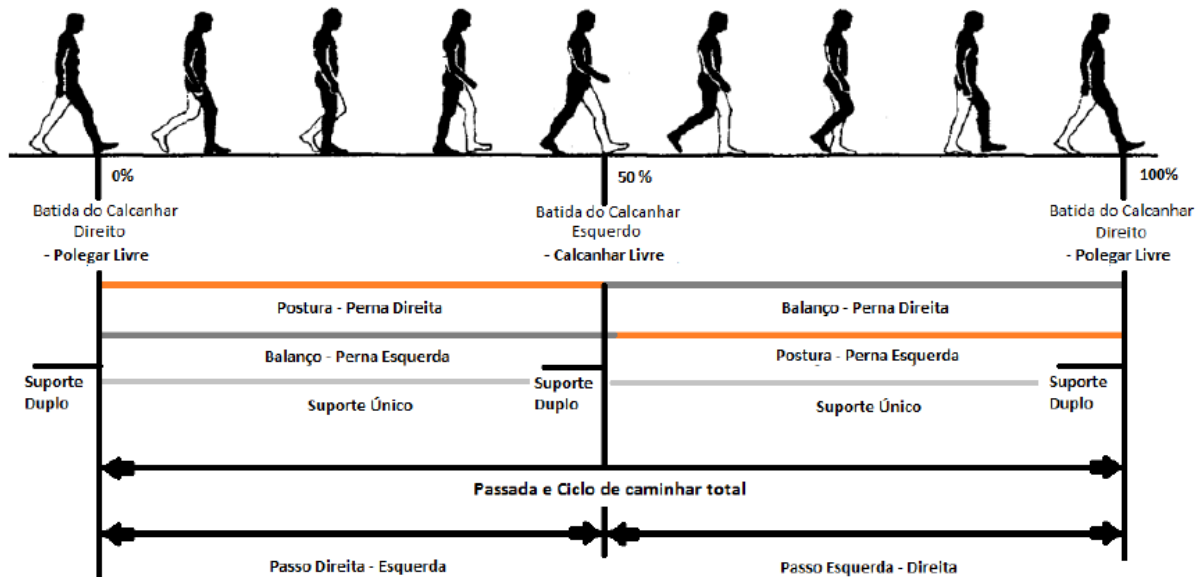


Figura 5 – Fases do ciclo do caminhar. Fonte: Adaptado por ANDERSSON (2014) de YAM; NIXON; CARTER (2004) e (YAM C. NIXON; CARTER, 2003).

Passada e cliço de caminhar total: É o intervalo de tempo entre sucessivas batidas de um mesmo calcanhar no chão (ANDERSSON, 2014).

Postura: É o tempo em que cada pé está em contato com o solo.

Balanço: Está é a fase em que o pé está suspenso no ar.

Suporte duplo e suporte único: A fase de postura possui um momento de suporte único, quando apenas um pé está em contato com o solo, e duas fases de suporte duplo, quando os dois pés estão em contato com o solo.

Comprimento do passo e da passada: Segundo (MURRAY; DROUGHT; KORY, 1964), o comprimento da passada é a distância entre sucessivos contatos de um mesmo pé

com o chão e o seu pé oposto. A medida é realizada de ambos os pé é realizada a partir de um ponto central no eixo longitudinal do pé, que coincide com o final do calcanhar.

2.2.3 Teoria dos seis determinantes da marcha

Uma das mais influentes e mais duradouras teorias do caminhar foi proposta por SAUNDERS; INMAN; EBERHART (1953), teoria conhecida como “os seis determinantes da marcha” (HEALTH, 2015). Segundo essa teoria, a marcha possui um conjunto de mecanismos que são considerados determinantes no padrão da marcha (SOUZA; TAVARES, 2015)

Humberto nos diz que, “a marcha é uma tarefa motora que envolve um padrão complexo de contrações musculares em diversos segmentos do corpo. A análise da marcha é feita dentro do evento definido como ciclo da marcha, que é a descrição da sequência de eventos que ocorrem em um membro inferior entre dois contatos iniciais consecutivos do mesmo pé”(NETO, 2015).

A teoria dos seis determinantes da marcha lida com a conservação de energia durante a locomoção do ser humano. De acordo com Saunders, a combinação de seis padrões de movimento diminui o deslocamento do centro de gravidade do corpo para aproximadamente 50%, conservando assim energia e favorecendo o movimento de locomoção (ORTHOTISTS; PROTHETISTS, 2015).

A teoria de Saunders sugere que os movimentos horizontais e verticais do corpo são dispendiosos e portanto ele desejava reduzir o deslocamento do centro de massa (CM) corporal para diminuir o gasto de energia.

As seis determinantes da marcha são (SOUZA; TAVARES, 2015):

- Rotação pélvica: Movimento da cintura pélvica que ocorre durante o caminhar;

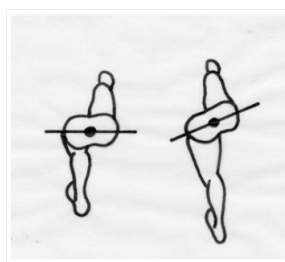


Figura 6 – Rotação pélvica. Fonte: NETO (2015).

- Inclinação da pélvis: Inclinação lateral da pélvis em oposto ao membro em apoio;
- Flexão do joelho da fase de apoio unipodálico: Flexão do joelho enquanto ele apoia o peso do corpo;
- Pé e joelho: Relação de deslocamentos angulares entre pé e joelho;
- Deslocamento lateral da pélvis: Deslocamento da pélvis para os lados durante o caminhar;

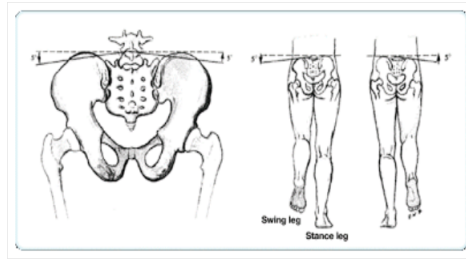


Figura 7 – Inclinação da pélvis. Fonte: NETO (2015).



Figura 8 – Deslocamento lateral da pélvis. Fonte: NETO (2015).

Por 40 anos essa teoria foi considerada como um fato, até que pelos anos de 1990 os pesquisadores começaram a medir e a quantificar características cinemáticas (HEALTH, 2015). Embora a teoria ainda seja aceita por alguns, ela carece de evidência experimental, estudos mostram que algumas determinantes não possuem grande impacto na relação de deslocamento vertical do centro de massa (CM) como proposto por Saunders (SOUZA; TAVARES, 2015).

2.2.4 Teoria do pêndulo invertido

Em oposição a teoria de Saunders, encontramos a teoria do pêndulo invertido, proposta por YAM C. NIXON; CARTER (2003). Neste modelo, o deslocamento do corpo, ao contrário do que propõe Saunders, auxilia na locomoção. Este modelo propõe que as energias cinéticas e potencial mudam em oposição de fase durante o contato do solo na fase de apoio unipodal (fase em que somente um pé está em contato com o solo), ou seja, enquanto uma energia está no valor mínimo a outra está no valor máximo, o que permite troca entre essas duas energias (CARVALHO; ANDRADE; PEYRÉ-TARTARUGA, 2015).

A troca destas duas energias significa que a energia cinética é convertida em potencial e vice-versa, conservando mais de 60% a 70% da energia mecânica requerida (SOUZA et al., 2009).

Nesse modelo, durante o passo da pessoa, o centro de massa do corpo descreve uma trajetória curvilínea (similar a um pêndulo invertido). Este mecanismo de flutuação de energia reduz o trabalho mecânico do sistema muscular pela conversão da energia e essa redução é proporcional a capacidade do organismo em converter energia cinética em potencial e vice-versa (CARVALHO; ANDRADE; PEYRÉ-TARTARUGA, 2015).

Em resumo, a teoria do pêndulo invertido considera energeticamente menos dispendioso se, durante a fase aérea, o membro se comportar como um pêndulo, descrevendo um arco, enquanto a teoria dos seis determinantes da marcha baseia-se na premissa de que os deslocamentos verticais e horizontais são energeticamente dispendiosos. (SOUZA et al., 2009)

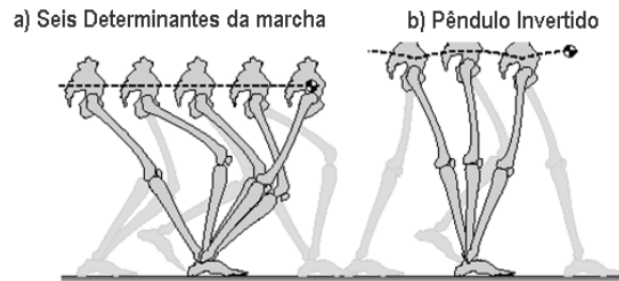


Figura 9 – Duas teorias do caminhar. Fonte: SOUZA (2009).

2.3 Análise e Modelagem do Caminhar Humano

1. Baseada em modelo (*Model-based*): Adquirir os parâmetros a partir do modelo de esqueleto construído (exemplos: ângulos das juntas, dimensões das partes do esqueleto, etc).
2. Livre de modelo (*Model-free*): Realiza medidas diretamente nos dados sem se preocupar com o modelo de esqueleto e suas especificidades (exemplos: forma da imagem, textura, cor, etc).

Nesta seção nós estudaremos análises dentro destas duas classificações.

2.3.1 Análise baseada em modelo

Nesta subseção abordaremos a análise baseada em modelo. Este tipo de análise obtém uma série de parâmetros dinâmicos ou estáticos corporais através da modelagem ou rastreamento de componentes do corpo, como pernas, coxas, braços e membros. Este tipo de análise possui a vantagem de ser independente de escalas, porém possui as desvantagens de necessitar capturar sequências do caminhar que sejam de alta qualidade e necessitar de maior processamento computacional (WANG et al., 2010).

Tipicamente, uma análise baseada em modelo possui os seguintes componentes (YAM; NIXON, 2015) (BALAZIA, 2013):

- **Captura do caminhar:** Os dados são capturados por um sensor ou câmera;
- **Extração de características:** São extraídas características estruturais (estáticas), como por exemplo a altura, e características de movimento, como por exemplo, os ângulos das trajetórias que os joelhos fazem durante o caminhar.

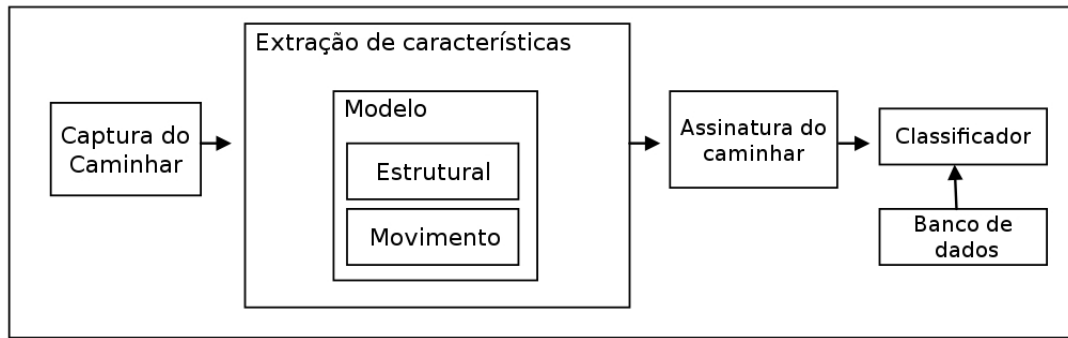


Figura 10 – Componentes de uma análise baseada em modelo. Fonte: YAM; NIXON (2015).

- **Assinatura do caminhar:** Os dados são analisados e uma assinatura para identificar o caminhar é criada;
- **Classificador:** É utilizado algum algoritmo para comparar os dados capturados com o exemplos armazenados em banco de dados.

A seguir, são apresentadas algumas análises baseadas em modelo.

2.3.1.1 Figuras de Movimento

YOO; NIXON (2011) desenvolveram um método de análise automatizada do caminhar humano, objetivando descrever, analisar, e classificar o movimento do caminhar.

Neste trabalho foi utilizada uma base de vídeos de gravações chamada de SOTON (SHUTLER et al., 2002). Nesta base estão gravações de pessoas caminhando dentro de ambientes fechados e em ambientes abertos, porém Yoo e Nixon utilizaram apenas as gravações em ambientes fechados. As capturas desta base foram realizadas com câmeras digitais, com FPS em 25 e com as pessoas caminhando em velocidade normal.

Utilizando a base SOTON, Yoo e Nixon utilizaram nos testes 100 pessoas diferentes (16 mulheres e 84 homens), para cada pessoa foram utilizadas 7 gravações e em cada gravação foi analisado 1 ciclo do caminhar percorrido pela pessoa. Embora a base de vídeos contivesse gravações que filmavam a pessoa na lateral e na diagonal, apenas as gravações da lateral da pessoa foram utilizadas. A pessoa era capturada enquanto caminhava em linha reta seguindo um caminho delimitado no chão.

A partir da base de vídeos, foram extraídos o corpo das pessoas e o seu contorno de forma a construir figuras que descrevem as pessoas. Estas figuras foram chamadas de Figuras de Movimento, e elas são formadas utilizando 8 traços ligados através de 6 pontos. Conforme ilustra a Figura 11:

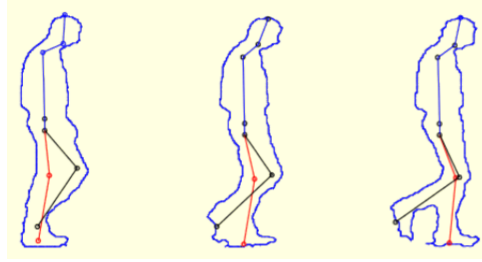


Figura 11 – Figuras de Movimento. Fonte: Adaptado de YOO; NIXON (2011)

Com base nestas figuras, parâmetros de movimento são calculados e medidos para caracterizar o caminhar das pessoas.

Pode-se resumir o sistema em três estágios:

1. Detecção e extração do movimento do corpo e seu contorno em uma sequência de imagens;
2. Extração de Figuras de Movimento pelos ângulos das juntas e pelos pontos do corpo;
3. Análises de parâmetros de movimento e extração de característica para classificar o caminhar humano.

A Figura 12 ilustra estes estágios:

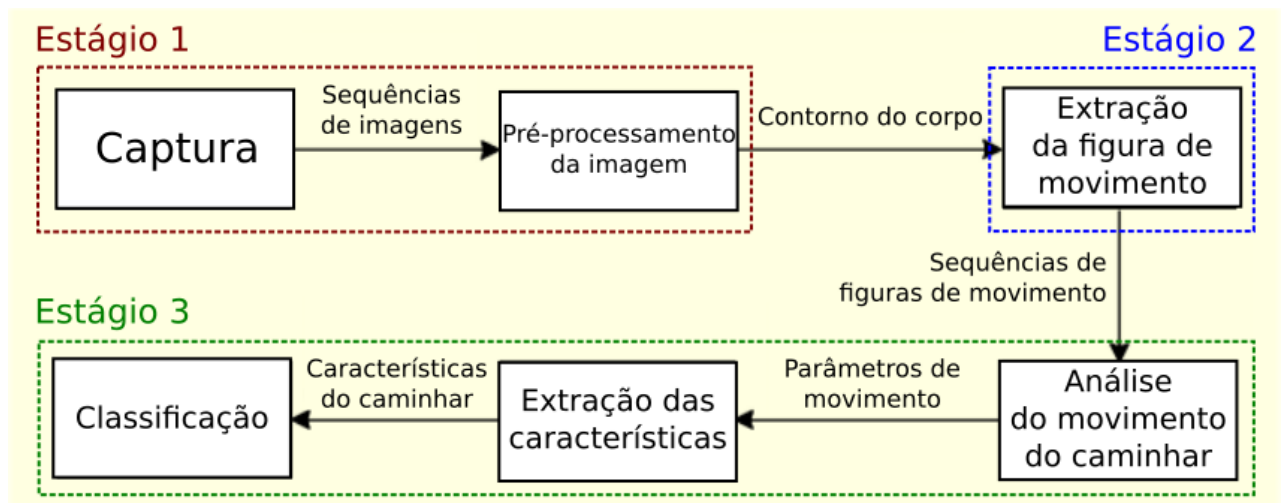


Figura 12 – Figuras de Movimento. Fonte: Adaptado de YOO; NIXON (2011)

Yon e Nixon utilizaram neste trabalho o modelo de pêndulo invertido. Este trabalho é conhecido como Figuras de Movimento (*stick figures*) ou como esqueletização (*skeletonization*) (ANDERSSON, 2014).

As seguintes características foram analisadas a fim de identificar as pessoas: tamanho da passada, tempo do ciclo, a velocidade do caminhar, coordenadas centrais do ciclograma do quadril Joelho, coeficiente de correlação entre os ângulos formados pelo joelho da perna direita

e esquerda e a média geral dos ângulos das articulações inferiores do corpo (YOO; NIXON, 2011).

O algoritmo KNN (K-nearest Neighbors) foi aplicado nos dados extraídos com o objetivo de identificar pessoas, e foi possível obter uma taxa de acerto de 96.7% utilizando 30 pessoas e 84% utilizando 100 pessoas.

2.3.1.2 *Análise de parâmetros espaço-temporais*

BENABDELKADER; CUTLER; DAVIS (2015), apresentaram um método para automaticamente estimar os parâmetros espaço-temporais do caminhar (comprimento do passo e cadência) através da análise de vídeo capturado de pessoa caminhando. O passo e a cadência são funções da altura do corpo, peso, gênero, e são usadas para identificação de pessoas.

Através de um sistema de câmera calibrada, a cadência é estimada utilizando a periodicidade do caminhar da pessoa e o comprimento do passo é estimado analisando a distância percorrida sobre um período de tempo.

Este método consiste em três módulos principais:

1. O primeiro módulo rastreia a pessoa caminhando em cada quadro, extrai a silhueta e estima a posição 2D da pessoa na imagem;
2. Após alguns frames capturados, o segundo módulo estima o período (T , em quadros por ciclo) e a distância (W , em metros) do caminhar e então calcula a cadência (C , em passos por minuto) e o comprimento do passo (L , em metros), conforme equações 1 e 2:

$$C = \frac{120 \cdot F_s}{T} \quad (1)$$

$$L = \frac{W}{n/T} \quad (2)$$

3. No terceiro módulo a identidade da pessoa é verificada utilizando uma classificação parametrizada Bayesiana.

Nos experimentos, foi criado um banco de dados com 17 pessoas capturadas com uma câmera digital da Sony, onde cada pessoa tinha em média 8 exemplos capturados, totalizando um total de 131 exemplos. Cada pessoa foi instruída a caminhar em torno de 30 metros em uma linha reta a uma velocidade fixa. A Figura 13 ilustra o ângulo de captura:



Figura 13 – Ângulo de captura do método de identificação de Benabdelkader. Fonte: BENABDELKADER; CUTLER; DAVIS (2015).

Como resultado na identificação de pessoas, obteve-se uma probabilidade de identificação correta de 40%. Os testes foram realizados utilizando *all but one* dos 131 exemplos. Este método funciona com imagens em baixa resolução e é robusto para mudanças de luz, roupa e erros de rastreamento.

2.3.1.3 Reconhecimento através de parâmetros estáticos do caminhar

BOBICK; JOHNSON (2001), desenvolveram uma técnica que captura parâmetros estáticos do corpo da pessoa específicos da ação do caminhar, como por exemplo, a distância entre os pés enquanto os dois estão tocando o chão durante o caminhar. Este método primeiramente utiliza um conjunto formado por quatro distâncias :

- A distância vertical entre a cabeça e o pé (d_1);
- A distância entre a cabeça e a pélvis (d_2);
- A distância entre o pé e a pélvis (d_3);
- A distância entre o pé esquerdo e o pé direito (d_4).

E também foi utilizado um segundo conjunto, formado pelas distâncias d_1 e d_3 .

A Figura 14 ilustra estas distâncias:

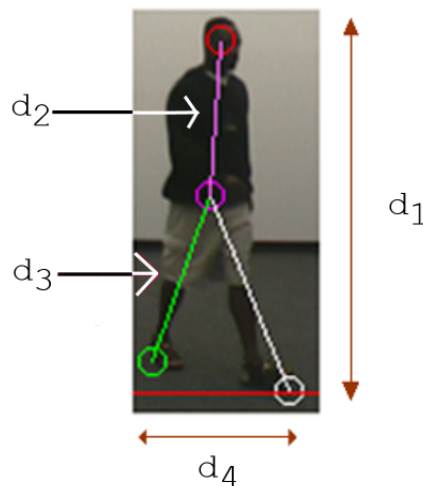


Figura 14 – Distâncias do corpo utilizadas como parâmetros estáticos. Fonte: BOBICK; JOHNSON (2001).

Estas distâncias são medidas durante o caminhar e somente quando os dois pés estão na distância máxima de separação durante o momento em que ambos estão tocando o chão e são concatenadas na forma de um vetor de 4 dimensões: $w = \langle d1, d2, d3, d4 \rangle$ e um vetor de 2 dimensões $w = \langle d1, d3 \rangle$.

Como resultado, verificou-se que o primeiro conjunto $w = \langle d1, d2, d3, d4 \rangle$ é mais sensível ao ponto de vista do que o segundo formado por $w = \langle d1, d3 \rangle$ (WANG et al., 2010; BOBICK; JOHNSON, 2001).

2.3.1.4 Identificação Biométrica com Antropometria e Caminhar Humano Utilizando o Kinect

ANDERSSON (2014), em sua dissertação de mestrado, aplicou algoritmos de aprendizado de máquina para classificar indivíduos através de atributos antropométricos e do caminhar humano utilizando o Kinect 360. Foram capturadas 164 pessoas para realizar o treinamento dos modelos onde cada indivíduo contém em média, 5 exemplos de caminhar, totalizando 702 exemplos. Dos indivíduos capturados, 113 eram homens e 51 eram mulheres.

Das capturas realizadas foram extraídos dados antropométricos, cinemáticos (baseados na teoria do pêndulo invertido de YAM C. NIXON; CARTER (2003)) e espaço-temporais (baseados no trabalho de YOO; NIXON (2011)). Para a análise antropométrica, foram calculado a média e o desvio padrão da distância dos segmentos em todos os quadros capturados. Para a análise cinemática foram calculados a média e o desvio padrão dos ângulos formados pelos quadris, joelhos, tornozelos e abertura do pé durante a caminhada. E como características espaço-temporal foram calculados o tamanho do passo, tamanho da passada, duração do ciclo e velocidade.

Os algoritmos utilizados para a classificação dos indivíduos foram o KNN, Multilayer Perceptron (MLP) e o Suport Vector Machines (SVM).

Ao aplicar aprendizado de máquina para identificar os indivíduos, o algoritmo que obteve o melhor desempenho foi o KNN com 87,74% de acurácia utilizando $k = 4$, seguido pelo SVM

com 87,3% de acurácia.

Para a coleta dos dados foi capturado o ângulo lateral da pessoa, com a mesma percorrendo uma trajetória semi-circular no sentido horário e anti-horário, conforme demonstra a Figura 15: (ANDERSSON, 2014)

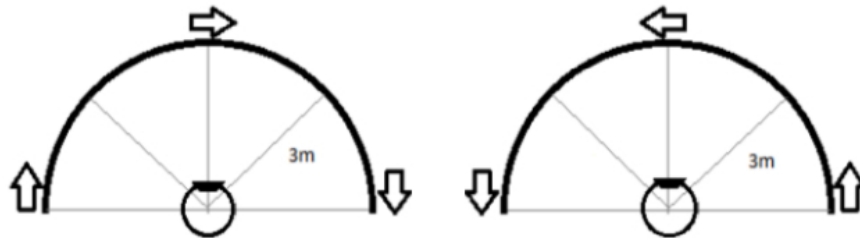


Figura 15 – Trajetória semi-circular percorrida pela pessoa. Fonte: Adaptado de ANDERSSON (2014).

2.3.1.5 Identificação de pessoas através de redes neurais

YOO JANG-HEE; MOON; HWANG DOOSUNG; NIXON (2008) utilizaram o algoritmo Multi-layer Perceptron (MLP) para reconhecer pessoas pelo seu caminhar. O movimento do caminhar é descrito como ações rítmicas e periódicas e um esqueleto 2D é extraído da silhueta da pessoa caminhando através de informações de movimento com análises topológicas guiadas pelo conhecimento da anatomia.

Um conjunto sequencial de esqueletos 2D, construído a partir de nove coordenadas capturadas, é usado como dados primários para a extração de características baseadas em parâmetros de movimento (WANG et al., 2010; YOO JANG-HEE; MOON; HWANG DOOSUNG; NIXON, 2008).

A Figura 16 ilustra uma sequência de esqueletos 2D descrevendo o caminhar.

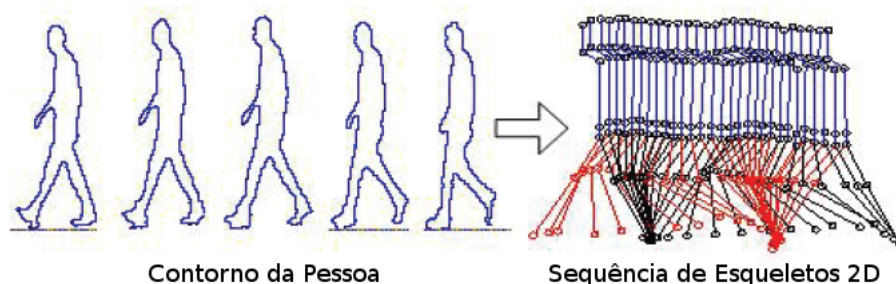


Figura 16 – Sequência de esqueletos 2D. Fonte: Adaptado de YOO JANG-HEE; MOON; HWANG DOOSUNG; NIXON (2008).

A velocidade da pessoa, assim como a média e a variação dos ângulos durante uma sequência do caminhar também foram utilizadas como características para identificação do indivíduo. Com todas as características coletadas, uma rede neural com *back-propagation* é utilizada para identificar as pessoas.

Nos experimentos, foram utilizadas 90 pessoas com 7 sequências cada uma de caminhada capturadas. O ângulo de captura dos indivíduos foi o lateral e eles foram categorizados em 3 grupos diferentes com 30 indivíduos em cada, de acordo com a qualidade dos dados. Dentre as 7 sequências de cada indivíduos, 4 foram utilizadas para treinar a rede neural e 3 para testes. O número de neurônios na camada invisível e na camada de saída da rede neural são 28 e 13, respectivamente. Como resultado, verificou-se que no grupo com a melhor qualidade nos dados foi possível obter uma taxa de acerto de 90%, no grupo com qualidade média e no de pior qualidade obteve-se em ambos 83.3% de acerto.

2.3.1.6 Análise da trajetória dos ângulos do quadril e do joelho

TANAWONGSUW R.; BOBICK (2015), desenvolveram uma pesquisa no sentido de reconhecer o indivíduo através da análise de características dinâmicas do caminhar dela, mais especificamente, a trajetória dos ângulos das juntas da parte inferior do corpo. Para capturar os dados, foi utilizado um sistema eletro-magnético de captura de movimento.

São colocados no usuário 16 sensores, distribuídos da seguinte maneira: em cada um dos tornozelos, pernas, coxas, mãos, antebraços, braços, um na cabeça, um na pélvis e dois no torso. Os dados são capturados a 30Hz e gravados no mesmo momento.

A Figura 17 ilustra uma pessoa utilizando os sensores e o resultado dos dados ilustrados em um plano:

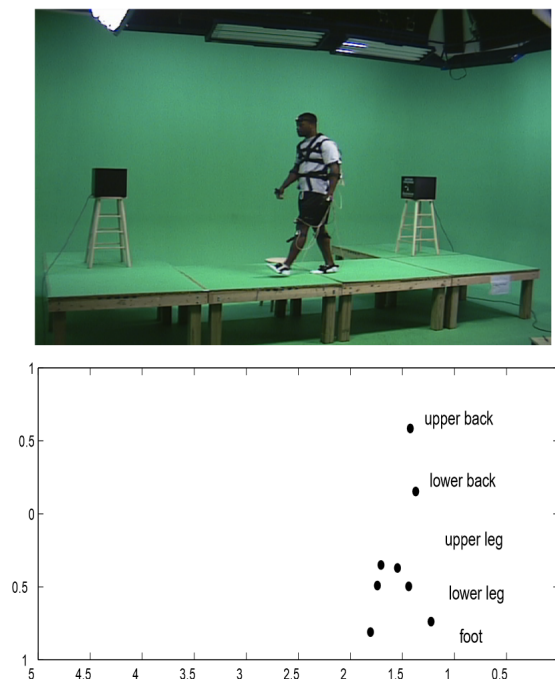


Figura 17 – Usuário utilizando sensores de captura. Fonte: Adaptado de TANAWONGSUW R.; BOBICK (2015).

2.3.1.7 *Análise de movimento através de velocity Hough transform*

NASH; CARTER; NIXON (1997) desenvolveram uma pesquisa com o objetivo de extrair parâmetros do corpo humano utilizando uma sequência de imagens. A análise é feita através de uma variação do algoritmo Hough Transform (HT). O algoritmo HT é utilizado para detectar círculos, linhas e outros tipos de curvas (SOLBERG, 2015).

Inicialmente, observa-se na literatura, que o algoritmo HT foi utilizado apenas para extrair características estáticas, sem análise de uma sequência de imagens, ou analisando parâmetros de movimentos de uma sequência de imagem sem considerar as características estacionárias. Características estáticas geralmente se referem a medidas baseadas na geometria, como a altura do corpo por exemplo, e características de movimento referem-se a parâmetros dinâmicos, como por exemplo, trajetórias de ângulos formados por juntas do corpo.

Na época em que esta pesquisa se desenvolveu, não havia trabalhos que haviam analisado ambas as características (estacionárias e de movimento) ao mesmo tempo. A implementação do HT, por analisar ambas as características, recebeu o nome de VHT (Velocity Hough Transform).

Uma das aplicações do VHT nesta pesquisa foi extrair a estrutura e movimento das coxas, e mostrou que o algoritmo, por considerar uma sequência de imagens como um todo, analisando tanto parâmetros estacionários como de movimento, tem performance melhor em lidar com imagens com ruído ou obstrução parcial do que o algoritmo HT original (NASH; CARTER; NIXON, 1997; WANG et al., 2010).

2.3.1.8 *Algoritmo de Wang, Ning e Hu*

WANG et al. (2005) realizaram um trabalho no sentido de desenvolver um algoritmo para analisar parâmetros estáticos e dinâmicos da biometria do corpo humano. Para a captura de parâmetros estáticos, representações da silhueta da pessoa caminhando foram extraídas e aplicadas em vetores. Esses vetores foram aplicados ao método *Procrusters shape analysis* para obter uma representação mais compacta, chamada de informação estática do corpo (WANG et al., 2005).

Para obter os parâmetros dinâmicos do caminhar, foi utilizado um *framework* chamado *Condensation*. Os parâmetros capturados foram as trajetórias dos ângulos das juntas dos membros inferiores.

Após as informações estáticas e dinâmicas serem capturadas e processadas elas são aplicadas separadamente ao algoritmo *nearest exemplar classifier (ENN)* e depois os resultados desse algoritmo foram aplicados a um classificador de combinação.

A Figura 18 ilustra uma visão geral do algoritmo:

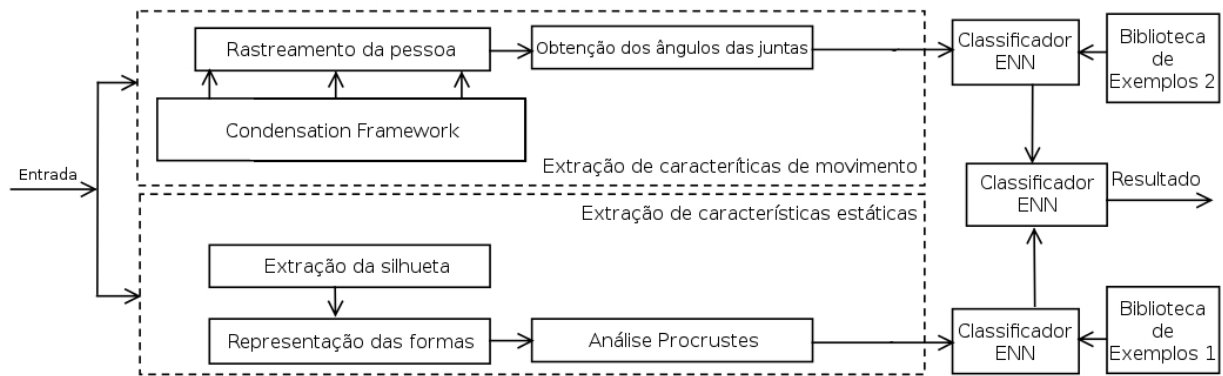


Figura 18 – Fluxograma do algoritmo de Wang. Fonte: Adaptado de WANG et al. (2005).

Nos experimentos foram coletados 4 exemplos de 20 pessoas, totalizando 80 exemplos. Em cada exemplo foi capturado uma imagem lateral da pessoa caminhando. Todas as imagens foram capturadas por uma câmera fixa com uma taxa de 25 quadros por segundo. Como resultado, verificou-se que utilizando as características estáticas e de movimento em conjunto trazem um resultado melhor que utilizar apenas uma delas separadamente. Com o método proposto o autor conseguiu obter uma acurácia máxima de 97.50%.

2.3.1.9 Análise de trajetórias de coordenadas 3D

BALAZIA (2013), em sua dissertação de mestrado, desenvolveu uma pesquisa no sentido de analisar as trajetórias de coordenadas 3D de pontos do corpo humano (ombros, mãos, cotovelos, lombos, joelhos e pés) e compará-las com dados armazenados em banco de dados com o objetivo de identificar pessoas.

Os dados que são analisados são a distância entre os pares dos pontos do corpo, ou seja, a distância entre um ombro e o outro, entre uma mão e outra, e assim por diante. Com essas informações, é possível obter métricas de distâncias relativas ao tempo, ou seja, qual a distância dos membros do corpo em cada momento, e assim, poder identificar o caminhar da pessoa.

A Figura 19 demonstra o processo do método implementado:

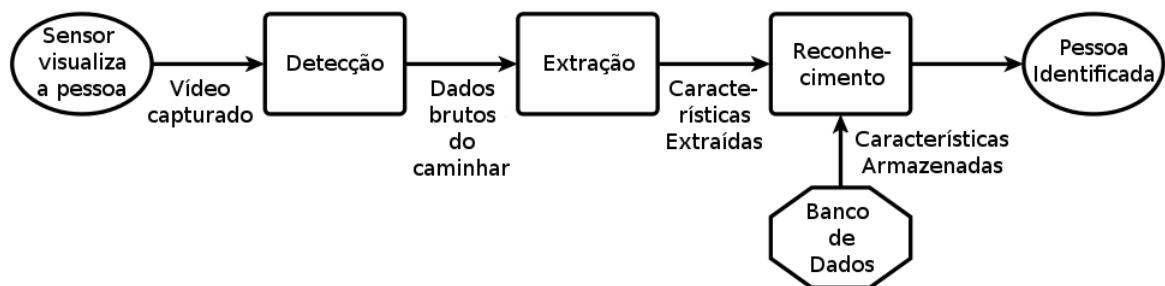


Figura 19 – Fluxograma da metodologia de Balazia. Fonte: Adaptado de BALAZIA (2013).

- **Sensor visualiza a pessoa:** Etapa em que a pessoa está caminhando em frente ao sensor;

- **Detecção:** O esqueleto da pessoa é identificado;
- **Extração:** São extraídos os pontos do corpo necessários para a análise;
- **Reconhecimento:** Os dados capturados são comparados com os exemplos em banco de dados;
- **Pessoa identificada:** É apresentado o resultado do algoritmo.

Na etapa de reconhecimento, quando os dados capturados são comparados com exemplos no banco de dados, foi utilizado um algoritmo implementado pelo autor para realizar a comparação. O algoritmo se baseia nas distâncias entre os membros do corpo e o peso que cada distância possui.

A base utilizada nos experimentos é a MoCap DB (MELLON, 2016), a qual possui 144 pessoas capturadas em diferentes tipos de movimentos (caminhar, dançar, etc). Desta base, o autor desta pesquisa utilizou 131 capturas correspondentes a 24 pessoas caminhando.

2.3.1.10 Método de segmentação

Proposto por YEOH et al. (2011), este método trabalha com a segmentação do corpo da pessoa em oito partes e infere matematicamente as posições do joelho e da forquilha (altura do chão até o osso púbis). São utilizados como parâmetros os ângulos descritos pelo joelho, forquilha, altura da forquilha, tamanho do passo, altura e largura da silhueta do indivíduo.

É aplicado o algoritmo *Fuzzy K-Nearest Neighbor (Fuzzy KNN)* como método de aprendizado de máquina para identificação da pessoa baseada nos dados processados.

Este método pode ser dividido em oito passos:

- O primeiro passo deste método é realçar a imagem que contém a silhueta da pessoa através de abertura morfológica;
- A altura e a largura da silhueta são medidas;
- A silhueta do corpo é dividida em oito partes de acordo com as medidas de altura e largura realizadas;
- A parte de baixo do corpo é automaticamente identificada;
- As trajetórias das juntas da parte inferior do corpo são computadas;
- O tamanho do passo e a altura da virilha são medidos;
- O algoritmo *Fuzzy K-Nearest Neighbor* é aplicado;
- Como resultado obtém-se a identificação da pessoa;

A Figura 20 demonstra a medição do corpo da pessoa e a divisão em 8 segmentos:

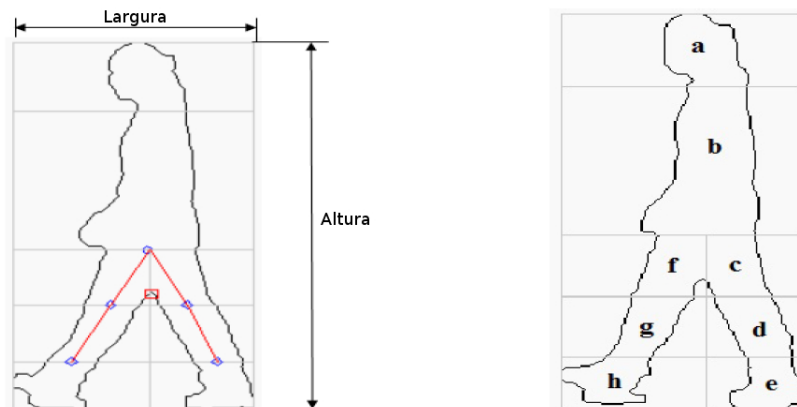


Figura 20 – Dimensões obtidas do indivíduo e a sua segmentação. Na esquerda são ilustradas a altura e largura da pessoa e na direita é apresentada a divisão em oito partes. Fonte: Adaptado de YEOH et al. (2011).

Nos testes realizados, foram utilizadas 11 pessoas caminhando em paralelo a uma câmera estática. Cada pessoa foi capturada vestindo uma variedade de calçados (chinelos, botas, meias, pés descalços, etc), roupas e carregando mochila, mala ou bolsa. Elas também foram capturadas caminhando em velocidades diferentes (devagar, normal e rápido). Para cada pessoa, foram capturados em torno de 20 sequências, indo da direita pra esquerda e vice-versa e vindo de frente. Foi possível uma identificação correta das pessoas com uma taxa de acerto de 87,96% com uma configuração de k para o *Fuzzy KNN* de $k = 6$ ou $k = 7$.

2.3.2 Análise livre de modelo

A análise livre de modelo não considera os parâmetros básicos e fundamentais do caminhar humano, porém analisa características aparentes extraídas durante o ciclo. Este método possui a vantagem de ser geralmente menos complexo e menos custoso computacionalmente, porém tende a ser mais dependente da pose de captura, ou seja, se restringir a funcionar apenas no ângulo de visão em que a pessoa foi capturada (ANDERSSON, 2014; SIVAPALAN et al., 2011).

2.3.2.1 Análise do corpo através de sub-partes

VENKAT; WILDE (2010), estudaram o problema de reconhecimento de pessoas através do caminhar usando um modelo probabilístico de interpretação de sub-partes do caminhar. Quando nós vemos à distância uma pessoa caminhando, nós podemos perceber algum padrão no balançar dos ombros ou no mover do quadril como características do andar da pessoa, o objetivo do algoritmo é capturar essas partes que formam o padrão do caminhar. Uma sub-parce do caminhar é definida como uma parte da silhueta do corpo se movimentando.

A Figura 21 demonstra a divisão da silhueta da pessoa em sub-partes.

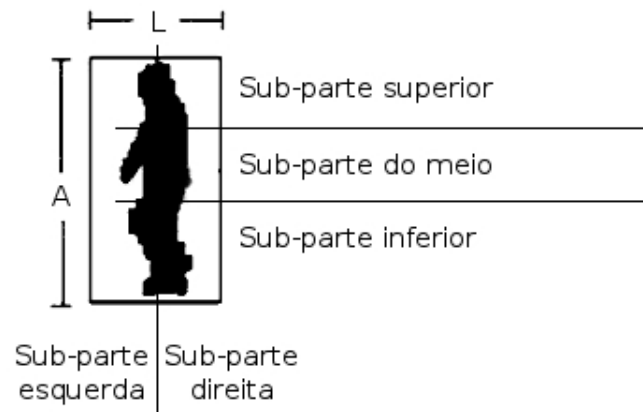


Figura 21 – A silhueta típica de uma pessoa e suas sub-partes. Fonte: Adaptado de VENKAT; WILDE (2010).

Para descobrir as sub-partes do caminhar que formam o padrão da caminhada, é utilizado aprendizado de máquina através de Rede Bayesiana e é aplicado inferência probabilística para comparar os dados com a base de exemplos com o fim de identificar a pessoa (VENKAT; WILDE, 2010).

Nos experimentos foram utilizadas duas bases para servir de exemplos de pessoas caminhando: a base *HumanID gait challenge*, que foi utilizada em uma competição realizada pela Universidade do Sul da Florida (USF) (SARKAR et al., 2005), e uma base com gravações de vários ângulos fornecida pela Academia Chinesa de Ciências (CASIA, 2006).

A base de dados da USF possui 122 pessoas capturadas em ambiente externo. Da base de dados da Academia Chinesa de Ciências foi utilizado um conjunto de 124 pessoas capturadas em 11 ângulos diferentes.

Utilizando a base de dados da USF, foi possível atingir um máximo de acurácia na identificação dos indivíduos de 90% e com a base da Academia Chinesa de Ciências foi possível obter 89.6% de acurácia.

2.3.2.2 *Gait Energy Images*

Gait Energy Images (GEI) é um método proposto por KUSAKUNNIRAN et al. (2009) e foi desenvolvido para reconhecimento do caminhar. GEI representa em uma única imagem, as características do caminhar através de múltiplas silhuetas de uma pessoa realizando um ciclo do caminhar. As silhuetas são normalizadas e alinhadas, formando uma representação compacta da ocupação espacial da pessoa durante um ciclo do caminhar, levando informações de suas características do caminhar, bem como aparência, permitindo que seja feita uma identificação. A Figura 22 ilustra a captura de diversos ângulos (SIVAPALAN et al., 2011).

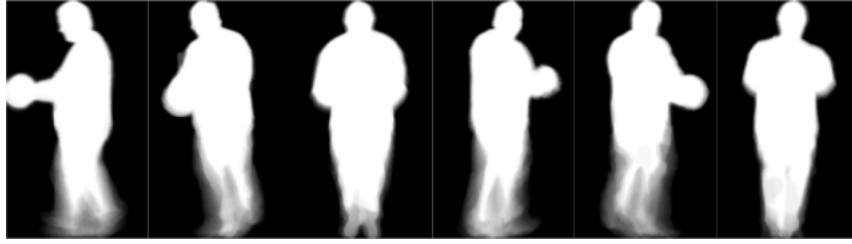


Figura 22 – GEIs extraídos de múltiplos ângulos. Pixels mais claros representam maior ocupação no ciclo do caminhar. Fonte: SIVAPALAN et al. (2011).

Várias extensões do conceito do GEI original foram propostas, incluindo o *Enhanced Gait Energy Image* (EGEI) (LIN; WANG, 2010) e o *Shifted Energy Image* (SEI) (HUANG; Boulgouris, 2010), que procuram lidar com problemas como, por exemplo, as roupas que as pessoas usam, porém não fornecem grandes melhoras em situações como ângulos de captura diferentes (SIVAPALAN et al., 2011).

2.3.2.3 Gait Energy Volumes

Gait Energy Volumes (GEV) possui várias vantagens sobre o *Gait Energy Images*, principalmente por trabalhar no espaço tri-dimensional. Ele contorna o problema da dependência de ângulo, não tem ambiguidade de pose ou oclusão causada pelo próprio corpo da pessoa e permite uma segmentação mais fácil de regiões não desejadas (como movimentos da mão, por exemplo). Porém este método também possui desvantagens, como por exemplo, exigir um computador com grande capacidade de processamento, fazendo com que fique impraticável dependendo da aplicação (SIVAPALAN et al., 2011).

Ao invés das silhuetas utilizadas no GEI, no GEV são utilizados voxels, semelhantes a matrizes, capazes de armazenar dados tri-dimensionais. Os voxels são espacialmente alinhados sobre um ciclo do caminhar de acordo com a equação 3:

$$GEV(k) = \frac{1}{n} \sum_{t=1}^n V(t) \quad (3)$$

Onde n é o número de quadros no K e V são os volumes de voxels alinhados. A Figura 23 ilustra o modelo:

Para testar os algoritmos propostos foi utilizado a base CMU MoBo (GROSS; SHI, 2001) e uma outra criada pelos autores do trabalho com gravações em ambiente interno.

A base de dados CMU MoBo possui 25 pessoas capturadas por 6 câmeras simultaneamente. As pessoas foram gravadas enquanto caminhavam em uma esteira. Quatro classes de capturas foram realizadas: caminhando devagar, caminhando rápido, caminhando segurando uma bola e caminhando em superfície inclinada. A última classe não foi utilizada nesta pesquisa porque não foi possível extrair de forma clara a silhueta da pessoa.

A Figura 24 ilustra os ângulos de captura das 6 câmeras utilizadas para as capturas da CMU MoBo:

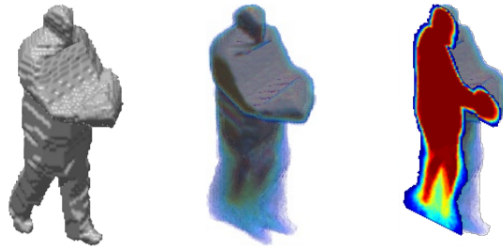


Figura 23 – Na esquerda está um exemplo de um modelo de reconstrução de voxel, na imagem do meio é apresentado um GEV e na direita está uma representação cortada do GEV exibida para melhor ilustrar as densidades internas. Fonte: SIVAPALAN et al. (2011).



Figura 24 – Demonstrativos dos ângulos de captura da base *CMU MoBo Database*. Fonte: SIVAPALAN et al. (2011).

A segunda base utilizada para os testes, a qual foi criada pelos autores do trabalho, consiste de 15 pessoas caminhando em direção ao Kinect, ou seja, apenas o ângulo frontal foi capturado. As pessoas foram capturadas caminhando em duas velocidades diferentes, normal e rápida. Cinco sequências foram gravadas para cada indivíduo, cada sequência possuía uma média de 2 a 3 ciclos. A captura ocorreu com aproximadamente 30 quadros por segundo e embora a imagem de cor fornecida pelo Kinect tenha sido capturada, foram utilizadas apenas as informações de profundidade.

Objetivando a classificação das pessoas, em ambas as bases foram testados dois tipos de identificação: utilizando capturas de mesma classe e a identificação utilizando classes diferentes. Na identificação das pessoas utilizando a mesma classe, cada captura foi dividida pela metade, onde uma parte era para treinamento e a outra para teste. Já para a classificação das pessoas utilizando classes diferentes, foram utilizadas as capturas por inteiro.

Em todos os testes, para comparação e classificação, foi utilizado o algoritmo de Boulgouris et al. (2006).

Na identificação das capturas utilizando a base Cmu MoBo, foi possível uma acurácia de 100% sem falsos alarmes para testes em capturas de mesma classe e para a classificação de classes diferentes foi possível alcançar 70% de acurácia sem falsos alarmes.

Utilizando a base criada pelos autores da pesquisa, tanto nos testes inter-classes como com os utilizando a mesma classe foi possível obter taxa de acerto de 100%.

2.3.2.4 Caixa Delimitadora

Este método, chamado de Caixa Delimitadora (*Bundle Rectangle*), foi proposto por GUILLEN et al. (2009a) e extrai características da pessoa ao delimitar um retângulo ao redor dela.

Uma vez que o indivíduo é capturado por uma câmera convencional, o plano de fundo é subtraído, as características da pessoa são extraídas e o indivíduo é reconhecido comparando as características obtidas com as armazenadas em banco de dados.

A Figura 25 ilustra a caixa que é delimitada ao redor dos indivíduos:



Figura 25 – Retângulo demarcando a silhueta da pessoa. Fonte: GUILLEN et al. (2009a).

Nos experimentos foi utilizada uma base contendo gravações em vídeo de 100 pessoas e o ângulo de captura foi a lateral da pessoa. Nesta base para cada pessoa foram capturados 5 vídeos de 8 segundos de duração. (GISSIC, 2016).

O algoritmo utilizado para identificar as pessoas através desta metodologia foi desenvolvido utilizando uma árvore de decisão. Para uma pessoa ser considerada como autenticada ela deveria passar por todas as etapas da árvore, as seguintes características foram analisadas para cada caixa delimitada ao redor dos indivíduos: altura, largura, área, ângulo diagonal da área e potência espectral total (dada pela soma de cada aplicação da Transformada Rápida de Fourier (FFT) nas áreas calculadas em cada quadro). Este algoritmo se baseia na seguinte equação:

A avaliação dos resultados dos testes foi feita através da equação 4:

$$R = 100\% - FRR - FAR \quad (4)$$

onde *FRR* representa a probabilidade de um impostor ser identificado como uma pessoa autorizada e *FAR* descreve a probabilidade de uma pessoa autorizada não ser reconhecida. Os resultados mostraram uma eficácia do método proposto como tendo uma confiança entre 95% a 100%.

2.3.2.5 Histograma de Gradiente em Imagem Gaussiana

Neste trabalho, os autores incorporaram o *Histogram of Oriented Gradients* (HOG) em imagens Gaussianas para melhorar a identificação das pessoas através do caminhar. Este tipo de representação espaço-temporal é chamado de *Gradient Histogram Gaussian Image* (GHGI)

(ARORA et al., 2015a).

Este método é similar ao *Gait Energy Image* (GEI), porém a utilização de função Gaussiana e a aplicação de HOG melhora a eficiência e reduz os ruídos. O GHGI preserva mais informações das imagens do que o GEI, onde apenas as linhas das silhuetas são preservadas. No fim do processo, os GHGI são utilizados no algoritmo K-Nearest Neighbor para a identificação da pessoa. Podemos descrever os passos deste método da seguinte maneira: (ARORA et al., 2015a)

1. Pre-processamento: As imagens que compõem o ciclo do caminhar são pré-processadas para alinhar o corpo da pessoa no meio, normalizar e dimensionar a imagem no tamanho de 128x88.
2. Período de detecção do caminhar: Neste momento é extraído das imagens o ciclo do caminhar;
3. Aplicação da função Gaussiana: Características gaussianas são extraídas para cada pixel das imagens que compõe o caminhar;
4. Utilização do HOG em imagens Gaussianas: As imagens Gaussianas são transformadas em múltiplos *templates* de Histogramas (GHGI) através da extração de dados usando HOG.
5. Identificação: Os GHGI são aplicados ao algoritmo KNN para identificar a pessoa.

A Figura 26 demonstra a etapa quatro, onde as imagens gaussianas são transformadas em GHGI.

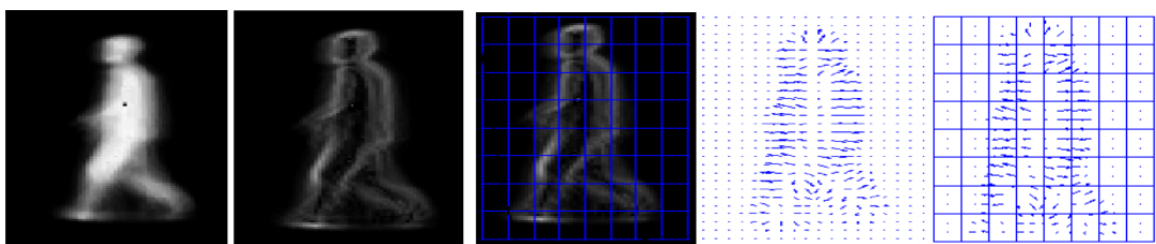


Figura 26 – Imagens gaussianas sendo transformadas em GHGI. Fonte: ARORA et al. (2015a).

Nos experimentos foi utilizado as bases de dados A e B do Instituto de Ciência e Pesquisa Industrial da Universidade de Osaka (OSAKA, 2016). Em ambas as bases de dados as pessoas foram gravadas caminhando em uma esteira e o ângulo de captura foi o da lateral da pessoa. A base de dados A possui 34 indivíduos caminhando em velocidades de 2 Km/h a 7 Km/h, embora o autor tenha utilizado apenas as variações de 5 Km/h a 7 Km/h. A base B contém 68 pessoas e embora as pessoas caminhem na mesma velocidade existe uma maior variação de roupas.

Nos testes realizados com a base A, foi possível obter 94.1% de acurácia na identificação das pessoas caminhando com velocidade de 5 Km/h e 6 Km/h e foi possível obter 82.3% de acurácia na velocidade de 7 Km/h. Já nos testes com a base B, a acurácia ficou em 77%.

2.4 Pesquisas com o Kinect na Análise Biométrica

O sensor Kinect deixou de ser utilizado apenas em jogos para ser aplicado em outras áreas, principalmente na identificação biométrica. Desde o seu lançamento no ano de 2010 CNET (2016), inúmeros estudos foram realizados procurando aplicar os dados fornecidos pelo sensor para além da área de entretenimento. Nesta seção serão mostradas diversas aplicações realizadas utilizando o sensor.

2.4.1 Análise biométrica através do caminhar

ARAUJO; GRAÑA; ANDERSSON (2013), pesquisaram a respeito de extrair pontos do esqueleto de pessoas caminhando para usar em identificação biométrica. Foi calculado o tamanho de várias partes do corpo e utilizado para treinar algoritmos de aprendizado de máquina: Multi-layer perceptrons (usando backpropagation com 10 neurônios na camada oculta), Decision Tree (usando C4.5 e Random Forests) e KNN (com k igual a 1).

O treinamento foi realizado com 8 pessoas (2 mulheres e 6 homens), todos da mesma faixa etária (25 a 35 anos), cada pessoa realizou 10 caminhadas em frente ao sensor, 5 paralelas e 5 perpendiculares, totalizando 1400 frames capturados para cada indivíduo e média de 4 segundos em cada caminhada. As partes do corpo usadas para treinar os modelos foram ambos os braços, antebraços, pernas, coxas, espinha toraxica, espinha cervical e o cálculo da altura. O melhor resultado foi obtido através do KNN, com 99,6% de acurácia.

PREIS et al. (2012), da Universidade de Munique Ludwig Maximilians, não utilizou diretamente os dados de profundidade e de cor fornecidos pelo Kinect, porém em seu trabalho, fez uso das posições x e y fornecidas pelo Kinect 360.

O sistema proposto consiste em três componentes:

1. O primeiro componente grava as informações do esqueleto da pessoa;
2. No segundo passo, são definidas 13 características biométricas obtidas através dos 20 pontos do esqueleto humano capturados pelo Kinect.
3. Por fim, é aplicado algoritmo de aprendizado de máquina para identificar a pessoa.

As 13 características definidas no segundo passo, são: Altura do corpo, comprimento das pernas, tronco, canelas, coxas, parte superior dos braços, ante-braços, largura do passo e velocidade. Os dados foram capturados com a pessoa caminhando de lado para o sensor.

Nos testes, foi utilizado validação cruzada de 7 subconjuntos. A base é constituída de 8 pessoas, caminhando em ângulo perpendicular (de lado) ao sensor, instruídas a caminhar em sua velocidade normal.

Como resultado, verificou-se que quatro características: Altura, comprimento das pernas e do torso e a parte de cima do braço esquerdo foram suficientes para corretamente classificar a pessoa em 91% dos casos utilizando classificação Bayesiana. Foi testado também os algoritmos 1R e C4.5, porém apresentaram resultados inferiores, 62,7% e 76,1%, respectivamente.

SINHA; CHAKRAVARTY; BHOWMICK (2013), utilizaram um método semelhante ao de PREIS et al. (2012), eles não se basearam diretamente nos dados de profundidade ou de imagem do Kinect, mas utilizaram os dados dos 20 pontos do esqueleto humano que são fornecidos pelo sensor.

Primeiramente, foram gravados os dados capturando as pessoas caminhando de lado em relação ao sensor. Após, foram extraídos características do caminhar analisando metade de cada ciclo do caminhar capturado das pessoas, e contruindo configurações com as distâncias entre as juntas do corpo. No terceiro momento, foi aplicado o algoritmo de rede neural Multi-layer Perceptron (MLP), para identificar o indivíduo. O MLP foi aplicado utilizando uma camada oculta com 25 neurônios.

A rede neural MLP usada neste trabalho foi adaptada para o aprendizado ocorrer no modo Batch, onde todos os exemplos são aplicados a rede neural e o erro de cada camada são acumulados para calcular o valor dos pesos. Também foi aplicado antes dos neurônios de entrada mais 2 neurônios chamados de neurônios de seleção, que tem por objetivo aumentar os valores de propriedades mais favoráveis e diminuir o de propriedades menos favoráveis.

Utilizando 5 indivíduos, com aproximadamente 140 ciclos de caminhadas para cada um, foi possível obter uma acurácia de 86%.

MILOVANOVIC; MINOVIC; STARCEVIC (2013), implementaram um método que transforma os pontos do esqueleto da pessoa, capturados através do Kinect 2, enquanto a pessoa caminha, em uma imagem que representa o ciclo do caminhar dela. Para analisar os dados transformados em imagem, foi utilizado a técnica de Content-Based Image Retrieval (CBIR) através da biblioteca Lire.

Nos experimentos realizados, foram capturados dados de 50 pessoas caminhando em direção ao sensor (33 homens e 17 mulheres). Cada pessoa foi capturada 5 vezes, criando 5 exemplos independentes. O caminho em que a pessoa deveria caminhar foi demarcado utilizando 2 linhas brancas. A pessoa começava a caminhar fora da área de cobertura do sensor e no momento em que ela estava dentro da área os dados da pessoa começavam automaticamente a serem gravados quando o Kinect identificava todos os pontos do esqueleto dela, isto foi feito objetivando que a pessoa adquira um ritmo natural de caminhar antes de ser capturado pelo sensor.

Para testar a eficácia dos dados capturados, foi utilizado as técnicas de *Genuine Acceptance Rate* (GAR) e *False Match Rate* (FMR). Na técnica GAR/FAR cada exemplo é comparado independentemente com todos os outros do banco de dados, e o resultado mostrou o potencial do trabalho proposto, sendo possível obter um GAR entre 80% e 85% com um FAR entre 10% e 15%.

AHMED; PAUL; GAVRILOVA (2015) capturaram 25 pontos do esqueleto humano através do Kinect 2 e aplicaram duas técnicas, Joint Relative Distance (JRD) e Joint Relative Angle (JRA), que calculam distâncias de pontos do corpo e os ângulos entre elas, para representar o caminhar humano. Os dados fornecidos pelo JRD e JRA foram aplicados ao algoritmo Dynamic Time Warping (DTW), que gerou como saída a diferença entre cada exemplo e a identificação da pessoa capturada.

Nos testes realizados, foi utilizado uma base de dados contendo 20 pessoas (14 homens e 6 mulheres), para cada indivíduo 3 vídeos foram capturados. Cada cena de vídeo contém a pessoa entrando na sala (de frente para o sensor), caminhando em direção a uma cadeira, e sentando nela. Foi utilizada validação cruzada com 3 subconjuntos, como o banco de dados possui 3 vídeos por pessoas, foram utilizados 2 vídeos para treinamento e 1 para teste. Foi possível obter uma taxa de acerto de 91%

KASTANIOTIS et al. (2015), utilizaram as características dinâmicas do caminhar para identificar as pessoas. Inicialmente, os dados capturados pelo Kinect são processados e um grupo de características é extraído, resultando em uma representação apropriada. As características que formam este grupo são a direção de ambos os braços e pernas, que são codificadas utilizando dois ângulos Euler, conforme Figura 27:

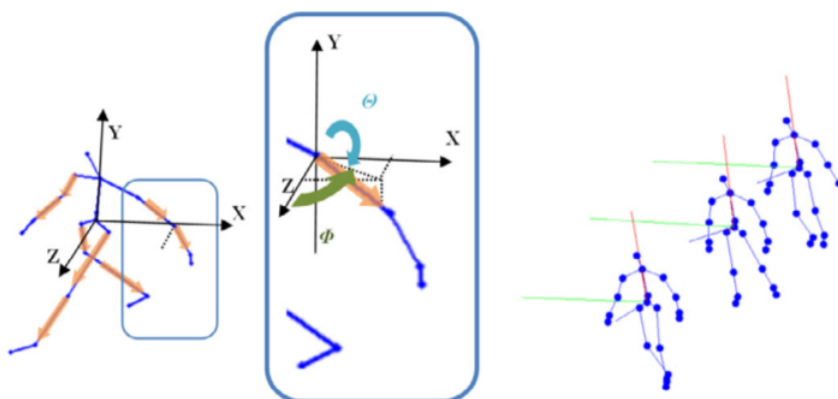


Figura 27 – Ângulos Euler utilizados para identificar o indivíduo. Fonte: (KASTANIOTIS et al., 2015).

Em cima da representação codificada através dos ângulos Euler, é calculado um valor de dissimilaridade, entre ela e outras representações de um conjunto de treinamento, utilizando o algoritmo Wald-Wolfowitz (WW). Este valor de dissimilaridade é transformado em uma representação de dissimilaridade esparsa para poder ser classificado.

Nos testes, foi utilizado uma base com 30 pessoas, 15 homens e 15 mulheres, com idade variando entre 23 e 55 anos. A Figura 27 também ilustra o ângulo de captura, o sensor estava colocado a 1,70m do chão, a esquerda da linha onde as pessoas caminharam, com um ângulo de 30° relativos a essa linha. As pessoas foram instruídas a caminharem em uma direção reta, em sua velocidade normal, sem o auxílio de algo delimitando o caminho. Foram capturadas 5 sequências para cada pessoa, variando de 55 a 120 quadros cada sequência. Foi possível obter

uma taxa de acerto de 93% na identificação das pessoas.

CHATTOPADHYAY; SURAL; MUKHERJEE (2015), utilizaram a informação de profundidade e a localização dos pontos do esqueleto, fornecida pelo Kinect, para resolver o problema de identificação da pessoas através do caminhar quando existe algo bloqueando parte da visão do sensor. Nesta pesquisa foram utilizados dois sensores Kinect, um capturando a pessoa de frente, do qual foram obtidos os dados relativos aos pontos do esqueleto, e o outro de costas, do qual foram extraídos dados de profundidade.

Do sensor que estava capturando a parte de frente da pessoa, objetivando formar as características para identificar as pessoas, foram analisados os movimentos de 5 pontos do esqueleto da pessoa: centro do quadril, parte esquerda do quadril, ambos joelhos e ambos pés.

Também, a partir destes 5 pontos do esqueleto, foi computado todas as combinações possíveis contendo 3 pontos. Destas combinações encontradas, foi calculado a curvatura e o centro do círculo passando através desses 3 pontos, e esses dados também foram utilizados como características.

Para auxiliar na identificação da pessoa, foram utilizadas características baseadas na silhueta da pessoa, que foram obtidas através dos dados de profundidade do sensor que estava capturando a parte de trás da pessoa.

Nos captura dos dados, foram utilizadas 29 pessoas e os sensores Kinect foram posicionados a uma altura de 2.5 metros com um ângulo de -27 graus.

Para montar a base de exemplos, foi utilizado somente o ângulo frontal e a captura ocorreu com velocidade de 30 quadros por segundo.

Para montar a base de testes, os dados foram obtidos utilizando sequências capturadas através de dois ângulos: frente (F) e atrás (A), duas velocidades do caminhar: normal (CN) e rápida (CR), duas situações de oclusão: com oclusão (CO) e sem oclusão (SO) e com duas velocidades de quadros por segundo, 30 fps (F1) e 15 fps (F2), a sigla FPS refere-se a *frames per second* (quadros por segundo, do inglês).

Para cada fps F1 e F2, existem quatro conjuntos para cada ângulo de captura (frente e atrás), correspondendo a cada um dos 29 indivíduos. A base completa possui, portanto, 232 sequências para cada fps F1 e F2.

Em cada teste é utilizado um par de sequências, relativos a visão de frente e a visão de trás, pois os dados de ambos os sensores são aplicados em conjunto para a classificação.

Utilizando validação cruzada com 3 subconjuntos e uma métrica de similaridade, foi possível obter uma taxa de acerto de 78.8%.

CHATTOPADHYAY et al. (2013), utilizaram as informações de profundidade e cor fornecidas pelo Kinect, para extrair características do caminhar da pessoa, capturando o ângulo frontal dela. Cada pixel das informações de profundidade foram mapeadas de forma a se encaixarem nos pixels capturados através da câmera RGB, sendo possível obter a silhueta da pessoa formada pelos dados de cor e profundidade, este método foi nomeado Pose Depth

Volume (PDV).

Para identificar a pessoa, os dados de PDV foram comparados através da distância Euclidiana com a base de exemplos. Para analisar a eficácia do método, foi utilizada uma base contendo 30 pessoas, o Kinect foi colocado a uma altura de 2m em um ângulo de -27 graus. As capturas foram feitas em diferentes quadros por segundo, sendo 10, 15, 20, 25 e 30 fps. Um dos ciclos da pessoa foi usado para treinamento e os outros para testes.

Para cada indivíduo, foram capturados 4 conjuntos de exemplos, em 3 deles (denotados como Ca, Cb, e Cc) cada pessoa usava a mesma roupa em cada captura, embora entre indivíduos diferentes as roupas fossem diferentes, e em 1 conjunto (denotado Cd) as pessoas vestiam roupas diferentes a cada captura.

Como resultado, verificou-se que através deste método foi possível obter 80% de acurácia na identificação através do caminhar e que a mudança de roupa não trouxe mudanças significativas na acurácia.

MUNSELL et al. (2012), da Carolina do Sul (EUA), desenvolveu uma pesquisa visando a identificação biométrica utilizando dados capturados do corpo inteiro da pessoa, através de análise do movimento do caminhar ou correr. O ângulo utilizado para capturar os dados foi a lateral da pessoa.

Este trabalho ao analisar os dados capturados pelo Kinect, ao contrário do método de pesquisa tradicional, não tentou isolar o ciclo do caminhar, porém ao analisar todos os dados capturados procurou identificar os dados que tinham maior relevância para o reconhecimento da ação sendo executada.

Antes de identificar a pessoa, ocorre um estágio onde é utilizado o algoritmo Support Vector Machine (SVM) para identificar se ação que está sendo realizada é o caminhar ou o correr. Após reconhecida a ação, os dados são enviados para um classificador que analisa padrões de movimentos e medidas antropométricas para identificar a pessoa.

O conjunto de treinamento é composto por 10 pessoas (6 homens e 4 mulheres), com idade variando entre 25 e 40 anos, onde cada pessoa realiza duas vezes a ação de correr e duas vezes a ação de caminhar, totalizando 40 vídeos.

Para os testes, o conjunto formado utilizou as mesmas 10 pessoas, porém foi capturado para cada pessoa 3 vezes a ação de correr e 3 vezes a ação de caminhar, totalizando 60 vídeos. Dentro das 3 capturas de cada ação, em uma delas a pessoa carregava uma mochila, na outra a pessoa usava a mesma mochila, porém carregava um objeto na mão direita e por fim, na outra, a pessoa se deslocava entrelaçando os pés, formando um "S".

Como resultado dos testes, foi verificado um ROC EER médio de 13% e um CMC Rank-1 médio de 90%.

Os testes inter-classes foram feitos comparando exemplos da classe das pessoas caminhando em velocidade normal com a classe das pessoas caminhando rapidamente. Tanto nos testes inter-classes como com os utilizando a mesma classe, foi possível obter taxa de acerto de 100%.

PROCHÁZKA et al. (2015) desenvolveram um trabalho, que através da análise do caminhar, procura diagnosticar doenças de coordenação motora, como o mal de Parkinson.

Para a captura dos dados, o sensor Kinect foi instalado a uma altura de mais ou menos 60 cm acima do chão. Foram utilizados 3 grupos de indivíduos na aquisição dos dados:

1. 18 pacientes, com idade entre 52 e 87 anos e com a doença de Parkinson;
2. 18 pacientes saudáveis, com idade entre 32 e 81 anos;
3. 15 estudantes saudáveis, com idade entre 23 e 25 anos.

O sensor foi utilizado capturando a lateral das pessoas. Cada indivíduo foi capturado 5 vezes, onde em cada captura a pessoa andou cerca de 4 metros (5 passos).

A seguir são apresentados as taxa de acertos na identificação da doença de Parkinson nas pessoas, onde foram aplicadas aos algoritmos de aprendizado de máquina as seguintes características de cada indivíduo: **a** refere-se a largura do passo, **b** é a velocidade do caminhar e **c** é a idade.

Tabela 1 – Resultado na identificação do mal de Parkinson

Algoritmo	Taxa de Acerto (%)		
	a e b	b e c	a e c
Rede Bayesiana	92.2	92.2	94.1
Decision Tree	94.1	92.2	94.1
3-NN	90.2	92.2	92.2
7-NN	90.2	86.3	86.3
10-NN	90.2	84.3	84.3

RAHEJA et al. (2014) desenvolveram um trabalho no sentido de analisar a anormalidade do caminhar, em ambiente hospitalar, de pacientes com esclerose múltipla. As características utilizadas como propriedades do indivíduo foram os ângulos, a velocidade e a profundidade de ambas as mãos e pernas e o ângulo de captura foi o frontal.

Os dados foram aplicados ao algoritmo SVM para classificar cada indivíduo como com saúde estável ou necessitando de atendimento. O método foi testado utilizando 75 pessoas, sendo 40 delas saudáveis. Das 35 pessoas que não eram saudáveis, apenas 1 foi classificada como normal e das 40 pessoas saudáveis, apenas 2 foram classificadas como tendo um caminhar anormal, resultando em 96% de acurácia ao todo.

GHOLAMI et al. (2015), em sua pesquisa, trabalharam na análise do caminhar em ambiente hospitalar, buscando avaliar a qualidade da saúde dos pacientes. O ângulo de captura das pessoas foi o frontal e cada pessoa caminhou de 5 a 10 vezes, onde foram selecionados as 5 melhores capturas. As seguintes características, extraídas através do caminhar, foram utilizadas para avaliar os pacientes:

- Características temporais: velocidade do caminhar, comprimento do passo, distância percorrida em um ciclo e a porcentagem de tempo em que a pessoa está com os dois pés em contato com o solo;
- Ângulo das juntas: refere-se a amplitude do movimento do quadril e joelhos;
- Padrão do caminhar: Obtido ao aplicar o algoritmo Dynamic Time Warping (DTW) aos dados referentes as posições em que o quadril e os joelhos estiveram enquanto a pessoa caminhava;

No estudo proposto, foram utilizadas 10 pacientes com desordem neurológica (esclerose múltipla) e 10 pessoas saudáveis. No dia da captura dos dados, foi fornecida duas pontuações, dadas por um médico, para cada paciente com esclerose múltipla, uma pontuação refere-se ao grau da esclerose e a outra com relação a uma análise clínica que observa o caminhar da pessoa.

Estes dados foram comparados com as informações obtidas através do Kinect e verificou-se grande similariedade entre ambos, onde as características avaliadas com o Kinect se atenuavam de modo similar as pontuações dadas pelo médico. As exceções ocorreram nas características referente a porcentagem de tempo em que a pessoa está com os dois pés em contato com o solo, durante o caminhar, e com relação ao comprimento do passo.

Comparando as características obtidas com os dados do Kinect de pacientes saudáveis e pacientes com esclerose múltipla, verificou-se grande diferença entre ambos e com a aplicação do teste t pareado obteve-se valores p menores que 0.05.

2.4.1.1 Resumo das pesquisas de análise do caminhar

Nesta subseção, são apresentadas diversas tabelas sumarizando as características e os resultados dos trabalhos discutidos nesta seção e no capítulo referente ao estado da arte.

Entre as características apresentadas, estão o ângulo de captura, o algoritmo utilizado para classificação, a versão do Kinect e a quantidade de pessoas utilizadas nas capturas (especificando a quantidade em cada gênero, quando citado no trabalho).

A tabela 2 apresenta o resumo dos ângulos em relação a pessoa, das capturas através do Kinect, citadas nesta seção.

Na última linha da tabela 2, o trabalho de SIVAPALAN et al. (2011) é apresentado como Multi-ângulo por utilizar 6 sensores ao mesmo tempo para cobrir diversos ângulos da pessoa sendo capturada.

Tabela 2 – Ângulos das capturas realizadas através do Kinect

Ângulo	Qtd	Trabalhos
Frente	8	(CHATTOPADHYAY et al., 2013), (SIVAPALAN et al., 2011), (MILOVANOVIC; MINOVIC; STARCEVIC, 2013), (AHMED; PAUL; GAVRILOVA, 2015), (RAHEJA et al., 2014), (GHOLAMI et al., 2015), (ARAUJO; GRAÑA; ANDERSSON, 2013)
Lado	13	(ARORA et al., 2015a), (ANDERSSON, 2014), (ARAUJO; GRAÑA; ANDERSSON, 2013), (GUILLEN et al., 2009a), (PREIS et al., 2012), (MUNSELL et al., 2012), (SINHA; CHAKRAVARTY; BHOWMICK, 2013), (PROCHÁZKA et al., 2015), (VENKAT; WILDE, 2010), (YEOH et al., 2011), (YOO; NIXON, 2011), (WANG et al., 2005), (YOO JANG-HEE; MOON; HWANG DOOSUNG; NIXON, 2008)
Diagonal	1	(KASTANIOTIS et al., 2015)
Frente e atrás simultaneamente	1	(CHATTOPADHYAY; SURAL; MUKHERJEE, 2015)
Multi-ângulo	1	(SIVAPALAN et al., 2011)
Outros	1	(BENABDELKADER; CUTLER; DAVIS, 2015)

A tabela 3 mostra os algoritmos utilizados como classificador nos trabalhos citados, quando mais de um algoritmo foi utilizado foi contado aquele que apresentou o melhor resultado.

Tabela 3 – Algoritmos que obtiveram melhores resultados (* a última linha relaciona trabalhos que utilizaram algoritmos próprios de classificação.)

Algoritmo	Qtd	Trabalhos
DTW	1	(AHMED; PAUL; GAVRILOVA, 2015)
GAR/FMR	1	(MILOVANOVIC; MINOVIC; STARCEVIC, 2013)
GEV/Boul-gouris	1	(SIVAPALAN et al., 2011)
KNN	4	(ARORA et al., 2015a), (ANDERSSON, 2014), (ARAUJO; GRAÑA; ANDERSSON, 2013), (YOO; NIXON, 2011)
Fuzzy KNN	1	(YEOH et al., 2011)
MLP	2	(SINHA; CHAKRAVARTY; BHOWMICK, 2013), (YOO JANG-HEE; MOON; HWANG DOOSUNG; NIXON, 2008)
Rede Bayesiana	3	(BENABDELKADER; CUTLER; DAVIS, 2015), (PREIS et al., 2012), (PROCHÁZKA et al., 2015)
SVM	2	(CHATTOPADHYAY et al., 2013), (RAHEJA et al., 2014)
*Outros	8	(KASTANIOTIS et al., 2015), (CHATTOPADHYAY et al., 2013), (CHATTOPADHYAY; SURAL; MUKHERJEE, 2015), (GHOLAMI et al., 2015), (GUILLEN et al., 2009a), (MUNSELL et al., 2012), (VENKAT; WILDE, 2010), (WANG et al., 2005)

A tabela 4 sumariza a versão do Kinect utilizada nos trabalhos. Observa-se através dos dados apresentados a seguir, que existe uma lacuna na pesquisa com a versão 2 do Kinect,

mostrando a importância de novas pesquisas com este sensor.

Tabela 4 – Versão do Kinect utilizada nos trabalhos

Kinect	Qtd	Trabalhos
Versão 1	18	(ARORA et al., 2015a), (MILOVANOVIC; MINOVIC; STARCEVIC, 2013), (SIVAPALAN et al., 2011), (ANDERSSON, 2014), (ARAUJO; GRAÑA; ANDERSSON, 2013), (SINHA; CHAKRAVARTY; BHOWMICK, 2013), (PREIS et al., 2012), (PROCHÁZKA et al., 2015), (CHATTOPADHYAY et al., 2013), (RAHEJA et al., 2014), (KASTANIOTIS et al., 2015), (CHATTOPADHYAY; SURAL; MUKHERJEE, 2015), (GHOLAMI et al., 2015), (GUILLEN et al., 2009a), (MUNSELL et al., 2012), (VENKAT; WILDE, 2010), (YEOH et al., 2011), (WANG et al., 2005)
Versão 2	1	(AHMED; PAUL; GAVRILOVA, 2015)

As tabelas 5 e 6 resumem a acurácia obtida nos trabalhos, também apresentam a quantidade de indivíduos utilizados nas capturas em cada pesquisa, especificando também o gênero dos indivíduos, o algoritmo utilizado, o método de análise e a versão do Kinect.

A tabela 5 relaciona os trabalhos que procuram identificar a pessoa através do caminhar e a tabela 6 relaciona os trabalhos que analisam o estado de saúde dos indivíduos. Quando o gênero das pessoas não é relacionado, ele não foi citado no trabalho. A coluna 'Md' relaciona se a técnica utilizada no trabalho é baseado em modelo (B) ou livre de modelo (L), explicações da diferença entre uma técnica e outra estão apresentadas na Seção 2.3.

A partir dos dados sumarizados, podemos verificar que para os trabalhos em que o ângulo de captura foi a lateral do indivíduo, a maior parte dos trabalhos utilizou algoritmos de aprendizado de máquina (KNN, MLP e Rede Bayesiana) para a classificação dos dados, enquanto que quando o ângulo de captura não era o lateral, a maior parte dos trabalhos utilizou outros tipos de algoritmos para classificação.

Também é possível observar que existe uma lacuna na pesquisa com a versão 2 do Kinect, mostrando a importância de novas pesquisas com este sensor. A pesquisa que utilizou a versão 2 do Kinect não utilizou algoritmos de aprendizado de máquina como classificador, além de apenas ter realizados testes com o ângulo frontal das pessoas. Isto mostra a importância de novas pesquisas com este sensor, objetivando estudar a diferença que ele traz na acurácia dos resultados comparado com a versão 1 do Kinect e a sua relevância no campo de pesquisa.

Verificou-se também que o ângulo de captura mais utilizado é o frontal (8 trabalhos), seguido do perpendicular (6 trabalhos).

O trabalho que apresentou a maior acurácia na identificação de pessoas utilizou o algoritmo GEV, com 100% de acurácia, seguida da pesquisa que utilizou o KNN, com acurácia de 96.6%. Porém deve-se ressaltar que a boa acurácia nestes trabalhos pode ter sido ao fato da base de exemplos destes trabalhos ter sido pequena, pois quanto menor é o número de exemplos a serem classificados, maior é a acurácia.

Tabela 5 – Resumo dos trabalhos na identificação do caminhar

Acurácia	Masc	Fem	Total	Algoritmo	Ângulo	Vs.	Md	Autor
100%	-	-	15	GEV/Boulogouris	Frente ou multi-ângulo	1	L	(SIVAPALAN et al., 2011)
99,6%	6	2	8	KNN	Frente e Lado	1	B	(ARAUJO; GRAÑA; ANDERSSON, 2013)
97,5%	-	-	20	Próprio	Lado	1	B	(WANG et al., 2005)
96.7%	-	-	30	Próprio	Lado	-	B	(YOO; NIXON, 2011)
95%	-	-	100	Próprio	Lado	1	L	(GUILLEN et al., 2009a)
94.1%	-	-	34	KNN	Lado	1	L	(ARORA et al., 2015a)
93%	15	15	30	Próprio	Diagonal	1	B	(KASTANIOTIS et al., 2015)
91%	14	6	20	DTW	Frente	2	B	(AHMED; PAUL; GAVRILOVA, 2015)
91%	-	-	8	Bayes	Lado	1	B	(PREIS et al., 2012)
90%	6	4	10	Próprio	Lado	1	B	(MUNSELL et al., 2012)
90%	-	-	122	Próprio	Lado	1	L	(VENKAT; WILDE, 2010)
90%	-	-	30	MLP	Lado	-	B	(YOO JANG-HEE; MOON; HWANG DOO-SUNG; NIXON, 2008)
87,96%	-	-	11	Fuzzy KNN	Lado	1	B	(YEOH et al., 2011)
87,74%	51	64	133	KNN	Lado	1	B	(ANDERSSON, 2014)
86%	8	2	10	MLP	Lado	1	B	(SINHA; CHAKRAVARTY; BHOWMICK, 2013)
85%	33	17	50	GAR/FMR	Frente	1	L	(MILOVANOVIC; MINOVIC; STARCEVIC, 2013)
80%	-	-	30	Próprio	Frente	1	B	(CHATTOPADHYAY et al., 2013)
78,16%	-	-	29	Próprio	Frente e atrás juntos	1	L	(CHATTOPADHYAY; SURAL; MUKHERJEE, 2015)
40%	-	-	17	Rede Bayesiana	Superior diagonal	-	B	(BENABDELKADER; CUTLER; DAVIS, 2015)

Tabela 6 – Resumo dos trabalhos na análise da saúde

Acurácia	Masc	Fem	Total	Algoritmo	Ângulo	Vs.	Md	Autor
96%	-	-	75	SVM	Frente	1	B	(RAHEJA et al., 2014)
92.2%	-	-	51	Bayes	Lado	1	B	(PROCHÁZKA et al., 2015)
	2	18	20	Próprio	Frente	1	B	(GHOLAMI et al., 2015)

Portanto analisando os resultados, Notou-se uma lacuna na utilização do algoritmo de classificação Bayesiana e no algoritmo MLP, pois os trabalhos que utilizaram estes algoritmos apenas abordaram o ângulo lateral dos indivíduos.

Os resultados dos trabalhos sugerem que algoritmos baseados em modelo possuam melhor acurácia do que os livre de modelo, com exceção do trabalho de SIVAPALAN et al. (2011), que utilizou o algoritmo GEV e alcançou 100% de acurácia, os demais não apresentaram resultados acima de 90%.

Os algoritmos GEV e GAR/FMR foram utilizados apenas no ângulo frontal, porém esta utilização justifica-se por estes trabalhos utilizarem abordagens livre de modelo, onde o ângulo frontal é importante para evitar ambiguidade de membros do corpo.

2.5 Algoritmos de Aprendizado de Máquina

Como visto nos capítulos anteriores, diversas características podem ser extraídas através do caminhar do ser humano. E conforme foi apresentado, quando uma pessoa caminha ela desenvolve um ciclo neste caminhar e dentro deste ciclo existem padrões. Os algoritmos de aprendizado de máquina se beneficiam destas características e padrões para aprenderem a como identificar as pessoas.

As técnicas de aprendizado de máquina empregam um princípio de inferência chamado de indução, no qual obtém-se conclusões genéricas a partir de um conjunto particular de exemplos (LORENA; CARVALHO, 2015).

Neste capítulo serão apresentados os algoritmos de aprendizado de máquina que foram utilizados neste trabalho: K-nearest Neighbor (KNN), Random Forest, Rede Bayesiana Naive Bayes e Rede Neural Multi-layer Perceptron (MLP). A utilização do KNN, Naive Bayes e MLP justifica-se devido a trabalhos relacionados utilizarem estes algoritmos e a utilização do Random Forest se justifica por ter sido a escolha da Microsoft para ser implementada no *software* do Kinect com o fim de obter o esqueleto dos indivíduos (SHOTTON et al., 2011).

2.5.1 K-nearest Neighbor

O K-nearest Neighbor é um conhecido algoritmo de aprendizado de máquina. Ele é um algoritmo bastante simples se comparado com outros e exige pouco processamento computacional.

O KNN utiliza como base de dados exemplos que estão separados em classes para classificar a entrada de um novo exemplo. O novo exemplo é classificado de acordo com a classe que mais se parece com ele. O modo como o algoritmo decide quais dos exemplos na base é mais parecido com o novo é se basear nos k exemplos mais próximos (SUTTON, 2015). Se k for igual a 1, então a classe do exemplo mais parecido será atribuída ao novo exemplo, se k for igual a 2, a classe dos dois exemplos mais parecidos será atribuída ao novo exemplo, e assim por diante, o k pode ser configurado para qualquer valor.

O KNN é geralmente baseado na distância Euclideana entre o novo exemplo e os k exemplos mais próximos. Supondo que x_i seja um exemplo de entrada com as características $p(x_{i1}, x_{i2}, \dots, x_{ip})$, n sendo o total de exemplos na base ($i = 1, 2, \dots, n$) e p o número total de características, a distância Euclideana entre x_{i1} e x_{il} ($l = 1, 2, \dots, n$) é definida por (PETERSON, 2009): $d(x_i, x_l) = \sqrt{(x_{i1} - x_{l1})^2 + (x_{i2} - x_{l2})^2 + \dots + (x_{ip} - x_{lp})^2}$

A Figura 28 mostra um gráfico X e Y com 19 exemplos para serem comparados a um novo exemplo. No exemplo desta figura, o k está configurado para o valor 4. Os 4 exemplos mais parecidos são comparados ao exemplo de entrada, e o exemplo de entrada será classificado com a classe dos exemplos em vermelho, por ser a classe mais frequente próxima ao exemplo de entrada (PETERSON, 2009):

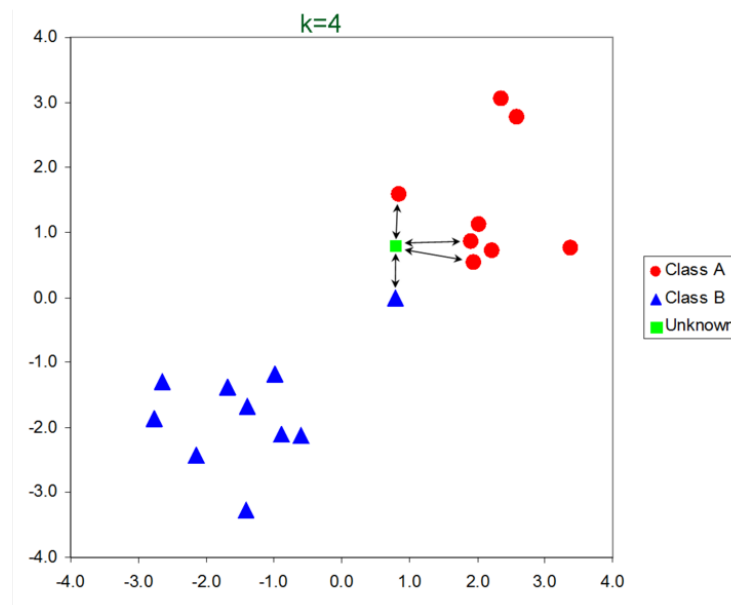


Figura 28 – Ilustração do funcionamento do KNN. Fonte: (PETERSON, 2009).

Por vezes, com o fim de melhorar a precisão e também de evitar empate no cálculo dos exemplos mais parecidos, utiliza-se média ponderada na equação, onde os exemplos mais próximos possuem peso maior.

2.5.2 Random Forest

Para ser possível entender o algoritmo Random Forest, é preciso antes compreender outro algoritmo de aprendizado de máquina, o algoritmo de Árvores de Decisão, pois o Random

Forest utiliza-se de árvores de decisão.

As Árvores de Decisão são treinadas de acordo com um conjunto de treino (exemplos que são classificados previamente) e depois outros exemplos são classificados de acordo com essa mesma árvore (GFBIOINFO, 2017).

Para ilustrar uma árvore de decisão, vamos utilizar uma conhecida ilustração formulada por MITCHELL (1997): supomos que o objetivo seja decidir se a pessoa vai jogar tênis. Para isto, alguns parâmetros devem ser levados em conta, como o aspecto do céu, a temperatura, a umidade e o vento. Cada um destes atributos possuem diversos valores, por exemplo, para a temperatura pode-se ter os valores ameno, fresco, quente, entre outros. A decisão se a pessoa irá ou não jogar tênis é o resultado da classificação de acordo com estes parâmetros. (GFBIOINFO, 2017).

Supomos que para construir esta árvore de decisão para jogar tênis, são utilizados exemplos de dias anteriores, conforme a tabela 7:

Tabela 7 – Exemplo de dados para árvore de decisão (GFBIOINFO, 2017).

Dia	Aspecto	Temp.	Humidade	Vento	Jogar Tênis
D1	Sol	Quente	Elevada	Fraco	Não
D2	Sol	Quente	Elevada	Forte	Não
D3	Nuvens	Quente	Elevada	Fraco	Sim
D4	Chuva	Ameno	Elevada	Fraco	Sim
D5	Chuva	Fresco	Normal	Fraco	Sim
D6	Chuva	Fresco	Normal	Forte	Não
D7	Nuvens	Fresco	Normal	Fraco	Sim
D8	Sol	Ameno	Elevada	Fraco	Não
D9	Sol	Fresco	Normal	Fraco	Sim
D10	Chuva	Ameno	Normal	Forte	Sim
D11	Sol	Ameno	Normal	Forte	Sim
D12	Nuvens	Ameno	Elevada	Forte	Sim
D13	Nuvens	Quente	Normal	Fraco	Sim
D14	Sol	Ameno	Elevada	Forte	Não

A árvore de decisão, de acordo com os exemplos apresentados na tabela anterior, ficaria desta forma:

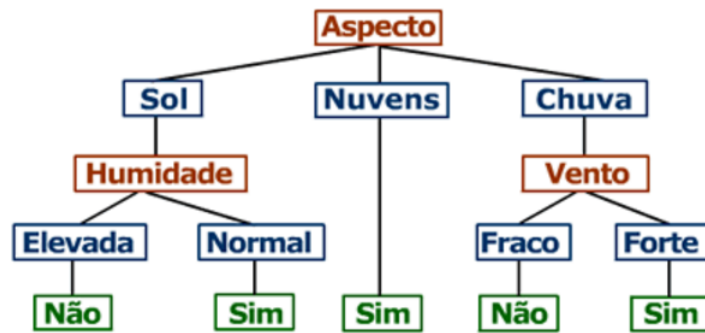


Figura 29 – Árvore de decisão para jogar tênis. Fonte: (GFBIOINFO, 2017).

A relação entre os elementos da árvore e os atributos, valores e classificações pode ser entendida na Figura 30:

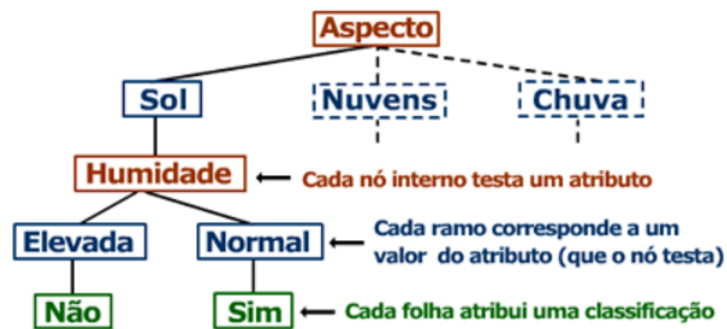


Figura 30 – Relações dos elementos da árvore de decisão para jogar tênis. Fonte: (GFBIOINFO, 2017).

Random Forest é um algoritmo de aprendizado de máquina que constrói árvores de decisão para classificar objetos (CENTRAL, 2017). Embora existam outros cálculos envolvidos, basicamente o que o algoritmo Random Forest faz é construir várias árvores de decisão e atribuir a classificação de acordo com o resultado que mais predominou entre as árvores, ou seja, cada árvore apresenta um resultado e aquele resultado que mais apareceu é o escolhido para classificar o novo exemplo (BREIMAN; CUTLER, 2017).

2.5.3 Rede Bayesiana

A rede bayseana trabalha com probabilidades, sendo assim através deste algoritmo é possível saber qual a probabilidade de ocorrência de cada valor na saída. Richard Neapolitan usa como ilustração para a rede bayseana um caso onde deseja-se saber se uma pessoa tem ou não uma determinada doença (NEAPOLITAN, 2004).

Vários fatores podem influenciar a decisão, e a rede Bayesiana utiliza-se de probabilidades para medir a importância de cada fator. Além de mostrar quais fatores tem influência direta entre si. Ele ilustra um caso onde tem-se o conhecimento se o indivíduo tem histórico de fumo, bronquite, fadiga, câncer na garganta e se o raio-x mostrou a presença da doença.

Neste exemplo, ele mostrou que os fatores tem diferentes relações entre si, por exemplo, se o paciente tem histórico de fumo isso tem uma ligação direta se ele tem bronquite e câncer de garganta. A rede bayseana então cria um mapa das relações entre cada fator e da a probabilidade dele ter a doença conforme a presença ou não de cada fator (NEAPOLITAN, 2004).

O algoritmo utilizado neste trabalho é o Naive Bayes, o qual é baseado no teorema da rede Bayesiana (MICROSOFT, 2017c).

2.5.4 Rede Neural Artificial

Segundo Haykin, “na sua forma geral, uma rede neural é uma máquina que é projetada para modelar a maneira como o cérebro realiza uma tarefa particular ou função de interesse” (HAYKIN, 2001, p. 28). As redes neurais artificiais tentam processar as informações de forma semelhante ao nosso cérebro, utilizando-se de neurônios que transmitem e recebem informações de outros neurônios. No aprendizado de máquina, os neurônios vão se ajustando conforme a experiência para melhor realizar determinadas tarefas. O neurônio artificial, de acordo com Haykin, possui pelo menos três elementos básicos, como (HAYKIN, 2001):

1. Um conjunto de sinapses ou elos de conexão, cada um caracterizado por um peso específico, que trabalham recebendo ou transmitindo informações (na rede neural artificial temos elos de conexão específicos para receber informações e outros para transmitir);
2. Um somador que soma os sinais de entrada;
3. Uma função de ativação que restringe a amplitude de saída do neurônio.

Conforme a rede neural vai ganhando experiência, ou seja, vai sendo treinada, os pesos das conexões sinápticas vão sendo ajustados para melhor resolver os problemas.

A Figura 31 ilustra um neurônio da rede neural artificial.



Figura 31 – Um neurônio artificial. Fonte: (GERSHENSON, 2015).

Diversas são as aplicações para as redes neurais, como reconhecimento de padrões, previsão, compressão de dados, entre outros. (GERSHENSON, 2015)

Um dos tipos mais famosos de treinamento de rede neural é a *backpropagation*, a qual é implementada na estrutura de rede neural MLP. A *backpropagation* parece ter sido inventada

originalmente por Paul Werbos em sua tese de doutorado em Harvard publicada no ano de 1974 e posteriormente reinventada no ano de 1982 por David Parker do *Stanford Linear Accelerator Center* (e reinventada várias vezes depois por diversas pessoas) (HECHT-NIELSEN, 1988). No tipo de treinamento backpropagation, são inseridos exemplos onde a saída é conhecida, e os pesos dos neurônios são ajustados até se parecerem ao máximo com àquela saída (ABERDEEN, 2015).

2.6 MICROSOFT KINECT

Nesta seção serão apresentadas as características de cada versão do sensor Kinect, incluindo informações sobre o *software* e o *hardware* de cada versão do sensor.

2.6.1 Kinect do Xbox 360

A Figura 32 apresenta a aparência do Kinect 360:



Figura 32 – Aparência do Kinect 360. Fonte: (CNET, 2017).

2.6.1.1 Hardware

O vídeo-game Xbox 360 foi o primeiro a vir com o sensor kinect, com um modelo desenvolvido especialmente para ele. O sensor Kinect proporcionou ao Xbox 360 algumas características, tais como (MICROSOFT, 2015a):

- Acessar o perfil de jogador apenas ficando em frente ao sensor;
- Permite interação com os jogos através de ações como pular, agachar e girar.
- Conectar-se facilmente com amigos através da identificação pelo sensor;

O sensor possui os seguintes componentes:

- Câmera RGB de 3 canais com resolução de 1280 x 960;
- Emissor infravermelho;

- Receptor infravermelho que recebe os feixes de luz do emissor para ser convertido em informações de profundidade com o fim de medir a distância entre os objetos e o sensor;
- Quatro microfones que conseguem identificar o local de origem do som;
- Motor para ajuste de ângulo;.

2.6.1.2 Software

Na inicialização, o sensor mapeia o ambiente e automaticamente ele ajusta as suas configurações para a melhor experiência possível. Após, ele procura por 48 pontos do corpo de cada jogador para criar uma réplica digital 3D necessária para a interação com o jogo (CANALTECH, 2015).

Através do mapeamento do jogador, o Kinect 360 consegue especificar a posição de 20 juntas do corpo humano, conforme Figura 33:

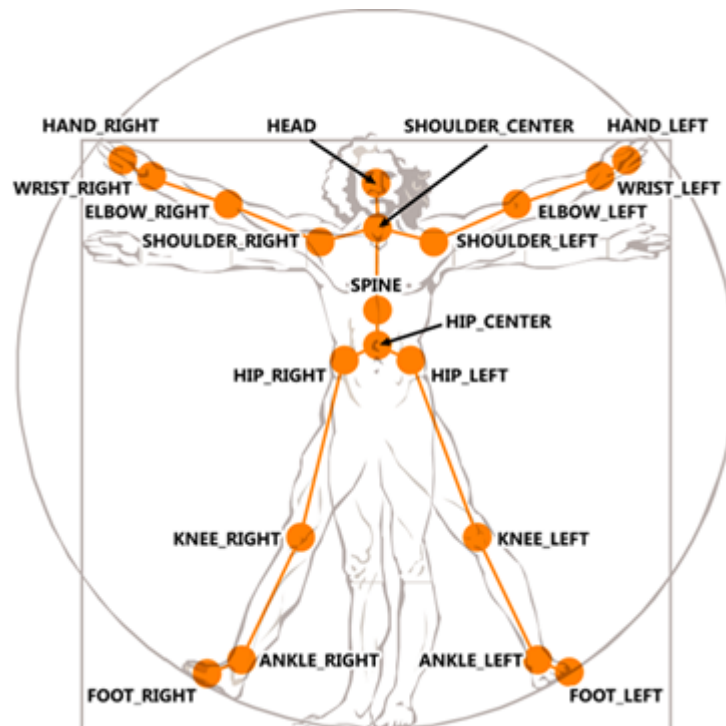


Figura 33 – Juntas mapeadas pelo Kinect 360. Fonte: (MICROSOFT, 2017b).

Foi utilizado aprendizado de máquina no software do Kinect. Os desenvolvedores do Kinect reuniram grande quantidade de dados de captura de movimento em cenários da vida real e utilizaram esses dados no algoritmo de aprendizado de máquina de Jamie Shotton, um pesquisador do Microsoft Research Cambridge, na Inglaterra. Foi possível mapear os dados capturados para modelos representando pessoas de diferentes idades, gênero, tipos de corpo, vestuário e classificar os movimentos desses modelos, enfatizando as juntas e as distâncias entre elas. (HOWSTUFFWORKS, 2015).

O software além de detectar e reagir ao que vê, também conseguiu distinguir os movimentos dos jogadores mesmo quando uma parte do corpo deles estiver parcialmente escondida,

pois o Kinect extrapola o que o resto do seu corpo está fazendo se ele puder detectar algumas partes dele (HOWSTUFFWORKS, 2015).

2.6.2 Kinect do Xbox One

Nesta seção serão apresentados dados técnicos e as melhorias do Kinect para o Xbox One, apresentando informações sobre o hardware e o software do novo sensor.



Figura 34 – Aparência do Kinect One. Fonte: (XBOX, 2015).

2.6.2.1 Hardware

A nova versão do sensor contém um *hardware* aprimorado, como por exemplo, o sensor de profundidade que possui maior capacidade de detectar objetos menores, uma melhor estabilidade de acompanhamento do corpo, visualização no escuro aprimorada e a câmera RGB que foi substituída por uma câmera de 1080p. (MICROSOFT, 2017d)

O processador e a memória do sensor também foram melhorados, foi utilizado o chip Microsoft X871141-001 e memória DDR3 SDRAM de 1GB.

O novo sensor possui os seguintes componentes:

- Câmera em cores de 1080p;
- Emissor e receptor infravermelho aprimorados;
- Quatro microfones que conseguem identificar o local de origem do som;
- Motor para ajuste de ângulo;

2.6.2.2 Software

Como no sensor anterior, o Kinect One também utiliza aprendizado de máquina no desenvolvimento de seu software (MICROSOFT, 2015c). Foram necessários 9 meses para modelar movimentos e ações de pessoas específicas, e treinar o sistema em como distinguir entre objetos e pessoas, quando movimentos começam e terminam e quais partes do corpo analisar quando calcular o movimento. Existe um enorme laboratório em Redmond onde duas

duzias de câmeras infravermelhas foram usadas para capturar movimentos específicos, como o balançar de um taco de baseball, por exemplo (EXTREMETECH, 2015).

O novo sensor Kinect possui melhorias com relação ao anterior, não somente em questões de hardware, mas também de software. O novo Kinect possui uma tecnologia chamada "*Time to Flight*", que consegue capturar o tempo que *photons* individuais levam para refletir em um objeto ou pessoa permitindo obter melhor precisão e acurácia. Por exemplo, o sensor possui precisão em entender que não apenas um polegar da mão está se mexendo, mas em qual direção ele está se movendo (VERGE, 2015).

O Kinect One consegue mapear 25 pontos do corpo humano, os 5 pontos a mais do que a versão anterior são: pescoço, dedos indicadores e polegares. A Figura 35 ilustra os pontos mapeados pelo Kinect One, na esquerda foi inserido os pontos especificados pelo Kinect 360 para efeitos de comparação:

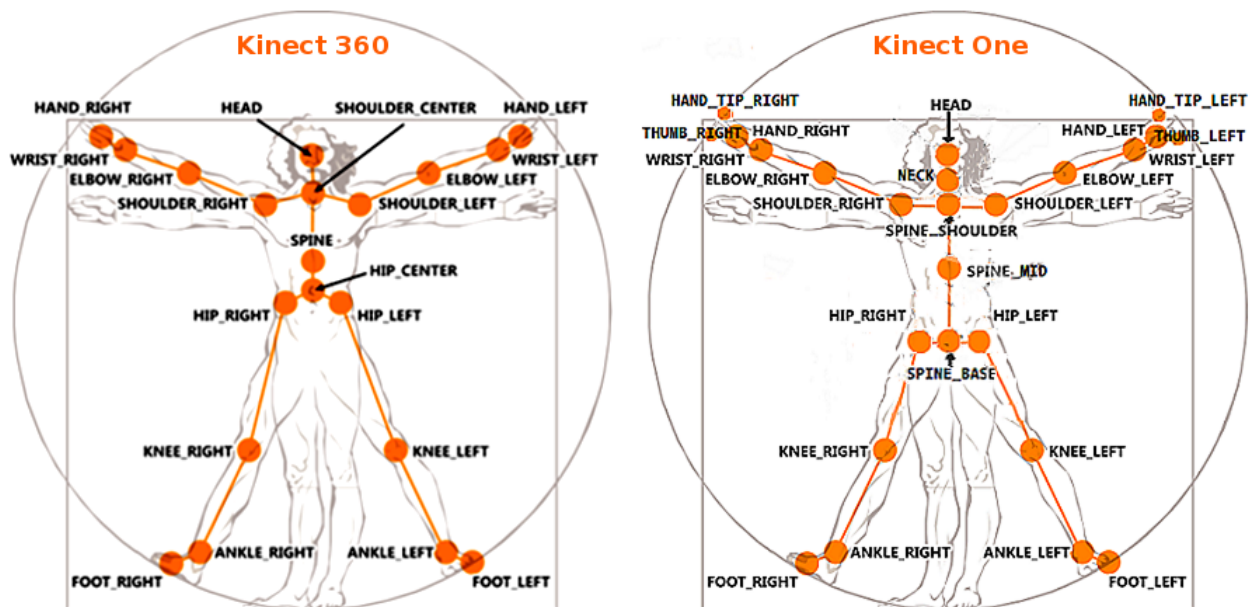


Figura 35 – Pontos mapeados pelos sensores Kinect. Fonte: (MASTER, 2017).

Na Figura 35 pode-se perceber, que além dos 5 pontos a mais, o Kinect One mapeia a região do quadril de forma diferente, pois além de trazer o quadril mais para baixo, o ponto da parte de baixo da coluna (*spine_base*) é alinhado na mesma altura do quadril.

3 OBJETIVOS E METODOLOGIA

3.1 Objetivos

A presente dissertação apresenta por objetivo central realizar um estudo comparativo entre a primeira e a segunda versão do Kinect na área da identificação de pessoas através do caminhar. Para realizar esta comparação serão aplicados diversos algoritmos de caracterização do caminhar. A utilização das duas versões do sensor Kinect, tem como objetivo preencher a lacuna que existe de pesquisas que mostrem a diferença nos resultados entre um sensor e outro na identificação das pessoas pelo modo como elas caminham.

Além do objetivo principal, este projeto possui como objetivos específicos:

1. Análise de quais atributos possuem maior contribuição na identificação pelo caminhar;
2. Propor melhorias nos processos existentes de caracterização dos indivíduos, através da utilização dos atributos que apresentam melhores resultados;
3. Analisar a diferença nos resultados de cada algoritmo de aprendizado de máquina e avaliar qual apresenta a melhor precisão;

3.2 Metodologia

Nesta seção será abordada a metodologia do trabalho que levou a identificação das pessoas. A Figura 36 ilustra o fluxograma do processo:

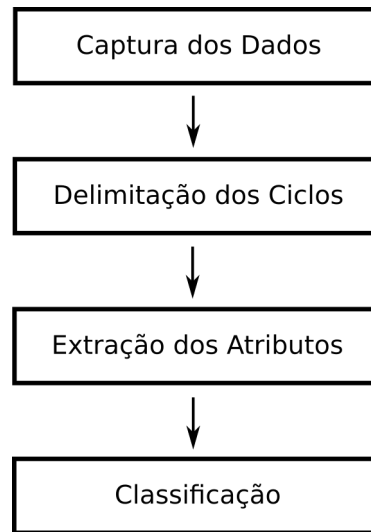


Figura 36 – Fluxograma da metodologia para identificação dos indivíduos.

Nas subseções seguintes é apresentada cada etapa do fluxograma da Figura 61, seguido do método de validação utilizado e da descrição do protocolo de captura dos indivíduos.

3.2.1 Captura

Foram coletados dados de 50 pessoas caminhando em frente aos sensores Kinect. O objetivo destas capturas de dados foi extrair as informações referentes ao caminhar dos indivíduos. Para que fosse possível realizar o estudo comparativo entre os sensores, no momento da gravação o sensor Kinect 360 foi colocado em cima do Kinect One, conforme ilustra a Figura 37:



Figura 37 – Posição dos sensores Kinect na captura dos dados.

A Figura 37 também apresenta a altura em que os sensores estavam do chão, o Kinect One estava em torno de 23,5 cm de altura e o Kinect 360 em torno de 30,5 cm de altura.

É possível também visualizar na Figura 37 que abaixo do suporte de sustentação dos sensores existe um disco. Este disco é giratório e permite girar os sensores durante as capturas de forma que eles fiquem sempre apontado para a pessoa que está sendo coletada. No presente trabalho a rotação do disco foi feita de forma manual, mas pode-se investigar formas de se implementar rotação automática para que os sensores fiquem sempre apontando para a pessoa que está sendo captura sem ser necessário alguém para girar o disco.

Outro ponto importante a destacar é o caminho percorrido pelas pessoas capturadas. As pessoas foram instruídas a caminharem 10 vezes em frente ao sensor, em uma trajetória semi-circular, partindo do ponto a esquerda dos sensores até o ponto a direita e depois retornando pelo mesmo caminho. O caminho que a pessoa deveria caminhar foi demarcado no chão para que a pessoa mantivesse uma distância de 2.40m dos sensores.

A Figura 38 ilustra o caminho feito pelas pessoas:

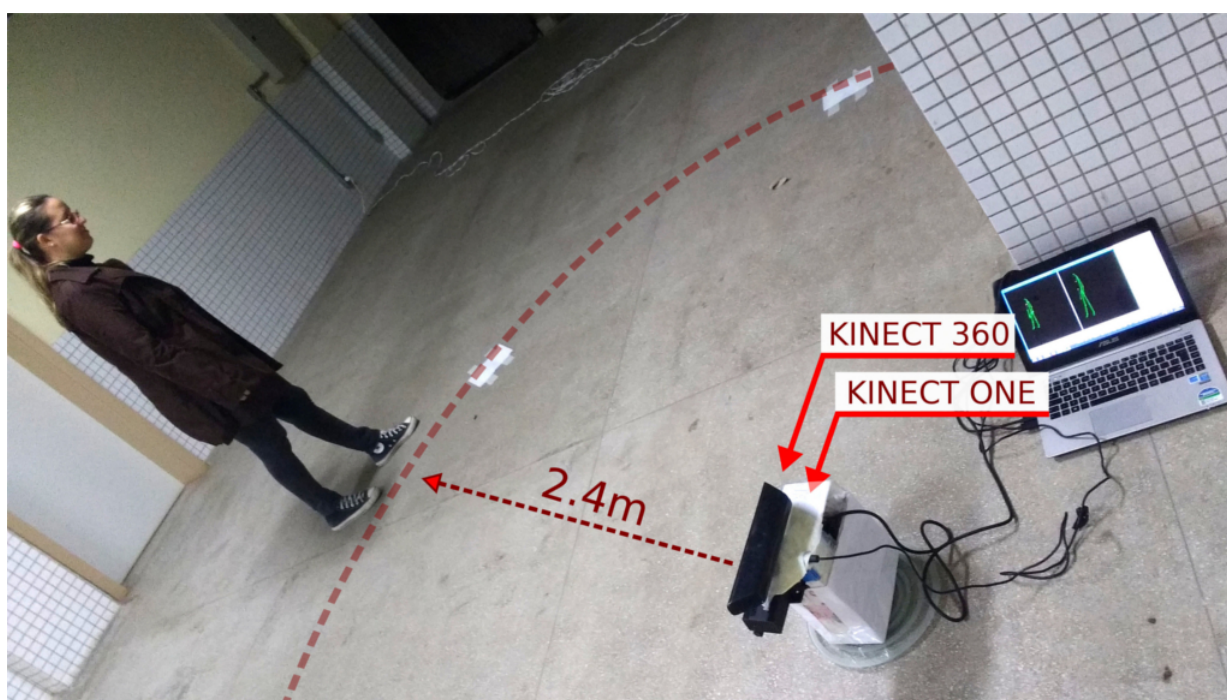


Figura 38 – Trajeto percorrido pelas pessoas nas capturas.

3.2.1.1 Alinhamento de FPS

A quantidade de quadros por segundo ou FPS (*frames per second*) das capturas realizadas ficou em torno de 20 quadros por segundo. Normalmente as capturas realizadas com os sensores Kinect ficam em 30 quadros por segundo, porém devido a este trabalho se propôr a realizar um estudo comparativo dos sensores e para isto ser necessário os dois Kinects estarem capturando a pessoa ao mesmo tempo, o computador que processava as capturas não conseguiu manter os 30 fps.

Foi necessário ter o cuidado de não permitir que o *software* que captura os dados dos sensores gravasse as informações de cada sensor em FPS diferentes um do outro. Para realizar o alinhamento de FPS entre os sensores o *software* foi configurado de forma que cada vez que um sensor enviava um pacote de dados, ele era gravado no arquivo somente se o outro sensor também já havia enviado um pacote, ou seja, se por exemplo, o Kinect 360 envia um pacote de dados, ele somente será gravado após o Kinect One enviar um pacote também, qualquer pacote enviado pelo sensor Kinect 360 entre este tempo de espera pelos dados do Kinect One é dispensado.

O fato da transmissão dos dados ter ficado em torno de 20 quadros por segundo não deve prejudicar os resultados de comparação dos sensores, pois mesmo que não houve-se esta redução de FPS, ambos teriam a mesma taxa, pois o número padrão de FPS de ambos os Kinects é 30.

3.2.1.2 Local de captura

Os dados foram capturados no campus Anglo da Universidade Federal de Pelotas. As capturas foram realizadas no final de um dos corredores de forma que fosse possível convidar pessoas para serem gravadas pelos sensores.

O espaço disponível era amplo de forma que foi possível traçar um trajeto de semi-círculo no chão, com uma distância de 2.4 metros dos sensores Kinect.

A figura 39 ilustra a distância percorrida pelas pessoas em relação ao sensor Kinect e também ilustra os trajetos da pessoa caminhando para a direita e para a esquerda.

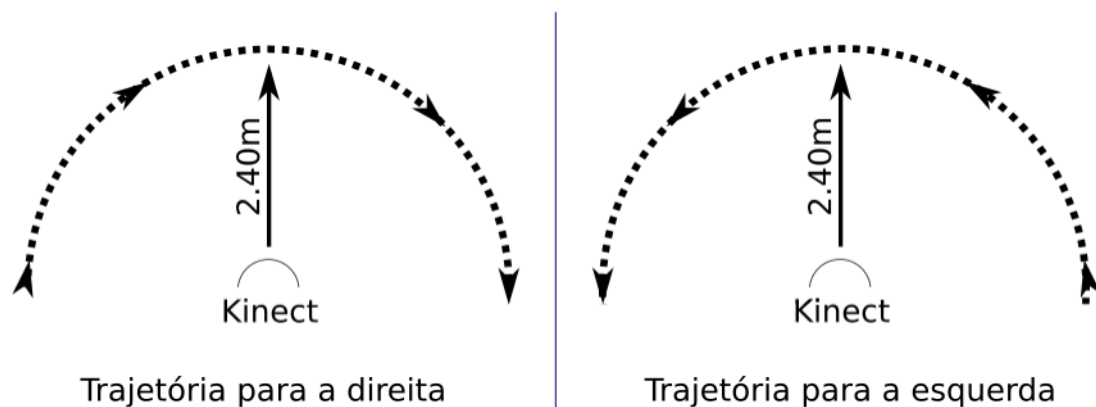


Figura 39 – Trajeto das capturas.

Nos testes realizados e apresentados no capítulo Resultados foi utilizado somente as capturas feitas das pessoas partindo do ponto a esquerda do sensor e indo para a direita.

Esta escolha de utilizar apenas uma trajetória foi feita pelo motivo dos algoritmos de aprendizado de máquina terem sido aplicados a tanto métodos de extração de atributos baseados em modelo como livres de modelo, e pelo fato dos métodos livres de modelo geralmente serem dependentes de pose, ou seja, o conjunto de treinamento e o conjunto de testes precisar ter as pessoas caminhando na mesma direção, optou-se por utilizar somente a trajetória para a

direita das gravações. Para uma análise do impacto dos resultados utilizando um conjunto de treinamento com pessoas caminhando em direção oposta em relação as pessoas do conjunto de teste, pode ser consultada a Seção 4.7. As capturas foram realizadas no período da noite com a iluminação vindo das lâmpadas no teto.

3.2.1.3 *Indivíduos capturados*

Das pessoas capturadas havia 27 homens e 23 mulheres, dentre as quais a maioria eram estudantes da Universidade Federal de Pelotas. Cada indivíduo capturado informou a altura, o peso e a idade. A altura das pessoas variou de 1,5m do indivíduo mais baixo para 1,93m da pessoa mais alta, sendo que a maior parte (28%) tinha altura entre 1,64 e 1,70. O peso das pessoas variaram de 47 kg a 120 kg, sendo que a maior parte dos indivíduos (26%) pesavam entre 71 a 82 kg. A idade dos indivíduos variou entre 18 e 49 anos, sendo que a maior parte (46%) tinha de 18 a 22 anos de idade.

Com base nos pesos fornecidos, foi possível calcular o Índice de Massa Corpórea (IMC) para cada pessoa. A maior parte (48%) estava com o peso normal e 14% estavam obesos, de acordo com a tabela fornecida por TECCIENCIA (2016).

Os gráficos das Figuras 40, 41, 42 e 43 apresentam as características dos indivíduos capturados:

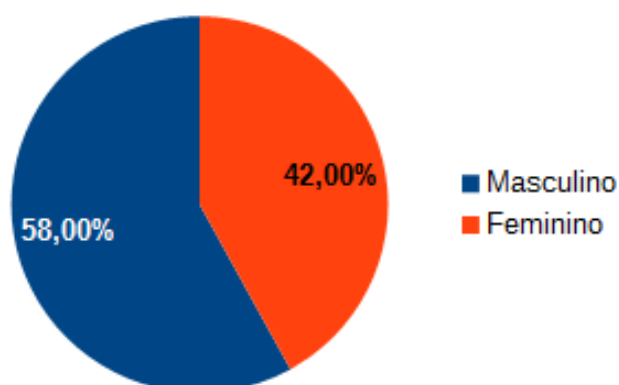


Figura 40 – Gênero das pessoas capturadas.

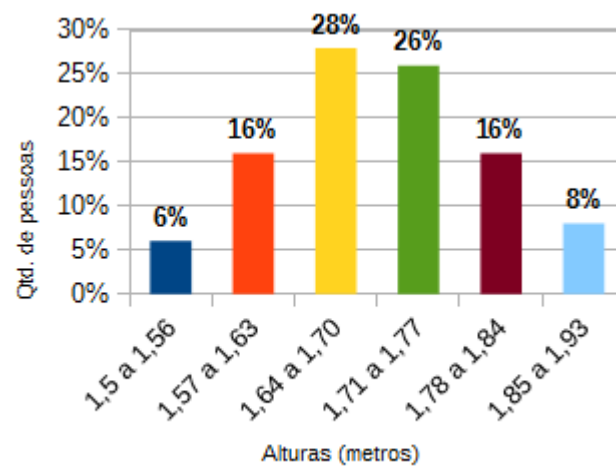


Figura 41 – Altura das pessoas capturadas.

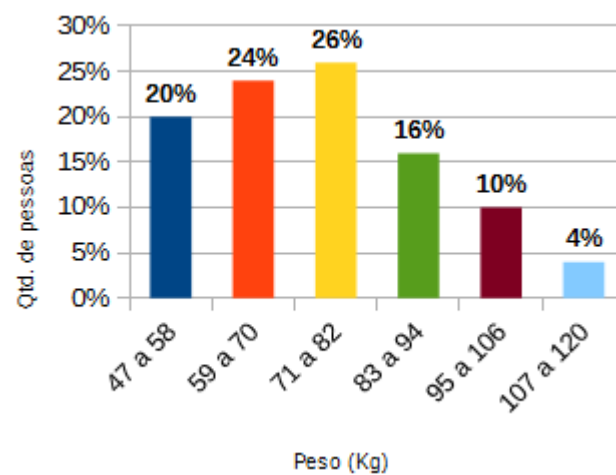


Figura 42 – Peso das pessoas capturadas.

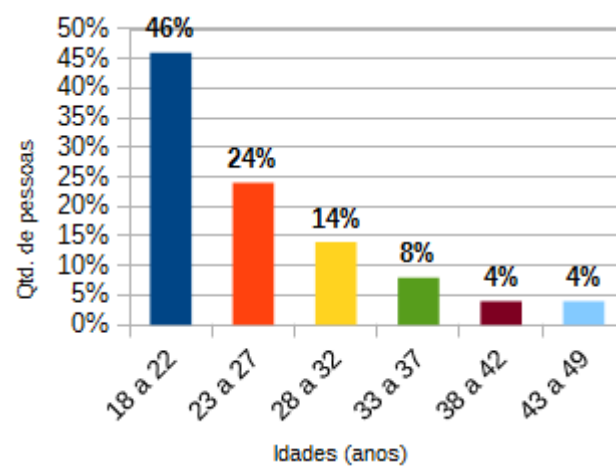


Figura 43 – Idade das pessoas capturadas.

3.2.2 Extração dos atributos

Para a extração e definição dos atributos que serão utilizados para a classificação dos indivíduos foram utilizadas as metodologias baseadas em modelo de ANDERSSON (2014), YOO; NIXON (2011) (Figuras de Movimento), assim como também foram utilizadas a abordagem livre de modelo de GUILLEN et al. (2009b) (Caixa Delimitadora) e uma proposta baseada na ideia de divisão em segmentos presente no trabalho de ARORA et al. (2015b). O objetivo da utilização destas metodologias foi utilizar abordagens tanto livre como baseada em modelo a fim de comparar as versões do Kinect.

Nas subseções seguintes, será apresentado o filtro utilizado na geração dos atributos e serão abordadas as implementações das metodologias.

3.2.2.1 Filtro

Conforme (MICROSOFT, 2016a) nos diz, na prática existe ruído presente nas posições das juntas e é importante usar um filtro de redução de ruído para remover o máximo possível de ruído dos dados. Este tipo de filtro é chamado de filtro de suavização, pois ele suaviza as posições no decorrer do tempo.

O filtro escolhido para ser aplicado foi um filtro linear sugerido pela (MICROSOFT, 2016b) chamado de Auto Regressive Moving Average (ARMA). Os filtros ARMA são uma classe de filtros que possuem várias categorias, a categoria de filtro ARMA aplicada neste trabalho chama-se Filtro de Movimento da Média Central.

A saída deste filtro é uma média ponderada da entrada atual, das N entradas anteriores e das M entradas futuras, conforme a equação 5:

$$\hat{X}_n = \sum_{i=-M}^N a_i X_{n-1} \quad (5)$$

O coeficiente a_i fornece o peso da equação e é configurado de acordo com equação 6:

$$a_1 = \frac{1}{M+N} \quad (6)$$

No presente trabalho foi utilizado para M e N o valor 8 ($M = 8$ e $N = 8$), com o coeficiente $a_i = 0,0625$, conforme utilizado na dissertação de (ANDERSSON, 2014).

3.2.2.2 Implementação da metodologia de Figuras de Movimento

Para extrair as características a fim de identificar as pessoas, foram analisados três parâmetros espaço-temporais: tamanho da passada, tempo do ciclo, a velocidade do caminhar e três parâmetros cinemáticos: coordenadas centrais do ciclograma do quadril Joelho, coefici-

ente de correlação entre os ângulos formados pelo joelho da perna direita e esquerda e a média geral dos ângulos das articulações inferiores do corpo (ANDERSSON, 2014) (YOO; NIXON, 2011).

O tamanho da passada foi obtido pelas dimensões físicas da imagem, onde foi calculado o movimento das pessoas analisando as coordenadas das figuras durante um ciclo do caminhar.

O tempo do ciclo foi obtido pela equação 7:

$$\text{tempo do ciclo (s)} = \frac{\text{duração do ciclo (quadros)}}{\text{FPS (quadros/s)}} \quad (7)$$

A velocidade foi obtida através da equação 8:

$$\text{velocidade (m/s)} = \frac{\text{tamanho da passada (m)}}{\text{tempo do ciclo (s)}} \quad (8)$$

O coeficiente de correlação cruzada foi obtida através da equação 9:

$$R = \frac{ss_{xy}^2}{ss_{xx} \times ss_{yy}} \quad (9)$$

Onde ss_{xx} , ss_{yy} e ss_{xy} são dados pelas equações 10, 11, 12:

$$ss_{xx} = \sum_{i=1}^N x^2 - Nx'^2 \quad (10)$$

$$ss_{yy} = \sum_{i=1}^N y^2 - Ny'^2 \quad (11)$$

$$ss_{xy} = \sum_{i=1}^N xy - Nx'y' \quad (12)$$

onde N é o número de pontos (x_i, y_i) e x', y' é a sua média (ANDERSSON, 2014).

A coordenada central do ciclograma quadril-jelho é definida de acordo a equação 13:

$$Centro_{QJ} = (Centro_{Quadril}, Centro_{Joelho}) \quad (13)$$

Onde $Centro_{Quadril}$ e $Centro_{Joelho}$ são os pontos centrais dos ângulos de 0 a N da articulação do quadril e também da articulação do joelho, dadas, respectivamente, pelas equa-

ções 14, 15:

$$Centro_{Quadril} = \frac{(\min(Q_0, \dots, Q_N) + \max([Q_0, \dots, Q_N]))}{2} \quad (14)$$

$$Centro_{Joelho} = \frac{(\min(J_0, \dots, J_N) + \max([J_0, \dots, J_N]))}{2} \quad (15)$$

3.2.2.3 Implementação da metodologia de Andersson

Seguindo a proposta de (ANDERSSON, 2014), utilizamos parâmetros antropométricos, cinemáticos e espaço-temporais para a identificação dos indivíduos.

Na análise antropométrica, para cada quadro capturado pelo sensor Kinect, foram calculadas as medidas de todos os segmentos do corpo dos indivíduos. Estas medidas foram calculadas através da distância Euclideana, conforme a equação 16:

$$segmento_{A,B} = \sqrt{\sum_{i=1}^3 (A_i - B_i)^2} \quad (16)$$

Para identificar a pessoa com base na antropometria, a média e o desvio padrão da distância dos segmentos em todos os quadros foram calculados.

Na análise cinemática, foram extraídos os ângulos das articulações inferiores do corpo humano, de cada quadro capturado, conforme o modelo de pêndulo proposto por YAM C. NIXON; CARTER (2003). Além dos ângulos dos quadris, joelhos e tornozelos, também foi utilizado o ângulo da abertura do pé descrito por MURRAY; DROUGHT; KORY (1964).

A Figura 44 ilustra os atributos cinemáticos utilizados:

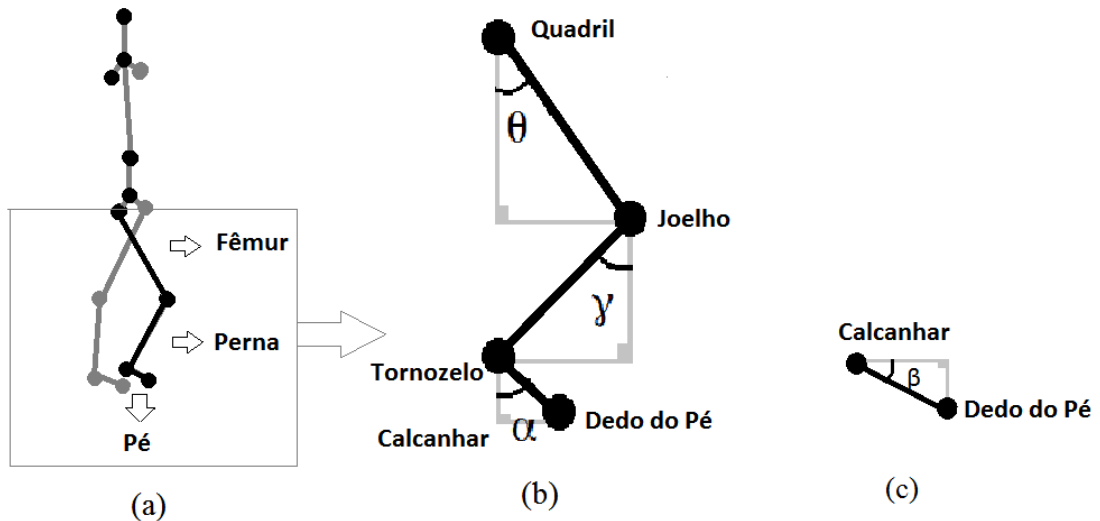


Figura 44 – Em (a) é apresentado os segmentos que foram utilizados, em (b) é descrito os ângulos destes segmentos e em (c) é apresentado o ângulo de abertura do pé. (ANDERSSON, 2014).

Os ângulos são calculados em cada quadro capturado e cada ângulo é formado por

3 pontos: duas articulações que correspondem ao segmento $A1$ e $A2$ e um terceiro ponto imaginário $p3$, o qual é formado pelas projeções x , y e z das duas articulações, formando um triângulo retângulo, conforme apresentado na Figura 44 (b). A equação 17 é utilizada para o cálculo dos ângulos:

$$angulo = \text{sen}^{-1} \left(\frac{A_2 - p_3}{A_2 - A_1} \right) \quad (17)$$

Conforme é possível visualizar na Figura 44 (b), o ângulo θ descreve a rotação do quadril durante o ciclo do caminhar, assim como o ângulo γ refere-se a rotação do joelho e o ângulo α refere-se a rotação do tornozelo. O ângulo β descreve a rotação de abertura do pé com eixo no calcanhar. Estes ângulos descrevem curvas periódicas durante o caminhar da pessoa que fornecem características úteis para o reconhecimento da pessoa.

Na medida em que a pessoa vai caminhando, as articulações descritas na Figura 44 se movem e as curvas periódicas formadas pelos ângulos destas articulações passam por fases de extensão e flexão, que podem ser notadas visualmente por picos (flexão) e vales (extensão), conforme apresentado na Figura 45.

Estes picos e vales foram utilizados como características para caracterizar o caminhar. Cada pico de flexão e cada vale de extensão possuem uma fase alta e também uma fase baixa. Foi gerado para cada fase alta e baixa uma média aritmética e um desvio padrão em relação a todo o caminhar capturado.

Para se encontrar os picos e os vales, foram calculados as diferenças de primeira ordem entre os ângulos, conforme equação 18:

$$diff(n) = \text{angulos}(n + 1) - \text{angulos}(n) \quad (18)$$

onde $diff$ é o resultado de todas as diferenças encontradas. Se n é maior que 0 e a próxima diferença representada por $diff(n + 1)$ é menor que 0 e a atual $diff(n)$ é diferente a próxima $diff(n + 1)$, essa diferença atual pode ser um pico. Somente será um pico se na posição atual a diferença entre seus valores originais for maior que 2.

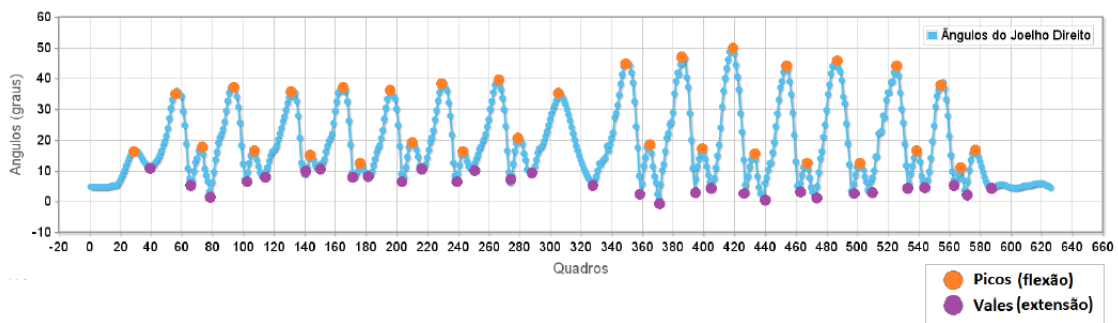


Figura 45 – Fases de extensão (vales) e flexão (picos) descritos pelos ângulos do joelho direito durante uma caminhada capturada (ANDERSSON, 2014).

Para encontrar os vales a mesma ideia é utilizada: se $diff(n)$ é menor que 0 e a próxima

diferença $diff(n + 1)$ é maior que 0 e a atual $diff$ é diferente da próxima $diff(n + 1)$, essa diferença pode ser um vale. Será um vale se a diferença entre os seus valores originais for maior que 2.

Na análise espaço-temporal, foram extraídos as seguintes características: tamanho do passo, tamanho da passada, duração do ciclo e velocidade.

O tamanho do passo foi obtido através da equação 19:

$$tamanhoPasso = \frac{\sum_{i=0}^n picos(i)}{n} \quad (19)$$

onde n são todos os picos dos ângulos do calcanhar.

O tamanho da passada (ou tamanho do ciclo do caminhar) $tamanhoCiclo$ é calculado como sendo aproximadamente $tamanhoPasso$ multiplicado por 2, para cada tamanho de passo calculado em cada quadro capturado. O tamanho do ciclo médio é portanto calculado através da média aritmética de cada valor, conforme a equação 20:

$$tamanhoCicloMedio = \frac{\sum_{i=0}^n tamanhoCiclo(i)}{n} \quad (20)$$

A duração do ciclo e a velocidade são calculados através das equações 21, 22:

$$duracaoCiclo = \frac{numQuadrosCicloMedio(quadros)}{30(quadros/s)} \quad (21)$$

$$velocidade = \frac{tamanhoCicloMedio}{duracaoCiclo} \quad (22)$$

onde $numQuadrosCicloMedio$ refere-se ao número de quadros, em média, que um ciclo de caminhar possui e $30 quadros/s$ a taxa de quadros por segundo da amostragem.

3.2.2.4 Implementação da metodologia de Divisão em Segmentos

ARORA et al. (2015a) em sua metodologia utilizou imagens divididas em segmentos para guardar um histórico do movimento do caminhar das pessoas. Neste trabalho foi implementado uma metodologia seguindo a mesma ideia, porém ao invés de utilizar as imagens das pessoas, os dados referentes as posições x, y e z dos indivíduos foram utilizadas.

Cada vez que uma articulação do corpo da pessoa entra na área de um segmento, o contador referente àquele segmento é incrementado e desta forma é possível armazenar um histórico de quais segmentos foram mais utilizados durante o caminhar e assim caracterizar o caminhar do indivíduo.

A área de captura dos indivíduos pode ser mapeada em quantidades diferentes de segmentos, por isso, diversas quantidades de segmentos foram testadas, partindo de 6 segmentos por dimensão (x, y e z) e indo até a quantidade de 16 segmentos. O que apresentou o melhor resultado foi a quantidade de 14 segmentos por dimensão e portanto no capítulo 4, referente aos resultados, esta é a configuração utilizada.

A Figura 46 ilustra a divisão da imagem em segmentos:

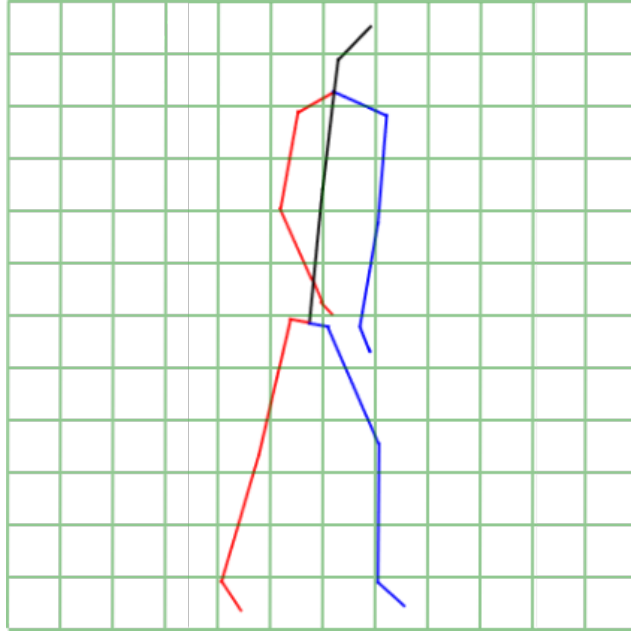


Figura 46 – Adaptação da metodologia de (ARORA et al., 2015a).

3.2.2.5 Implementação da metodologia da Caixa Delimitadora

A largura do retângulo é a representação da largura da pessoa. Ela é obtida através da média da diferença entre os pontos do lado esquerdo e direito no eixo x em cada momento do instante t, conforme equação 23:

$$W = \frac{\sum_{i=1}^{N_W} [x_{RP}(t) - x_{LP}(t)]_i}{N_W} \quad (23)$$

Onde N_W representa o total da quantidade de dados para a largura. A largura da medida da cadência das mãos e pernas com sua variação.

A área do retângulo é determinada pelo ciclo do caminhar, conforme ilustra a figura 47: A área A é obtida pela média da multiplicação entre a altura do retângulo e a largura a cada

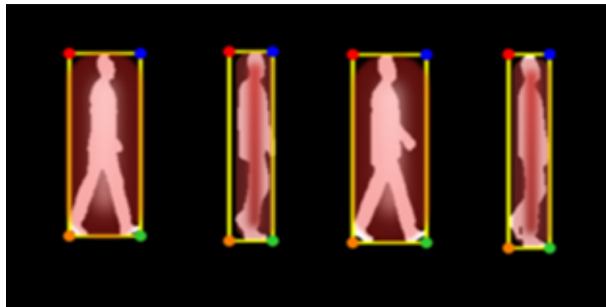


Figura 47 – Área do retângulo (GUILLÉN et al., 2009a).

momento t , dado pela equação 24:

$$A = \frac{\sum_{i=1}^{N_A} [H(t) \cdot W(t)]_i}{N_A} \quad (24)$$

Onde N_A representa o total de dados acumulados para a área.

O ângulo diagonal do retângulo varia durante o ciclo do caminhar, conforme ilustra a figura 48:

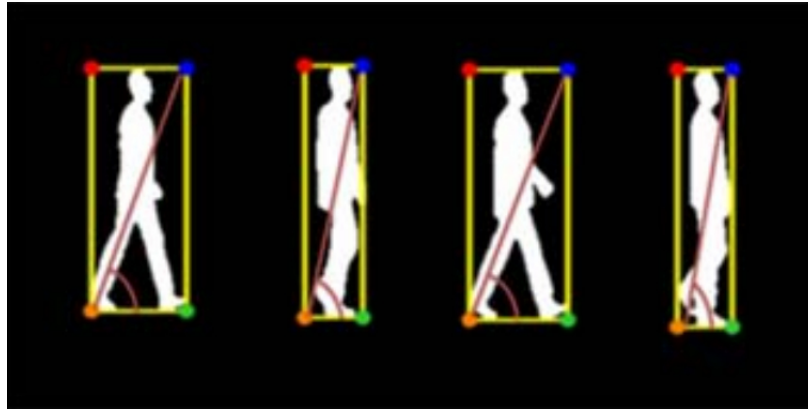


Figura 48 – Ângulo diagonal do retângulo (GUILLEN et al., 2009a).

Este ângulo médio diagonal α é obtido pela média da diferença, em graus, entre o eixo x e a diagonal (ilustrada na figura anterior) em cada instante t , conforme equação 25:

$$\alpha = \frac{\sum_{i=1}^{N_\alpha} \left[\tan^{-1} \left(\frac{H(t)}{W(t)} \right) \right]_i}{N_\alpha} \quad (25)$$

Onde N_α o total da quantidade de dados coletados para a característica do ângulo diagonal. Também é extraído como característica a potência espectral total P , dada pelo somatório de cada aplicação da Transformada Rápida de Fourier (FFT) nas áreas calculadas em cada quadro $A(f)$, conforme equação 26: (GUILLEN et al., 2009a) (ANDERSSON, 2014)

$$P = \sum_{f=1}^{N_f} A(f) \quad (26)$$

3.2.3 Classificação

Conforme descrito na Seção 2.5, para classificar os indivíduos foram utilizados os seguintes algoritmos de aprendizado de máquina: KNN, MLP, Naive Bayes e Random Forest. A única exceção ocorre para a metodologia de extração de atributos de Divisão em Segmentos, pois neste método não foi utilizado o MLP devido a grande quantidade de atributos que esta metodologia gera, fazendo com que o tempo necessário para treinar esta rede neural seja muito grande no computador utilizado nesta dissertação (Intel Core i7 com 8 GB de RAM).

Os algoritmos foram testados variando as seguintes configurações: para o K-Nearest Neighbor foram testados valores de $K = 1$ a $K = 12$, para o algoritmo Multi-layer Perceptron foram testados os valores de 500 épocas a 6000 épocas, com intervalo de 500 épocas, ou seja, o número de épocas testadas no algoritmo foi 500, 1000, 1500, 2000 e assim por diante até chegar a 6000. O Naive Bayes foi testado com e sem *Kernel Estimator* e o algoritmo Random Forest foi testado com o número de épocas variando de 500 a 6000, com intervalo de 500 épocas.

A rede neural MLP foi aplicada utilizando 1 camada de entrada, 1 camada oculta e 1 camada de saída. O número de neurônios da camada oculta foi aplicado sendo metade da quantidade de neurônios da camada de entrada, para assim a rede neural ficar adaptada a quantidade de atributos de cada proposta. O Random Forest foi aplicado utilizando 100 árvores.

Os algoritmos foram implementados utilizando o WEKA (*Waikato Environment for Knowledge Analysis*) que foi criado pela Universidade de Waikato com o objetivo de facilitar a aplicação de algoritmos de aprendizado de máquina WEKA (2017).

A tabela 8 sumariza a implementação no WEKA dos algoritmos utilizados neste trabalho:

Tabela 8 – Implementação dos algoritmos de aprendizado de máquina no WEKA

Algoritmo	Implementação no WEKA
KNN	<code>weka.classifiers.lazy.IBk</code>
MLP	<code>weka.classifiers.functions.MultilayerPerceptron</code>
Naive Bayes	<code>weka.classifiers.bayes.NaiveBayes</code>
Random Forest	<code>weka.classifiers.trees.RandomForest</code>

3.2.4 Método de validação

Para aplicar os algoritmos de aprendizado de máquina utilizados neste trabalho, foi necessário separar os atributos extraídos das capturas realizadas em conjunto de treinamento e conjunto de teste.

Como existem 5 capturas para cada pessoa, separou-se 4 capturas para treinamento e 1 para teste. Em cima das 4 capturas para treinamento, utilizou-se validação cruzada n-Fold, a qual consiste em dividir o conjunto de dados de tamanho **n** em **m** partições com tamanho de **n/m** (BW, 2017).

O objetivo de realizar esta validação cruzada no conjunto de treinamento foi para obter os parâmetros dos algoritmos de aprendizado de máquina que seriam aplicados no conjunto de teste. Para cada proposta de extração de atributos foram testados através desta validação cruzada no conjunto de treinamento diversos parâmetros (conforme apresentados na Seção 3.2.3) e os melhores foram utilizados no conjunto de teste. Na seção de resultados, é apresentado a configuração que foi aplicada ao conjunto de teste em cada proposta e algoritmo utilizado.

3.2.5 Protocolo de captura

Visando garantir a consistência na captura dos indivíduos foi desenvolvido um protocolo a ser seguido. Nesta subseção é resumido as etapas do protocolo, uma descrição mais completa juntamente com a ilustração do formulário utilizado para anotar as informações dos indivíduos pode se encontrada no Apêndice B.

Os passos para realizar a captura dos indivíduos são os seguintes:

1. Encontrar o local ideal;
2. Posicionar adequadamente os sensores Kinects;
3. Marcar no chão o caminho que as pessoas devem percorrer;
4. Montar o equipamento;
5. Convidar pessoas;
6. Realizar as capturas.

O passo de número 1 é realizado uma única vez, pois após definido o local onde as capturas são realizadas é importante manter o mesmo local para evitar variações nas gravações por causa de ambientes diferentes.

Os demais passos são feitos todos os dias em que as capturas são realizadas, sendo que antes das coletas é realizado os passos 2, 3 e 4 e depois os passos 5 e 6 se repetem até ao máximo de pessoas que for possível capturar no mesmo dia.

4 RESULTADOS

Neste capítulo são apresentados e analisados os resultados obtidos utilizando quatro metodologias diferentes para extração de atributos. As configurações testadas nos algoritmos de aprendizado de máquina estão descritas na Seção 3.2.4.

O algoritmo de aprendizado de máquina e a configuração que obteve a melhor acurácia no conjunto de validação variou de sensor para sensor, por exemplo, na aplicação da metodologia da Caixa delimitadora, no Kinect 360 o algoritmo Naive Bayes com Kernel Estimator obteve a melhor acurácia e no Kinect One o Random Forest com 5000 iterações obteve a maior acurácia, por este motivo nas tabelas deste capítulo é apresentada para cada sensor a configuração e o resultado em que cada algoritmo obteve o melhor resultado.

4.1 Proposta de Andersson

Os seguintes resultados foram alcançados utilizando a metodologia proposta por ANDERSSON (2014). Esta metodologia está descrita na Seção 3.2.2.3.

Tabela 9 – Resultado utilizando a metodologia de Andersson com Kinect 360

Kinect 360		
Algoritmo	Configuração	Resultado
Random Forest	1000 iterações	94%
KNN	K = 1	60%
Naive Bayes	Sem Kernel Estimator	50%
MLP	500 iterações	84%

Tabela 10 – Resultado utilizando a metodologia de Andersson com Kinect One

Kinect One		
Algoritmo	Configuração	Resultado
Random Forest	6000 iterações	98%
KNN	K = 1	44 %
Naive Bayes	Sem Kernel Estimator	60%
MLP	3000 iterações	86%

O algoritmo que obteve a maior acurácia foi o Random Forest: com os dados do Kinect 360 foi possível obter acurácia de 94% e com o Kinect One obteve-se acurácia de 98%, resultando em uma diferença entre os sensores de 4,08%.

É interessante atentarmos ao fato de que o Random Forest obteve o melhor resultado em ambos os sensores, e que este foi, conforme destacado na Seção 2.5, o algoritmo utilizado pela Microsoft no *software* do Kinect para obter o esqueleto dos indivíduos.

As tabelas 11 e 12 apresentam os resultados quando apenas os atributos antropométricos são aplicados na classificação:

Tabela 11 – Atributos antropométricos com Kinect 360

Kinect 360		
Algoritmo	Configuração	Resultado
Random Forest	1000 iterações	92%
KNN	K = 1	82%
Naive Bayes	Sem Kernel Estimator	60%
MLP	500 iterações	88%

Tabela 12 – Atributos antropométricos com Kinect One

Kinect One		
Algoritmo	Configuração	Resultado
Random Forest	6000 iterações	98%
KNN	K = 1	88%
Naive Bayes	Sem Kernel Estimator	78%
MLP	3000 iterações	90%

Em ambos os sensores o Random Forest apresentou a maior taxa de acertos, com 92% de acurácia para o Kinect 360 e 98% de acurácia para o Kinect One, resultando em uma diferença entre os sensores de 6,12%. Utilizando o Kinect 360 a acurácia diminuiu apenas 2 pontos percentuais se comparado quando todos os atributos da proposta de ANDERSSON (2014) foram aplicados, e no Kinect One a acurácia se manteve igual.

A seguir, nas tabelas 13 e 14, são apresentados os resultados de quando cada atributo antropométrico é retirado e o impacto que isto gera no resultado final, descrevendo assim a importância que cada atributo possui no conjunto. Quando o resultado do atributo está com valor negativo significa que a acurácia melhorou sem a presença deste atributo.

O algoritmo utilizado para gerar os resultados nos dados do Kinect 360 foi o Random Forest com 1000 iterações e nos dados do Kinect One foi o Random Forest com 6000 iterações, pois estes algoritmos foram os utilizados na aplicação do conjunto com todos os atributos.

Tabela 13 – Impacto de cada atributo antropométrico no Kinect 360

Kinect 360		
Resultado com todos os atributos antropométricos: 92%		
Atributo	Resultado	Mudança
Cabeça a centro dos ombros	92%	0%
Centro dos ombros a espinha	90%	2,17%
Espinha a centro do quadril	92%	0%
Centro dos ombros e ombro direito	90%	2,17%
Ombro direito a cotovêlo direito	92%	0%
Cotovelo direito a punho direito	94%	-2,17%
Punho direito a mão direita	90%	2,17%
Centro dos quadris a quadril direito	94%	-2,17%
Quadril direito a joelho direito	94%	-2,17%
Joelho direito a calcanhar direito	92%	0%
Calcanhar direito a ponta do pé direito	94%	-2,17%
Centro dos ombros a ombro esquerdo	90%	2,17%
ombro esquerdo a cotovelo esquerdo	92%	0%
Cotovelo esquerdo a punho esquerdo	92%	0%
Punho esquerdo a mão esquerda	94%	-2,17%
Centro dos quadris e quadril esquerdo	92%	0%
Quadril esquerdo a joelho esquerdo	94%	-2,17%
Joelho esquerdo a calcanhar esquerdo	88%	4,34%
Calcanhar esquerdo a ponta do pé esquerdo	90%	2,17%

O atributo que apresentou o maior impacto no Kinect 360 foi a distância entre o joelho e o calcanhar esquerdos (4,34%).

É interessante notarmos que vários atributos obtiveram resultados negativos, ou seja, quando estes atributos foram retirados a acurácia melhorou. Este resultado negativo não significa necessariamente que estes atributos sejam ruins, mas provavelmente o Kinect 360 não conseguiu mapear corretamente estes pontos do esqueleto das pessoas.

Quase todos os atributos com resultado negativo não estão do lado oculto da pessoa em relação ao sensor, mas sim do lado que estava virado diretamente para os Kinects, mostrando a capacidade do sensor de conseguir inferir a posição das juntas do esqueleto mesmo quando elas estão ocultas.

Tabela 14 – Impacto de cada atributo antropométrico no Kinect One

Kinect One		
Resultado com todos os atributos antropométricos: 98%		
Atributo	Resultado	Mudança
Cabeça a pescoço	98%	0%
Pescoço a Centro dos ombros	96%	2,04%
Centro dos ombros a ombro direito	96%	2,04%
Ombro direito a cotovêlo direito	98%	0%
Cotovelo direito a punho direito	98%	0%
Punho direito a mão direita	98%	0%
Mão direita e dedo indicador	98%	0%
Mão direita a polegar	98%	0%
Centro dos ombros a ombro esquerdo	96%	2,04%
Ombro esquerdo a cotovêlo esquerdo	98%	0%
Cotovelo esquerdo a punho esquerdo	98%	0%
Punho esquerdo a mão esquerda	98%	0%
Mão esquerda e dedo indicador	98%	0%
Mão esquerda a polegar	96%	2,04%
Centro dos ombros a centro da coluna	96%	2,04%
Centro da coluna e centro dos quadris	96%	2,04%
Centro dos quadris a quadril direito	98%	0%
Centro dos quadris a quadril esquerdo	98%	0%
Quadril direito a joelho direito	96%	2,04%
Quadril esquerdo a joelho esquerdo	98%	0%
Joelho direito a calcanhar direito	96%	2,04%
Joelho esquerdo a calcanhar esquerdo	96%	2,04%
Calcanhar esquerdo a ponta do pé esquerdo	96%	2,04%
Calcanhar direito a ponta do pé direito	96%	2,04%

Para o Kinect One nenhum atributo apresentou resultado negativo, o que sugere uma melhor precisão da nova versão do sensor em obter os pontos do esqueleto dos indivíduos.

A Figura 49 ilustra o esqueleto de um indivíduo capturado em cada um dos sensores, onde em vermelho está destacado os atributos que obtiveram resultado negativo e em verde os que obtiveram resultado positivo, as partes na cor cinza representam os atributos em que o impacto foi 0%.

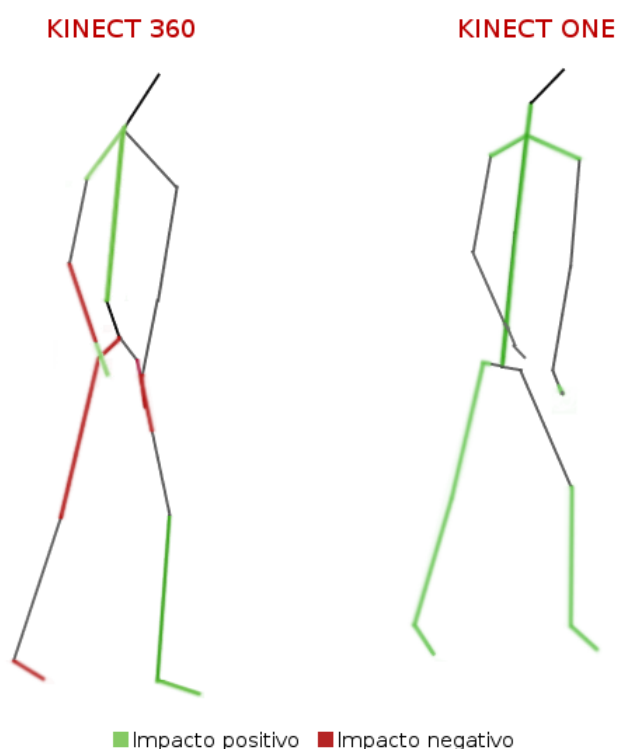


Figura 49 – Análise dos atributos antropométricos da metodologia de ANDERSSON (2014).

É possível perceber que a maior parte dos atributos com resultado negativo do Kinect 360 advém da parte inferior do esqueleto, sugerindo uma maior dificuldade da primeira versão do sensor em capturar os membros desta parte do corpo das pessoas. Também podemos verificar que a região do quadril do Kinect 360 trouxe impacto negativo, o que pode sugerir que a nova configuração do quadril apresentada no Kinect One seja melhor para caracterizar o indivíduo.

Um aspecto interessante a destacar é que o polegar esquerdo e o pescoço trouxeram resultados positivos na análise do Kinect One, mostrando que estes novos pontos que existem apenas nesta nova versão do sensor agregam valor aos resultados.

Utilizando somente os atributos cinemáticos, os resultados são apresentados nas tabelas 15 e 16:

Tabela 15 – Atributos cinemáticos com Kinect 360

Kinect 360		
Algoritmo	Configuração	Resultado
Random Forest	1000 iterações	54%
KNN	K = 1	26%
Naive Bayes	Sem Kernel Estimator	22%
MLP	500 iterações	36%

O Random Forest apresentou a maior acurácia em ambos os sensores, com acurácia de 54% para o Kinect 360 e de 68% para o Kinect One, resultando em uma diferença entre os

Tabela 16 – Atributos cinemáticos com Kinect One

Kinect One		
Algoritmo	Configuração	Resultado
Random Forest	6000 iterações	68%
KNN	K = 1	14%
Naive Bayes	Sem Kernel Estimator	18%
MLP	3000 iterações	42%

sensores de 20,58%.

São apresentados nas tabelas 17 e 18 os resultados de quando cada atributo cinemático é removido e a diferença que isto faz no número de acertos. O algoritmo utilizado para gerar os resultados foi o Random Forest com 1000 iterações nos dados do Kinect 360 e o Random Forest com 6000 iterações nos dados do Kinect One, pois foram os utilizados na aplicação dos conjuntos de atributos.

Tabela 17 – Impacto de cada atributo cinemático no Kinect 360

Kinect 360		
Resultado com todos os atributos cinemáticos: 54%		
Ângulo	Resultado	Mudança
Calcanhar e pé esquerdos	46%	14,81%
Calcanhar e pé direitos	48%	11,11%
Abertura do pé esquerdo	48%	11,11%
Abertura do pé direito	50%	7,4%
Quadril e joelho esquerdos	50%	7,4%
Quadril e joelho direitos	46%	14,81%
Joelho e calcanhar esquerdos	54%	0%
Joelho e calcanhar direitos	54%	0%

O ângulo entre o calcanhar e o pé esquerdos e o ângulo entre o quadril e o joelho esquerdos são os atributos que possuem maior impacto (14,81%), sendo que os ângulos formados entre o joelho e o calcanhar em ambas as pernas não agregaram valor, apresentando 0% de impacto.

Tabela 18 – Impacto de cada atributo cinemático no Kinect One

Kinect One		
Resultado com todos os atributos cinemáticos: 68%		
Ângulo	Resultado	Mudança
Calcanhar e pé esquerdos	56%	17,64%
Calcanhar e pé direitos	60%	11,76%
Abertura do pé esquerdo	60%	11,76%
Abertura do pé direito	68%	0%
Quadril e joelho esquerdos	66%	2,94%
Quadril e joelho direitos	60%	11,76%
Joelho e calcanhar esquerdos	62%	8,82%
Joelho e calcanhar direitos	66%	2,94%

No Kinect One o ângulo entre o calcanhar e pé esquerdos é o atributo de maior impacto (17,64%), sendo que a retirada do atributo referente ao ângulo de abertura do pé direito não surtiu diferença alguma (0% de impacto).

Avançando para a análise dos atributos espaço-temporais, nas tabelas 19 e 20 são apresentados os resultados obtidos com a aplicação destes atributos:

Tabela 19 – Atributos espaço-temporais com Kinect One

Kinect 360		
Algoritmo	Configuração	Resultado
Random Forest	1000 iterações	16%
KNN	K = 1	20%
Naive Bayes	Sem Kernel Estimator	24%
MLP	500 iterações	20%

Tabela 20 – Atributos espaço-temporais com Kinect 360

Kinect One		
Algoritmo	Configuração	Resultado
Random Forest	6000 iterações	20%
KNN	K = 1	20%
Naive Bayes	Sem Kernel Estimator	20%
MLP	3000 iterações	16%

O algoritmo que obteve o melhor resultado nos dados do Kinect 360 foi o Naive Bayes (24%) e nos dados do Kinect One o Random Forest, KNN e Naive Bayes empataram com o melhor resultado (20%), resultando em uma diferença entre os sensores de 16,66%, com o Kinect 360 obtendo a maior acurácia.

É apresentado nas tabelas 21 e 22 os resultados de quando cada atributo espaço-temporal é removido e a diferença que isto faz no número de acertos. O algoritmo utilizado nos resultados do Kinect 360 é o Naive Bayes sem Kernel Estimator e nos resultados do Kinect One é utilizado o KNN com $K = 1$.

Tabela 21 – Impacto de cada atributo espaço-temporal com Kinect 360

Kinect 360		
Resultado com todos os atributos espaço-temporais: 24%		
Atributo	Resultado	Mudança
Tamanho da Passada	24%	0%
Tamanho do Ciclo	22%	8,33%
Duração do Ciclo	20%	16,66%
Velocidade	20%	16,66%

Tabela 22 – Impacto de cada atributo espaço-temporal com Kinect One

Kinect One		
Resultado com todos os atributos espaço-temporais: 20%		
Atributo	Resultado	Mudança
Tamanho da Passada	18%	10%
Tamanho do Ciclo	14%	30%
Duração do Ciclo	22%	-10%
Velocidade	22%	-10%

No Kinect 360 o atributo de maior impacto foi a duração do ciclo e a velocidade, no Kinect One o melhor atributo foi o tamanho do ciclo.

Ao analisar os conjuntos de atributos da metodologia de ANDERSSON (2014), percebe-se que o conjunto de atributos antropométricos é o maior responsável pela alta taxa de acertos, seguido dos cinemáticos que trazem resultados próximos a 60% e por fim os atributos espaço-temporais que apresentam resultados próximos a 20%.

A Figura 50 apresenta um gráfico comparando os resultados em cada um dos sensores:

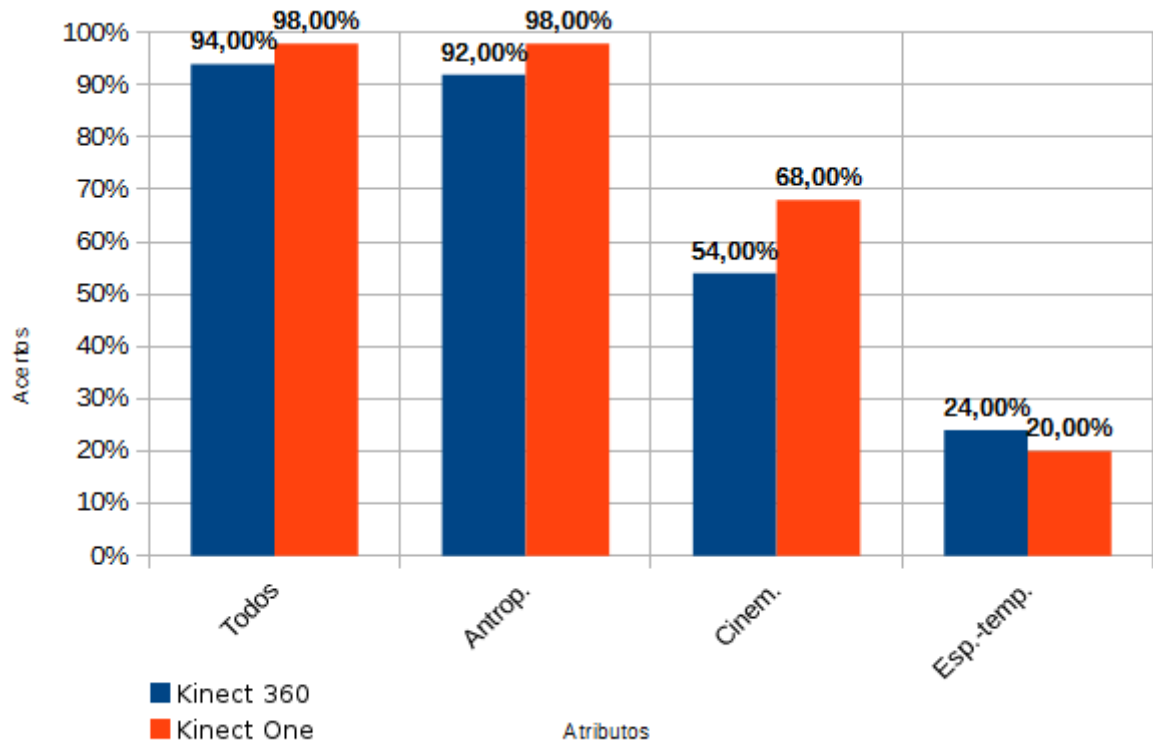


Figura 50 – Comparativo dos sensores utilizando a metodologia de ANDERSSON (2014).

A tabela 23 apresenta a diferença entre os sensores dos resultados apresentados na Figura 50:

Tabela 23 – Diferença entre os sensores na aplicação da proposta de ANDERSSON (2014)

Atributos	Diferença de acurácia
Todos	4,08%
Antropométicos	6,1%
Cinemáticos	20,58%
Espaço-temporais	-20%

Quando todos os atributos são aplicados, ou apenas os antropométricos, a diferença entre os sensores não é tão expressiva, porém quando somente os atributos cinemáticos ou espaço-temporais são aplicados, a diferença entre eles é grande, onde o Kinect One é 20,58% superior ao Kinect 360 nos atributos cinemáticos e inferior 20% nos atributos espaço-temporais.

Pelo motivo do Kinect 360 ter obtido melhor resultado apenas nos atributos espaço-temporais, conclui-se portanto que a menos que o objetivo seja utilizar apenas estes atributos, ao utilizar a metodologia de ANDERSSON (2014) é vantajoso utilizar o Kinect One.

4.2 Proposta das Figuras de Movimento

As tabelas 24 e 25 apresentam os resultados que foram alcançados utilizando a metodologia proposta por YOO; NIXON (2011).

Tabela 24 – Resultado utilizando a metodologia das Figuras de Movimento com Kinect 360

Kinect 360		
Algoritmo	Configuração	Resultado
Random Forest	1000 interações	36%
KNN	K = 1	28%
Naive Bayes	Com Kernel Estimator	38%
MLP	500 interações	36%

Tabela 25 – Resultado utilizando a metodologia das Figuras de Movimento com Kinect One

Kinect One		
Algoritmo	Configuração	Resultado
Random Forest	3500 iterações	38%
KNN	K = 1	30 %
Naive Bayes	Sem Kernel Estimator	32%
MLP	1000 iterações	34%

Analisando as tabelas 24 e 25 percebemos que os algoritmos de maior acurácia em ambos os sensores chegaram a 38%. A baixa acurácia desta metodologia, não estando acima de 38% em nenhum algoritmo, sugere que esta proposta não traz bons resultados ao aplicá-la a dados do Kinect.

As tabelas 26 e 27 apresentam o impacto de cada atributo. Foi utilizado o Naive Bayes com Kernel Estimator para os dados do Kinect 360 e o Random Forest de 3500 iterações para os dados do Kinect One, pois foram os algoritmos que apresentaram os melhores resultados para cada sensor no conjunto de validação quando foram aplicados todos os atributos da proposta das Figuras de Movimento.

Tabela 26 – Analisando o impacto de cada atributo das Figuras de Movimento no Kinect 360

Kinect 360		
Resultado com todos os atributos: 38%		
Atributo	Resultado	Mudança
Tempo do ciclo	42%	-10,52%
Tamanho da passada	28%	26,31%
Velocidade	32%	15,78%
Média dos ângulos das juntas	28%	26,31%
Coeficiente de correlação	36%	5,26%
Coordenadas centrais	30%	21,05%

Tabela 27 – Analisando o impacto de cada atributo das Figuras de Movimento no Kinect One

Kinect One		
Resultado com todos os atributos: 38%		
Atributo	Resultado	Mudança
Tempo do ciclo	38%	0%
Tamanho da passada	36%	5,55%
Velocidade	44%	-13,63%
Média dos ângulos das juntas	28%	35,71%
Coeficiente de correlação	42%	-9,52%
Coordenadas centrais	34%	11,76%

A Figura 51 apresenta o gráfico com o comparativo dos resultados:

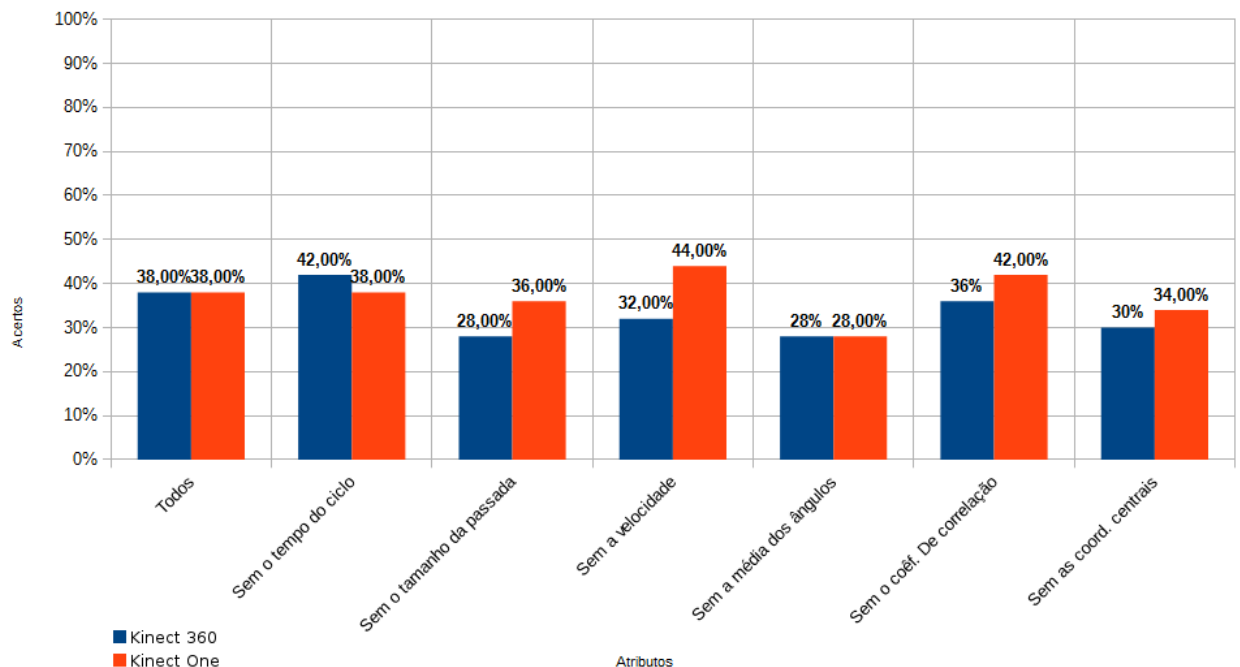


Figura 51 – Resultados das Figuras de Movimento.

Através do resultado individual de cada atributo, pode-se notar que o Kinect 360 teve dificuldades no atributo referente ao tempo do ciclo, pelo fato do tempo do ciclo utilizar a diferença entre a posição dos calcanhares durante o seu cálculo, é provável que a posição dos calcanhares durante a captura tenha sido obtida com melhor acurácia no Kinect One. Uma observação que colabora com esta sugestão é de que o ângulo dos indivíduos que agregaram maior valor aos resultados nos atributos cinemáticos da proposta de ANDERSSON (2014) foi o do calcanhar e o pé esquerdos e de que por meio da análise dos atributos antropométricos da metodologia de ANDERSSON (2014) verificou-se que a distância entre o calcanhar e a ponta do pé no Kinect 360 trouxe impacto negativo aos resultados.

O impacto negativo do atributo referente ao coeficiente de correlação no Kinect One

sugere que o fêmur de ambas as pernas no Kinect One produziram ângulos muito semelhantes entre eles, esta possível semelhança entre os ângulos pode ter relação com a nova configuração do quadril utilizada na segunda versão do sensor, onde a posição dos pontos no quadril estão alinhados com a base da coluna.

Embora a análise individual dos atributos da metodologia de Figuras de Movimento possa nos trazer dados interessantes, a acurácia desta metodologia não foi satisfatória para a caracterização dos indivíduos. Esta acurácia inferior a 45% se deve, entre outros fatores, ao motivo de que três dos parâmetros deste método são espaço-temporais (tempo do ciclo, tempo da passada e velocidade) e já na metodologia de ANDERSSON (2014) verificou-se que os parâmetros espaço-temporais agregam pouca melhora no número de acertos. Provavelmente todas as pessoas caminharam em velocidades similares, não permitindo que estes parâmetros trouxessem uma boa caracterização dos indivíduos.

4.3 Proposta da Caixa Delimitadora

Nas tabelas 28 e 29 é possível visualizar os resultados que foram alcançados utilizando a metodologia proposta por GUILLEN et al. (2009a).

Tabela 28 – Resultado utilizando a metodologia da Caixa Delimitadora com Kinect 360

Kinect 360		
Algoritmo	Configuração	Resultado
Random Forest	5500 interações	14%
KNN	K = 1	10%
Naive Bayes	Com Kernel Estimator	20%
MLP	1000 iterações	18%

Tabela 29 – Resultado utilizando a metodologia da Caixa Delimitadora com Kinect One

Kinect One		
Algoritmo	Configuração	Resultado
Random Forest	500 iterações	34%
KNN	K = 1	18%
Naive Bayes	Com Kernel Estimator	26%
MLP	3000 iterações	20%

Embora exista uma diferença significativa (58,82%) entre os resultados de maior acurácia obtidos em cada sensor, pois com os dados do Kinect 360 foi possível obter acurácia de 14% e com os dados do Kinect One a maior acurácia foi de 34%, os resultados nos mostram que esta metodologia não trouxe uma boa precisão na identificação dos indivíduos.

As tabelas 30 e 31 apresentam o impacto de cada atributo. Foi utilizado Random Forest de 5500 iterações para os dados do Kinect 360 e Random Forest de 500 iterações para os

dados do Kinect One por estas configurações terem sido as que obtiveram a maior acurácia quando todos os atributos da proposta foram aplicados no conjunto de validação.

Tabela 30 – Analisando o impacto de cada atributo da Caixa Delimitadora no Kinect 360

Kinect 360		
Resultado com todos os atributos: 20%		
Atributo	Resultado	Mudança
Altura média	12%	40%
Largura média	20%	0%
Área média	18%	10%
Ângulo médio	14%	30%
FFA	12%	40%

Tabela 31 – Analisando o impacto de cada atributo da Caixa Delimitadora no Kinect One

Kinect One		
Resultado com todos os atributos: 34%		
Atributo	Resultado	Mudança
Altura média	14%	58,82%
Largura média	38%	-11,76%
Área média	38%	-11,76%
Ângulo médio	34%	0%
FFA	32%	5,88%

A Figura 52 apresenta o gráfico com o comparativo dos resultados:

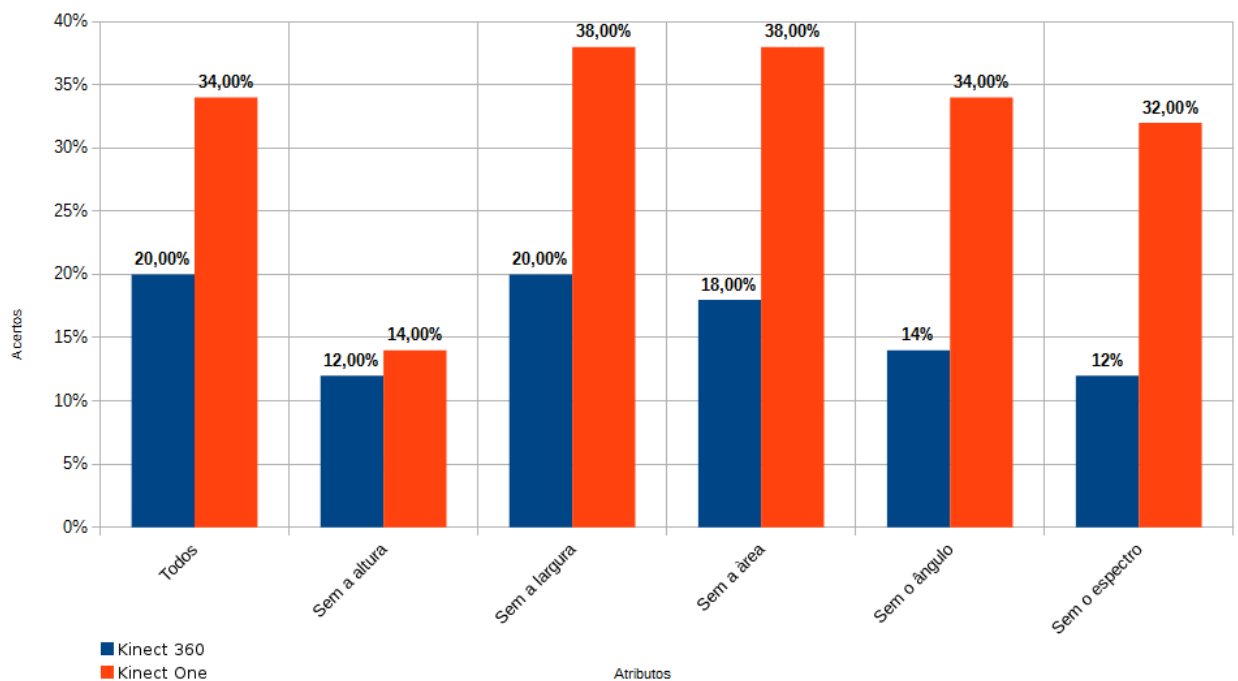


Figura 52 – Resultados da Caixa Delimitadora.

Pelos resultados, verifica-se que utilizar os Kinects aplicando a área como meio de caracterização dos indivíduos pelo caminhar não apresentou boa acurácia, sendo que os resultados foram ainda piores no Kinect One, pois embora o resultado dele tenha sido superior, pela análise dos atributos verifica-se que a altura do indivíduo foi o atributo que praticamente caracterizou os indivíduos, pois os demais, com exceção do FFA, ou impactaram negativamente os resultados ou não agregaram acurácia.

4.3.1 Aplicação do sistema de autenticação de Guillen

Nesta subseção são apresentados os resultados quando os atributos extraídos foram aplicados ao sistema de autenticação desenvolvido por GUILLEN et al. (2009a) que está descrito na Seção 2.3.2.4.

Os resultados são apresentados na tabela 32:

Tabela 32 – Resultado utilizando Sistema de Autenticação de Guillen

Sensor	Resultado
Kinect 360	5.6%
Kinect One	8.8%

Embora a diferença entre os sensores tenha tido uma porcentagem expressiva (36,36%), em nenhum dos dois a acurácia foi satisfatória, pois os resultados mostraram acurácia abaixo de 10%.

Verifica-se que ao aplicarmos a metodologia da Caixa Delimitadora nos algoritmos de aprendizado de máquina KNN, Naive Bayes, Random Forest e MLP os resultados foram significativamente melhores, o que sugere que este sistema de autenticação não traz bons resultados.

Ao investigar a razão da acurácia deste sistema ter sido tão baixa, verificou-se que quando um exemplo da base de teste é aplicado no modelo treinado, o sistema quase sempre encontrava mais de uma pessoa que passava pela árvore de decisão criada, o que aponta para uma necessidade do sistema conseguir decidir adequadamente qual exemplo a ser escolhido quando mais de um é autenticado.

4.4 Proposta da Divisão em Segmentos

Nas tabelas 33 e 34 é possível visualizar os resultados desta metodologia.

Tabela 33 – Resultado utilizando a metodologia da Divisão em Segmentos com Kinect 360

Kinect 360		
Algoritmo	Configuração	Resultado
Random Forest	1000 iterações	94%
KNN	K = 1	86%
Naive Bayes	Com Kernel Estimator	80%

Tabela 34 – Resultado utilizando a metodologia da Divisão em Segmentos com Kinect One

Kinect One		
Algoritmo	Configuração	Resultado
Random Forest	2000 iterações	98%
KNN	K = 1	92%
Naive Bayes	Sem Kernel Estimator	94%

Em ambos os sensores o algoritmo que apresentou a maior acurácia foi o Random Forest, sendo que no Kinect 360 a melhor configuração foi utilizando 1000 iterações e no Kinect One foi utilizando 2000 iterações. A diferença entre os sensores ficou em 4,08%.

Embora no Kinect 360 o Random Forest tenha tido uma discrepância maior em relação aos outros algoritmos do que nos resultados do Kinect One, todos os algoritmos de aprendizado de máquina aplicados apresentaram boa acurácia. É pertinente destacar a alta acurácia desta metodologia, pois ela representa uma metodologia livre de modelo que consegue obter boa acurácia mesmo quando o ângulo de captura do indivíduo é a lateral dele. Pois conforme discutido na Seção 2.4.1.1, a abordagem livre de modelo geralmente funciona melhor no ângulo frontal, e a através das tabelas 6 e 7 foi possível visualizar que a maior parte das pesquisas livres de modelo de fato utilizaram o ângulo frontal dos indivíduos.

A Figura 53 apresenta o comparativo do melhor resultado de cada sensor:

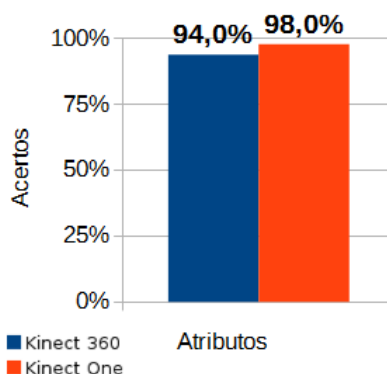


Figura 53 – Comparativo dos sensores na implementação da metodologia da Divisão em Segmentos.

4.5 Combinação de atributos

Nesta seção são apresentados os resultados de testes feitos quando foram combinados as propostas que obtiveram maior acurácia nos testes anteriores (ANDERSSON (2014) e da Divisão em Segmentos) e quando é aplicado um conjunto formado por atributos selecionados dentre todas as propostas. Para combinar as propostas e os atributos não é utilizada equação, os atributos das diferentes propostas são apenas inseridas no mesmo conjunto de treinamento e de teste.

4.5.1 Combinação das propostas de maior acurácia

Visando atingir uma quantidade maior de acertos, as duas propostas de atributos que obtiveram a melhor acurácia (proposta de ANDERSSON (2014) e da Divisão em Segmentos) foram combinadas e aplicadas aos algoritmos de aprendizado de máquina KNN, Naive Bayes e Random Fores. A rede neural MLP não foi utilizada devido ao tempo de treinamento ser muito grande em virtude da grande quantidade de atributos.

Dos três algoritmos de aprendizado de máquina utilizados, o Random Forest obteve o melhor resultado nos dados de ambos os sensores Kinect, sendo que nos dados do Kinect 360 a configuração de melhor resultado foi a de 1500 interações e nos dados do Kinect One a melhor configuração foi a de 500 interações. As tabelas 35 e 36 resumem os resultados da aplicação dos algoritmos:

Tabela 35 – Resultado da combinação das propostas de melhor acurácia no Kinect 360

Kinect 360		
Algoritmo	Configuração	Resultado
Random Forest	1500 interações	96%
KNN	K = 1	82%
Naive Bayes	Com Kernel Estimator	78%

Tabela 36 – Resultado da combinação das propostas de melhor acurácia no Kinect One

Kinect One		
Algoritmo	Configuração	Resultado
Random Forest	500 interações	98%
KNN	K = 1	82%
Naive Bayes	Com Kernel Estimator	82%

Os resultados da aplicação das configurações de maior acurácia (Random Forest de 1500 interações para os dados do Kinect 360 e Random Forest de 500 interações para os dados de Kinect One) são apresentados no gráfico da Figura 54:

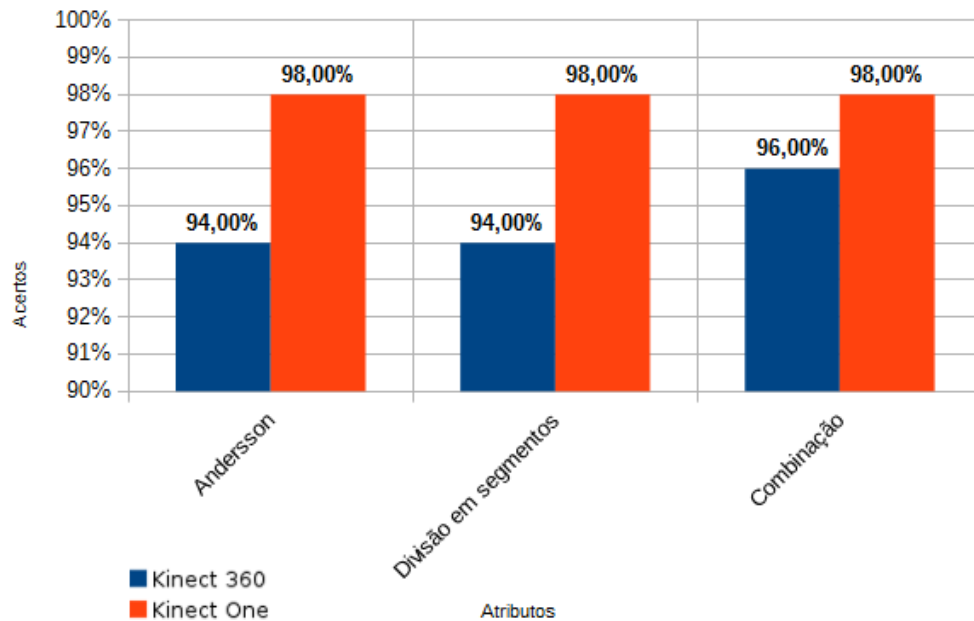


Figura 54 – Resultado da aplicação do conjunto de atributos formado pelas propostas de maior acurácia.

Comparando com a proposta da Divisão em Segmentos, é possível verificar que para o Kinect 360 a aplicação do combinação das duas metodologias fez com que a acurácia aumentasse em 2,08% e que para o Kinect One a acurácia se manteve igual. A diferença nos resultados entre um sensor e outro, na aplicação do conjunto das duas metodologias, ficou em 2,04%.

4.5.2 Combinação dos atributos que aumentam a acurácia

Neste capítulo pode-se perceber que foi analisado, após a apresentação dos resultados, o impacto que cada atributo possui dentro de cada proposta. Através destas análises, verificou-se que alguns atributos quando removidos faziam com que a acurácia melhorasse. Com base nesta constatação, foi criado um conjunto que reuniu os atributos de todas as propostas, porém sem aqueles que quando removidos faziam a acurácia aumentar.

Os algoritmos de aprendizado de máquina que foram aplicados foram o KNN, Naive Bayes e Random Forest, sendo que o MLP não foi utilizado devido ao tempo que levaria para treinar o modelo por causa da grande quantidade de atributos. As configurações aplicadas foram as mesmas descritas na Seção 3.2.3.

Na aplicação dos algoritmos de aprendizado de máquina no conjunto de validação, para o Kinect 360 o KNN apresentou a melhor acurácia utilizando $K = 1$, o Random Forest utilizando 6000 iterações e o Naive Bayes sem utilizar Kernel Estimator. Para o Kinect One, o KNN obteve a maior acurácia utilizando $K = 1$, o Random Forest utilizando 4000 iterações e o Naive Bayes sem utilizar Kernel Estimator, portanto estas foram as configurações utilizadas nesta seção.

Na tabela 37 é apresentado quais atributos foram removidos de cada proposta:

Tabela 37 – Atributos removidos

Método	Conjunto	Atributos removidos	
		Kinect 360	Kinect One
(ANDERSSON, 2014)	Antropométrico	Antebraço, lado direito do quadril, fêmur direito, fêmur esquerdo, pé direito, mão esquerda	-
(ANDERSSON, 2014)	Cinemático	-	-
(ANDERSSON, 2014)	Espaço-temporal	-	Duração do ciclo e velocidade
Caixa Delimitadora	-	-	Largura média e Área média
Divisão em Seg-mentos	-	-	-
Figuras de Movimento	-	Tempo do ciclo	Velocidade e coeficiente de correlação

As tabelas 38 e 39 apresentam os resultados:

Tabela 38 – Resultado com Kinect 360

Kinect 360		
Algoritmo	Configuração	Resultado
Random Forest	6000 iterações	74%
KNN	K = 1	56%
Naive Bayes	Sem Kernel Estimator	36%

Tabela 39 – Resultado com Kinect One

Kinect One		
Algoritmo	Configuração	Resultado
Random Forest	4000 iterações	92%
KNN	k = 1	48%
Naive Bayes	Sem Kernel Estimator	52%

É possível perceber que embora o Kinect One tenha obtido um resultado bom ao utilizar o Random Forest (92%), esta metodologia de selecionar os melhores atributos não trouxe resultados tão bons como outras propostas. O gráfico da Figura 55 apresenta a comparação entre as metodologias de extração de atributos com este conjunto que seleciona atributos de todas as propostas:

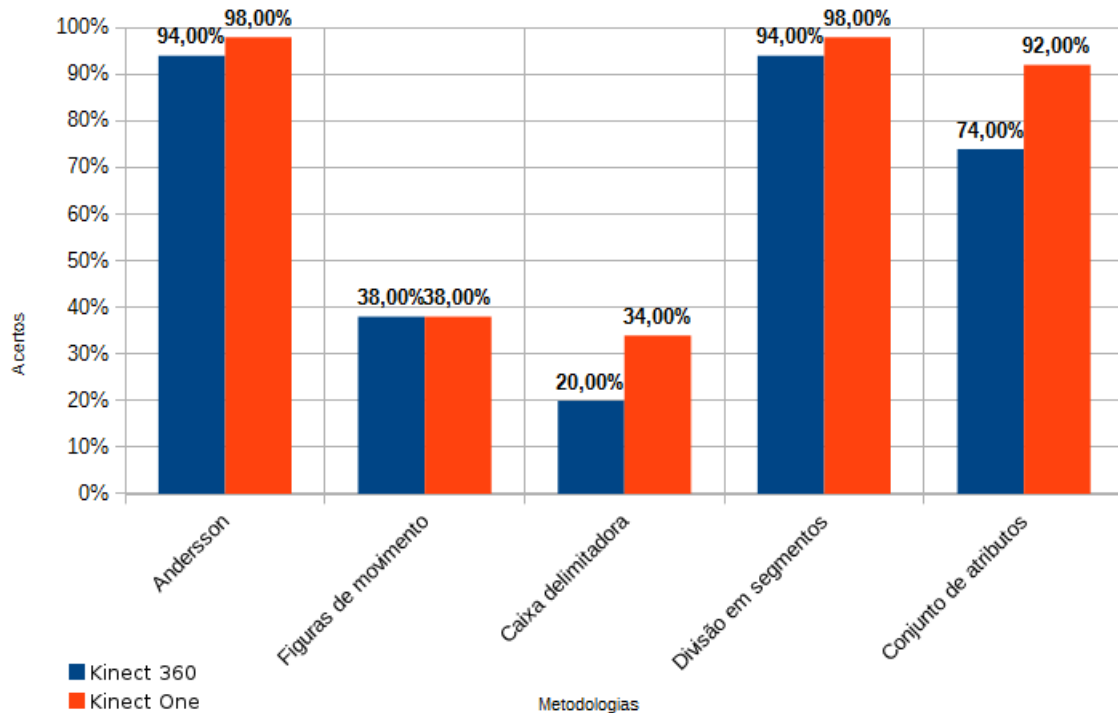


Figura 55 – Comparativo do conjunto que seleciona atributos de todas as metodologias.

4.6 Análise dos Resultados

A tabela 40 apresenta a diferença entre os sensores na implementação das quatro metodologias de extração de atributos e nas combinações das propostas.

Tabela 40 – Diferença nos resultados entre os sensores

	Kinect 360	Kinect One	Diferença
Andersson	94%	98%	4,08%
Figuras de Movimento	38%	38%	0%
Caixa Delimitadora	20%	34%	41,17%
Divisão em Segmentos	94%	98%	4,08%
Andersson + Figuras de Movimento	96%	98%	2,04%
Atributos de maior acurácia	74%	92%	19,56%

A Figura 56 apresenta o comparativo entre o Kinect 360 e o Kinect One abordado na tabela 40. A última proposta do gráfico, chamada de Conjunto de atributos, são os atributos selecionados de todas as propostas conforme descrito na Seção 4.5.2.

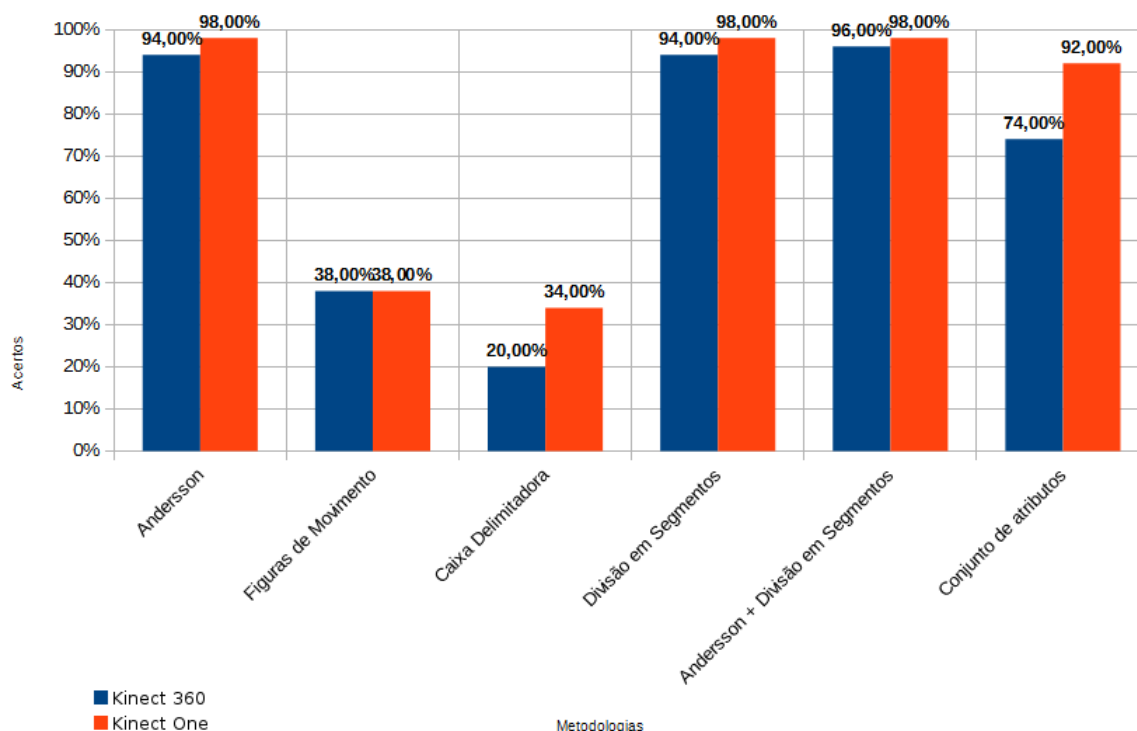


Figura 56 – Comparativo entre os resultados.

Nota-se que em todos os casos o Kinect One obteve maior quantidade de acertos. A maior diferença de resultado entre os sensores ocorre no algoritmo Caixa Delimitadora, onde o Kinect One possui uma acurácia 41,17% maior que o Kinect 360.

A proposta das Figuras de Movimento obteve o mesmo resultado em ambos os sensores (38%), sendo que depois desta proposta a metodologia em que houve menor diferença entre os sensores foi a metodologia que combinou as propostas de ANDERSSON (2014) e da Divisão em Segmentos.

A metodologia que obteve a pior quantidade de acertos foi a da Caixa Delimitadora, pois em ambos os sensores a acurácia foi inferior a 35%. Conforme abordado anteriormente, este baixo resultado se dá em grande parte pelo fato dos atributos utilizarem em seus cálculos as medidas da área que a pessoa desenvolve durante o caminhar. Devido a baixa acurácia nos resultados pode-se deduzir que os indivíduos desenvolveram medidas de área semelhantes entre eles.

A proposta de ANDERSSON (2014) e a proposta da Divisão em Segmentos foram as metodologias que individualmente obtiveram a melhor acurácia em ambos os sensores. Para o Kinect 360, ambas metodologias obtiveram 94% de acertos, e para o Kinect One, ambas metodologias obtiveram 98% de acertos. É importante notar que através da combinação destas duas metodologias foi encontrada a melhor acurácia para o Kinect 360 (96%), e para o Kinect One a melhor acurácia (98%) foi obtida tanto na combinação destas duas propostas, como na aplicação separada da metodologia de ANDERSSON (2014) e da Divisão em Segmentos.

Este resultado positivo para o Kinect One mostra que sua utilização em relação ao

Kinect 360 na identificação de pessoas é justificável, porém a baixa diferença entre os sensores pode carregar um aspecto de surpresa tendo em vista que o Kinect One possui um *hardware* relativamente melhor. Se formos comparar os resultados de maior acurácia em ambos os sensores, a diferença entre eles foi de apenas 2,04%.

4.7 Análise da dependência de pose

Aplicamos os algoritmos de aprendizado de máquina utilizando a base de treinamento com pessoas caminhando em uma direção e a base de testes com indivíduos caminhando em outra. O objetivo é saber se é possível identificar pessoas que estão caminhando em sentido diferente da base de exemplos. O método de validação e divisão dos dados é o mesmo descrito na Seção 3.3, a diferença ocorre que nos conjuntos de treinamento é utilizado exemplos de pessoas caminhando para a direita, enquanto no conjunto de testes estas mesmas pessoas estão caminhando para a esquerda.

4.7.1 Proposta de Andersson

Para esta metodologia foi aplicado um conjunto de dados contendo todos os atributos desta proposta (antropométricos, cinemáticos e espaço-temporais). As tabelas 41 e 42 apresentam os resultados de maior acurácia obtidos em cada algoritmo de aprendizado de máquina aplicado.

Tabela 41 – Análise da dependência de pose utilizando a proposta de ANDERSSON (2014) com Kinect 360

Kinect 360		
Algoritmo	Configuração	Resultado
Random Forest	2500 iterações	34%
KNN	K = 8	8%
Naive Bayes	Sem Kernel Estimator	6%
MLP	2000 iterações	18%

Tabela 42 – Análise da dependência de pose utilizando a proposta de ANDERSSON (2014) com Kinect One

Kinect One		
Algoritmo	Configuração	Resultado
Random Forest	2000 iterações	56%
KNN	K = 1	18%
Naive Bayes	Sem Kernel Estimator	8%
MLP	1000 iterações	30%

O resultado de maior acurácia para o Kinect 360 (34%) foi utilizando o Random Forest com 2500 iterações e para o Kinect One o resultado de maior acurácia foi utilizando o Random

Forest (56%) com 2000 iterações.

4.7.2 Proposta de Figuras de Movimento

Nas tabelas 43 e 44 é possível visualizar os resultados de maior acurácia obtidos em cada algoritmo de aprendizado de máquina aplicado utilizando a proposta de Figuras de Movimento.

Tabela 43 – Análise da dependência de pose utilizando a proposta de Figuras de Movimento com Kinect 360

Kinect 360		
Algoritmo	Configuração	Resultado
Random Forest	3500 iterações	16%
KNN	K = 8	18%
Naive Bayes	Com Kernel Estimator	20%
MLP	1000 iterações	10%

Tabela 44 – Análise da dependência de pose utilizando a proposta de Figuras de Movimento com Kinect One

Kinect One		
Algoritmo	Configuração	Resultado
Random Forest	500 iterações	14%
KNN	K = 10	10%
Naive Bayes	Com Kernel Estimator	10%
MLP	500 iterações	10%

No Kinect 360 o resultado de maior acurácia (20%) foi alcançado utilizando Naive Bayes com Kernel Estimator e no Kinect One o resultado de maior acurácia (14%) foi obtido utilizando Random Forest de 500 iterações.

4.7.3 Proposta da Caixa Delimitadora

São apresentados nas tabelas 45 e 46 os resultados de maior acurácia obtidos aplicando a metodologia da Caixa Delimitadora para a análise da dependência de pose.

Tabela 45 – Análise da dependência de pose utilizando a proposta da Caixa Delimitadora com Kinect 360

Kinect 360		
Algoritmo	Configuração	Resultado
Random Forest	5500 iterações	14%
KNN	K = 11	8%
Naive Bayes	Sem Kernel Estimator	14%
MLP	1000 iterações	18%

Tabela 46 – Análise da dependência de pose utilizando a proposta da Caixa Delimitadora com Kinect One

Kinect One		
Algoritmo	Configuração	Resultado
Random Forest	500 iterações	34%
KNN	K = 1	18%
Naive Bayes	Com Kernel Estimator	26%
MLP	3000 iterações	20%

Nos dados do Kinect 360 o resultado de maior acurácia (18%) foi obtido utilizando MLP com 1000 iterações e com os dados do Kinect One o resultado de maior acurácia (34%) foi obtido utilizando Random Forest com 500 iterações.

4.7.4 Proposta de Divisão em Segmentos

As tabelas 47 e 48 apresentam os resultados de maior acurácia obtidos em cada um dos algoritmos de aprendizado de máquina aplicados nos atributos extraídos através da metodologia de Divisão em segmentos.

Tabela 47 – Análise da dependência de pose utilizando a proposta de Divisão em segmentos com Kinect 360

Kinect 360		
Algoritmo	Configuração	Resultado
Random Forest	2500 iterações	28%
KNN	K = 2	16%
Naive Bayes	Sem Kernel Estimator	10%

Tabela 48 – Análise da dependência de pose utilizando a proposta de Divisão em segmentos com Kinect One

Kinect One		
Algoritmo	Configuração	Resultado
Random Forest	5000 iterações	34%
KNN	K = 7	8%
Naive Bayes	Com Kernel Estimator	14%

Para os dados do Kinect 360 o resultado de maior acurácia (28%) foi obtido utilizando Random Forest com 2500 iterações e com os dados do Kinect One o resultado de maior acurácia (34%) foi obtido utilizando Random Forest com 5000 iterações.

4.7.5 Análise dos resultados da dependência de pose

A Figura 57 apresenta os resultados de maior acurácia obtidos em cada uma das metodologias de extração de atributos aplicadas.

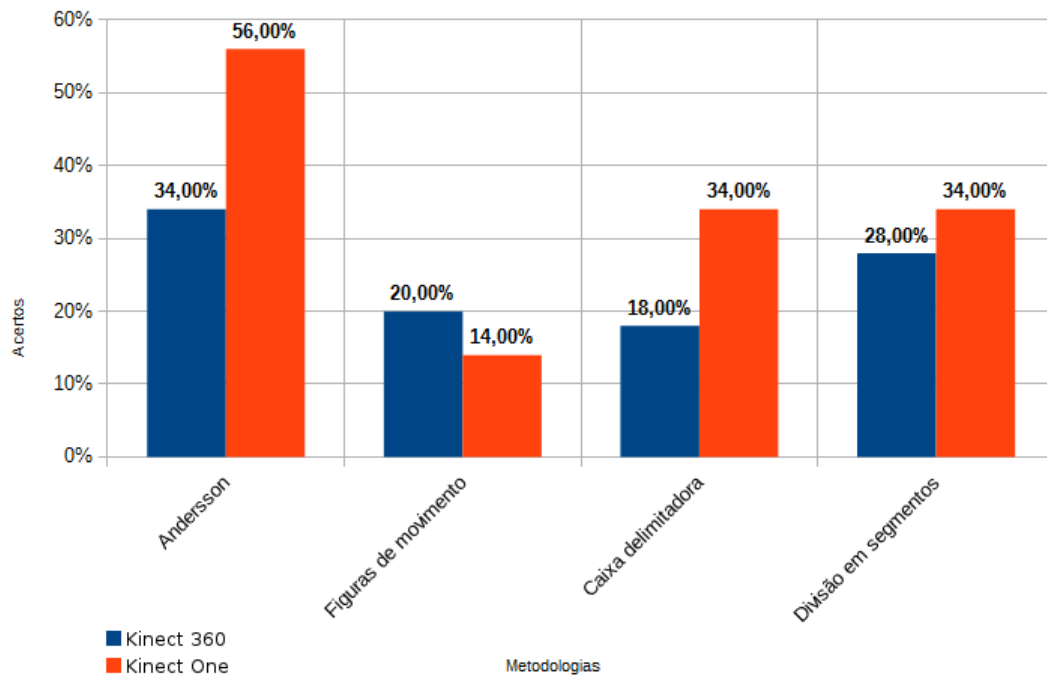


Figura 57 – Comparativo dos resultados da análise de dependência de pose.

Conforme pode-se perceber na baixa acurácia apresentada no gráfico, o número de acertos caiu bruscamente quando os conjuntos de treinamento continham pessoas caminhando em direção oposta a dos conjuntos de testes. O único algoritmo que obteve acurácia superior a 50% foi o da proposta de ANDERSSON (2014) e apenas quando aplicado nos dados do Kinect One.

Esta baixa acurácia nos resultados nos mostra que no caso implementado neste trabalho existe uma grande dependência de pose até mesmo nos algoritmos aplicados que são baseados em modelo (a metodologia de ANDERSSON (2014) e a das Figuras de Movimento), trazendo para o campo da discussão a busca por maneiras de diminuir esta necessidade de se ter pessoas no conjunto de treinamento e de teste caminhando na mesma direção. Uma das soluções, que pode ser investigada em trabalhos futuros, é incluir no conjunto de treinamento os indivíduos caminhando em ambas as direções.

5 CONCLUSÃO

Através da utilização de quatro propostas de extração de atributos foi possível obter algumas conclusões em relação a diferença entre as versões do Kinect na análise biométrica.

Por meio da análise dos atributos antropométricos da metodologia de ANDERSSON (2014), verificou-se que o Kinect 360 possui maior dificuldade em obter as posições da parte inferior do esqueleto das pessoas. Também através da análise destes atributos foi possível verificar que dois dos novos pontos do esqueleto que são capturados exclusivamente pelo Kinect One (pescoço e polegar esquerdo) agregaram valor aos resultados.

Com a aplicação da proposta das Figuras de Movimento, identificou-se que o Kinect One teve dificuldades em caracterizar os indivíduos pela diferença entre os ângulos de uma perna e outra, isto deve provavelmente por causa da nova configuração do quadril que o Kinect One apresenta.

Foi possível verificar que a utilização individual do conjunto de atributos espaço-temporais não traz boa acurácia, e que ao utilizá-los com outros tipos de atributos, como antropométricos e cinemáticos, eles agregam pouca melhora ao número de acertos.

Outro ponto observado, é a sugestão de que o Kinect One obtém com melhor precisão a posição dos calcanhares durante o caminhar. Esta sugestão ocorre porque verificou-se a dificuldade do Kinect 360 em calcular o tempo do ciclo na proposta das Figuras de Movimento, sendo que o tempo do ciclo é obtido por meio da posição dos calcanhares, além disto, observou-se que o ângulo de maior impacto nos resultados do Kinect One na proposta de ANDERSSON (2014) foi o formado pelo calcanhar e pé esquerdos e de que a distância do calcanhar e pé direitos no Kinect 360 traz impacto negativo aos resultados.

Por meio da metodologia da Caixa Delimitadora, verificou-se que caracterizar o caminhar dos indivíduos pela área não trouxe bons resultados, sendo que no Kinect One a acurácia foi ainda mais baixa. Esta baixa taxa de acertos sugere que para a identificação de pessoas com o Kinect utilizar a área como meio de caracterização não é uma opção razoável.

Em relação aos algoritmos de aprendizado de máquina, o Random Forest foi o algoritmo que mais se destacou, pois na maioria das propostas ele obteve a maior precisão.

Separadamente, as metodologias de ANDERSSON (2014) e da Divisão em Segmentos foram as que apresentaram a melhor acurácia em ambos os sensores. Sendo que através da

combinação destas duas propostas foi possível encontrar o maior resultado para o Kinect 360 (96%). A melhor acurácia obtida com os dados do Kinect One (98%) foi encontrada tanto na combinação destas duas propostas, como na aplicação separada da metodologia da Divisão em Segmentos.

Por meio da análise da maior quantidade de acertos obtidos no Kinect 360 (96%) e no Kinect One (98%), constata-se que para a identificação por meio do caminhar, o Kinect One é a melhor opção. Embora é possível observar através da pequena diferença entre os sensores, de apenas 2%, que a utilização do Kinect One não traz uma melhora muito significativa.

Através da investigação da dependência de pose, constatou-se que utilizar o conjunto de treinamento com pessoas caminhando em uma direção e o conjunto de testes com pessoas caminhando em outra trouxe resultados com acurácia baixa. O melhor resultado foi obtido utilizando a proposta de ANDERSSON (2014), onde o Kinect One obteve acurácia de 56% e o Kinect 360 de 34%. O Kinect One se saiu melhor em todas as metodologias testadas, com exceção da metodologia das Figuras de Movimento, onde o Kinect 360 obteve acurácia 30% superior.

Como trabalhos futuros, é proposto uma investigação da utilização de pessoas caminhando em ambas as direções no conjunto de treinamento, com o fim de tentar-se obter uma dependência de pose menor. Também propõe-se a formulação e aplicação de uma metodologia que utilize para identificar as pessoas os dois sensores em conjunto, verificando se é possível obter resultados melhores combinando atributos extraídos de ambas as versões do Kinect. Por fim, propõe-se realizar uma comparação dos sensores na identificação de gênero dos indivíduos.

REFERÊNCIAS

- ABERDEEN, R. G. U. **The Back Propagation Algorithm**. [Online; acessado em 26 de novembro de 2015], <https://www4.rgu.ac.uk/files/chapter3%20-%20bp.pdf>.
- AHMED, F.; PAUL, P.; GAVRILOVA, M. DTW-based kernel and rank-level fusion for 3D gait recognition using Kinect. **The Visual Computer**, v.31, n.6, p.915–924, 2015.
- ALEXANDER, R. N. Walking and running. **American Scientist**, v.72, p.438–354, 1984.
- ANDERSSON, V. O. **Identificação Biométrica com Antropometria e Caminhar Humano Utilizando o kinect**. 2014. Dissertação (Mestrado em Ciência da Computação) — Universidade Federal de Pelotas, Pelotas.
- ARAUJO, R.; GRAÑA, G.; ANDERSSON, V. Towards Skeleton Biometric Identification Using the Microsoft Kinect Sensor. **the 28th Annual ACM Symposium**, p.21–26, 2013.
- ARORA, P.; SRIVASTAVA, S.; ARORA, K.; BAREJA, S. Metabolic Cost of Generating Force During Human Leg Swing. **Second International Symposium on Computer Vision and the Internet (VisionNet'15)**, 2015.
- ARORA, P.; SRIVASTAVA, S.; ARORA, K.; BAREJA, S. Improved Gait Recognition using Gradient Histogram Gaussian Image. **Second International Symposium on Computer Vision and the Internet (VisionNet'15)**, v.58, p.408–413, 2015.
- BAKER, R. The history of gait analysis before the advent of modern computers. **Gait and Posture**, 2006.
- BALAZIA, M. **Human Gait Recognition Based on Body Component Trajectories**. 2013. Dissertação (Mestrado em Ciência da Computação) — Masaryk University, Porto, Portugal.
- BENABDELKADER, C.; CUTLER, R.; DAVIS, L. **Stride and Cadence as a Biometric in Automatic Person Identification and Verification**. [Online; acessado em 04 de novembro de 2015], http://www.cs.umd.edu/~chiraz/fg_stride.pdf.
- BEUNEN, G. Physical Growth, Maturation and Performance. **Kinanthropometry and Exercise Physiology Laboratory Manual**, v.1, n.3, 2009.

BOBICK, A. F.; JOHNSON, Y. A. Gait Recognition Using Static, Activity-Specific Parameters. **Computer Society Conference on Computer Vision and Pattern Recognition**, 2001.

BORELLI, G. **On the movement of animals**. Translated by Maquet P. Springer-Verlag, Berlin: [s.n.], 1989.

BREIMAN, L.; CUTLER, A. **Random Forests**. [Online; acessado em 28 de fevereiro de 2017], https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm.

BRUNIEIRA, C. A. V. **Andž"lise Biomecdž"nica da Locomodž"dž"o Humana**. [Online; acessado em 16 de outubro de 2015], <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2816028/>.

BW, R. **Método de Validação Cruzada**. [Online; acessado em 04 de março de 2017], <http://www.revistabw.com.br/revistabw/metodo-de-validacao-cruzada/>.

CANALTECH. **Como funciona o Kinect?** [Online; acessado em 19 de outubro de 2015], <http://canaltech.com.br/o-que-e/kinect/Como-funciona-o-Kinect/>.

CARVALHO, A. R. d.; ANDRADE, A.; PEYRÉ-TARTARUGA, L. A. High-level language abstraction for reconfigurable computing. **Revista Brasileira de Reumatologia**, v.55, n.1, 2015.

CASIA. **Gait database offered by Chinese Academy of Sciences**. <http://www.sinobiometrics.com>.

CENTRAL, D. S. **Random Forests Algorithm**. [Online; acessado em 28 de fevereiro de 2017], <http://www.datasciencecentral.com/profiles/blogs/random-forests-algorithm>.

CHATTOPADHYAY, P.; ROY, A.; SURAL, S.; MUKHOPADHYAY, J. Pose Depth Volume extraction from RGB-D streams for frontal gait recognition. **Journal of visual communication and image representation**, v.25, p.53–63, 2013.

CHATTOPADHYAY, P.; SURAL, S.; MUKHERJEE, J. Frontal gait recognition from occluded scenes. **Patterns Recognition Letters**, v.63, p.9–15, 2015.

CHEN, J. Gait Correlation Analysis Based Human Identification. **The Scientific World Journal**, v.2014, 2014.

CNET. **Timeline: A look back at Kinect's history**. [Online; acessado em 19 de dezembro de 2016], <https://www.cnet.com/news/timeline-a-look-back-at-kinects-history>.

CNET. **Microsoft Xbox 360 Kinect review**.. [Online; acessado em 17 de março de 2017], <https://www.cnet.com/uk/products/microsoft-xbox-360-kinect/review/>.

DCJS, N. **The New York State Bertillon Bureau**. [Online; acessado em 13 de outubro de 2015], http://www.criminaljustice.ny.gov/ojis/history/bert_ny.htm.

DELGADO, L. A. **Avaliação da Composição Corporal**. [Online; acessado em 05 de outubro de 2015]. Disponível em: <<http://www.ebah.com.br/content/ABAAAAdJAAF/avaliacao-fisica-composicao-corporal>>.

EXTREMETECH. **Microsoft begs devs to pay attention to Kinect 2 for Windows, even after Xbox One de-bundling**. [Online; acessado em 25 de outubro de 2015], <http://www.extremetech.com/gaming/183913-microsoft-begs-devs-to-pay-attention-to-kinect-2-for-windows-even-after-xbo>

GERSHENSON, C. **Artificial Neural Networks for Begginers**. [Online; acessado em 26 de novembro de 2015], <https://datajobs.com/data-science-repo/Neural-Net-%5BCarlos-Gershenson%5D.pdf>.

GFBIOINFO. **Árvores de Decisão**. [Online; acessado em 28 de fevereiro de 2017], <http://web.tecnico.ulisboa.pt/ana.freitas/bioinformatics.ath.cx/bioinformatics.ath.cx/indexf23d.html?id>.

GHOLAMI, F. et al. Gait Assessment for Multiple Sclerosis Patients Using Microsoft Kinect. **Computer Vision and Pattern Recognition**, v.abs/1508.02405, 2015.

GISSIC. **Grupo de Investigacion en Seguridad y Sistemas de Comunicaciones**. [Online; acessado em 18 de dezembro de 2016], <http://gissic.umng.edu.co>.

GROSS, R.; SHI, J. **The CMU Motion of Body (MoBo) Database**. Pittsburgh, PA: Robotics Institute, 2001. (CMU-RI-TR-01-18).

GUILLEN, E.; PADILLA, D.; HERNANDEZ, A.; BARNER, K. Gait Recognition System: Bundle Rectangle Approach. **World Academy of Science, Engineering and Technology**, v.58, p.696–702, 2009.

GUILLEN, E.; PADILLA, D.; HERNANDEZ, A.; BARNER, K. Gait Recognition System: Bundle Rectangle Approach. **International Science Index, Electrical and Computer Engineering**, v.3, n.10, p.1846–1852, 2009.

HAYKIN, S. **Redes Neurais: Princípios e prática**. Porto Alegre: Bookman, 2001. v.2nd Edition.

HEALTH, U. N. L. of Medicine National Institute of. **Dynamic Principles of Gait and Their Clinical Implications**. [Online; acessado em 13 de outubro de 2015], <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2816028/>.

HECHT-NIELSEN, R. Theory of the backpropagation neural network. **Neural Networks**, v.1, Supplement 1, p.445 –, 1988. {INNS} First Annual Meeting.

HOWSTUFFWORKS. **O software Kinect aprende com a experiência**. [Online; acessado em 19 de outubro de 2015], <http://tecnologia.hsw.uol.com.br/microsoft-kinect2.htm>.

HUANG, X.; Boulgouris, V. Gait Recognition Using Linear Discriminant Analysis with Artificial Walking Conditions. **Proc. IEEE Int. Conf. on Image Processing**, v.2, p.2461–2464, 2010.

KASTANIOTIS, D.; THEODORAKOPOULOS, I.; C., T.; ECONOMOU, G. A framework for gait-based recognition using Kinect. **Patterns Recognition Letters**, v.9, n.6, p.327–335, 2015.

KUSAKUNNIRAN, W.; WU, Q.; L., H.; ZHANG, J. Multiple views gait recognition using View Transformation Model based on optimized Gait Energy Image. **International Conference on Computer Vision Workshops**, p.1058–1064, 2009.

LEICESTER, U. of. **History of DNA Profiling**. [Online; acessado em 30 de novembro de 2015], <http://www2.le.ac.uk/departments/emfpu/genetics/explained/profiling-history>.

LIN, C.; WANG, K. A Behavior Classification Based on Enhanced Gait Energy Image. **Proc. Int. Conf. on Networking and Digital Society**, v.2, p.589–592, 2010.

LORENA, A. C.; CARVALHO, A. C. P. L. F. d. **Uma Introdução às Support Vector Machines**. [Online; acessado em 26 de novembro de 2015], www.seer.ufrgs.br/rita/article/download/rita_v14_n2_p43-67/3543.

MASTER, C. **Kinect SDK: Skeleton Tracking**. [Online; acessado em 08 de março de 2017], <http://www.contentmaster.com/kinect/kinect-sdk-skeleton-tracking/>.

MELLON, U. C. **CMU Graphics Lab Motion Capture Database**. [Online; acessado em 18 de dezembro de 2016], <http://mocap.cs.cmu.edu>.

MICROSOFT. **Kinect for Xbox 360**. [Online; acessado em 18 de outubro de 2015], <http://www.xbox.com/en-US/xbox-360/accessories/kinect>.

MICROSOFT. **Kinect for Xbox One**. [Online; acessado em 30 de outubro de 2015], <http://www.xbox.com/pt-BR/Xbox-One/accessories/kinect-for-xbox-one>.

MICROSOFT. **The Future of Kinect**. [Online; acessado em 25 de outubro de 2015], <http://news.microsoft.com/features/the-future-of-kinect/>.

MICROSOFT. **Skeletal Joint Smoothing White Paper**: Introduction. [Online; acessado em 05 de novembro de 2016], <https://msdn.microsoft.com/en-us/library/jj131429.aspx#ID4E5>.

MICROSOFT. **Skeletal Joint Smoothing White Paper**: Auto Regressive Moving Average (ARMA) Filters. [Online; acessado em 05 de novembro de 2016], <https://msdn.microsoft.com/en-us/library/jj131429.aspx#ID4EZHAC>.

MICROSOFT. **Kinect for Windows Sensor Components and Specifications**. [Online; acessado em 17 de março de 2017], <https://msdn.microsoft.com/en-us/library/jj131033.aspx>.

MICROSOFT. **JointType Enumeration**. [Online; acessado em 08 de março de 2017], <https://msdn.microsoft.com/en-us/library/microsoft.kinect.jointtype.aspx>.

MICROSOFT. **Algoritmo Naive Bayes da Microsoft**. [Online; acessado em 27 de março de 2017], <https://msdn.microsoft.com/pt-br/library/ms174806.aspx>.

MICROSOFT. **Hardware Kinect**. [Online; acessado em 17 de março de 2017], <https://developer.microsoft.com/pt-br/windows/kinect/hardware>.

MILOVANOVIC, M.; MINOVIC, M.; STARCEVIC, D. Walking in Colors: Human Gait Recognition Using Kinect and CBIR. **Multimedia**, v.20, p.28–36, 2013.

MITCHELL, T. M. **Machine Learning**. New York: McGraw-Hill, 1997.

MUEHLENBEIN, M. P. **Human Evolutionary Biology**. [Online; acessado em 05 de outubro de 2015], https://books.google.com.br/books?id=3NRf_8gwm08C&pg=PA92&lpg=PA92&dq=anthropometric+history+Johann+Sigismund+Elsholtz&source=bl&ots=BXS9fBgEGZ&sig=DQwpFks53Ev6j81LQAfFDQEDnms&hl=pt-BR&sa=X&ved=0CDQ6AEwBGovChMIjq0tiYusyAIVwRyQCh3HzAvC#v=onepage&q=Johann%20Sigismund%20Elsholtz&f=false.

MUNSELL, C. B.; TEMLYAKOV, A.; QU, C.; WANG, S. Person Identification Using Full-Body Motion and Anthropometric Biometrics from Kinect Videos. **ECCV 2012**, p.91–100, 2012.

MURRAY, M. P.; DROUGHT, A. B.; KORY, R. C. Walking Patterns of Normal Man. **Journal of Bone and Joint Surgery**, v.46-A, n.2, p.438–354, 1964.

NASH, M. J.; CARTER, J. N.; NIXON, M. S. Dynamic feature extraction via the velocity Hough transform. **Pattern Recognition Letter**, v.18, p.1035–1047, 1997.

NEAPOLITAN, E. R. **Learning Bayesian Networks**. New Jersey, EUA: Pearson Prentice Hall, 2004.

NETO, H. **Os Determinantes da marcha**. [Online; acessado em 16 de outubro de 2015], <http://fisioterapiahumberto.blogspot.com.br/2010/02/os-determinantes-dademarçoa.html>.

NEWTON, I. *The principia: mathematical principles of natural philosophy*. Translated by Cohen I., Whitman A. Berkeley. **University of California Press**, 1999.

ORTHOTISTS, A. A. of; PROTHETISTS. **Determinants of Gait**. [Online; acessado em 29 de novembro de 2015], http://www.oandp.org/olc/course_extended_content.asp?frmCourseId=29CC44FD-D215-48B7-A53D-F130AFAB8087&frmTermId=determinants_of_gait.

OSAKA, U. de. **OU-ISIR Biometric Database**. [Online; acessado em 18 de dezembro de 2016], <http://www.am.sanken.osaka-u.ac.jp/BiometricDB/GaitTM.html>.

PASQUALI, L.; ARAÚJO, M. E. C. **Histórico dos Processos de Identificação**. v.2nd Edition. [Online; acessado em 23 de março de 2017]. Disponível em: <<http://www.papiloscopistas.org/historico.doc>>.

PETERSON, L. E. K-nearest neighbor. **Scholarpedia**, v.4, n.2, p.1883, 2009. revision 136646.

PETROSKI, E. L. **Desenvolvimento e validação de equações generalizadas para a estimativa da densidade corporal em adultos**. 1995. Dissertação (Mestrado em Ciência da Computação) — Universidade Federal de Santa Maria, Santa Maria.

PREIS, J.; KESSEL, M.; WERNER, M.; LINNHOF-POPIEN, C. Gait Recognition with Kinect. **Proceedings of the First Workshop on Kinect in Pervasive Computing**, 2012.

PROCHÁZKA, A. et al. Bayesian classification and analysis of gait disorders using image and depth sensors of Microsoft Kinect. **Digital Signal Processing**, v.47, p.169–177, 2015.

RAHEJA, J.; CHAUDHARY, A.; K., N.; MAITI, S. Pre-consultation help necessity detection based on gait recognition. **Patterns Recognition Letters**, v.68, n.2, p.1357–1363, 2014.

RUIZ-GARZÓN, G. Bertillon and Galton. Probabilistic arguments related to the identification on criminals. **Boletín de Estadística e Investigación Operativa**, v.29, n.2, p.129–140, 2013.

SARKAR, S. et al. The HumanID Gait Challenge Problem: Data Sets, Performance, and Analysis. **Transactions on Pattern Analysis and Machine Intelligence**, v.27, n.2, p.162–177, 2005.

SAUNDERS, J. B.; INMAN, V. T.; EBERHART, H. D. The major determinants in normal and pathological gait. **General Surgery**, v.35, n.A(3), p.35–A(3):543–558, 1953.

SHOTTON, J. et al. Real-Time Human Pose Recognition in Parts from a Single Depth Image. In: CVPR, 2011. **Anais... IEEE**, 2011.

SHUTLER, J. D.; GRANT, M. G.; NIXON, M. S.; CARTER, J. N. On a large sequence-base human gait database. **Proc. of 4th International Conference on Recent Advances in Soft Computing**, v.4, p.66–71, 2002.

SINFIC. **História Geral da Biometria**. [Online; acessado em 31 de outubro de 2015], <http://www.sinfic.pt/SinficWeb/displayconteudo.do2?numero=25030>.

SINHA, A.; CHAKRAVARTY, K.; BHOWMICK, B. Person Identification using Skeleton Information from Kinect. **The Sixth International Conference on Advances in Computer-Human Interactions**, Nice, França, p.102–108, 2013.

SIVAPALAN, S. et al. Gait Energy Volumes and Frontal Gait Recognition using Depth Images. **Internation Joint Conference on Biometrics**, p.1–6, Oct 2011.

SOLBERG, A. **INF 4300 - Hough Transform**. [Online; acessado em 18 de novembro de 2015], <http://www.uio.no/studier/emner/matnat/ifi/INF4300/h09/undervisningsmateriale/hough09.pdf>.

SOUSA, A.; TAVARES, J.; E., M.; SOUSA, F. Análise Da marcha Baseada Numa Correlação Multifactorial. **Atas do 3º Congresso Nacional de Biomecânica**, p.11–16, 2009.

SOUZA, A. S. P. d. **Análise da marcha Baseada em Correlação Multifactorial**. 2009. Dissertação (Mestrado em Ciência da Computação) — Faculdade de Engenharia da Universidade do Porto, Porto, Portugal.

SOUZA, A.; TAVARES, J. M. R. S. **A marcha humana: uma abordagem biomecânica**. [Online; acessado em 14 de outubro de 2015], https://web.fe.up.pt/~tavares/downloads/publications/artigos/IHC_2010_AS.pdf.

STEVENAGE, S. V.; NIXON, M. S.; K., V. Visual analysis of gait as a cue to identity. **Applied Cognitive Psychology**, v.13, n.6, p.513–526, 1999.

SUTTON, O. **Introduction to k Nearest Neighbour Classification and Condensed Nearest Neighbour Data Reduction**. [Online; acessado em 25 de novembro de 2015], http://www.math.le.ac.uk/people/ag153/homepage/KNN/OliverKNN_Talk.pdf.

TANAWONGSUW R.; BOBICK, A. **Gait Recognition from Time-normalized Joint-angle Trajectories in the Walking Plane**. [Online; acessado em 07 de novembro de 2015], http://www.cc.gatech.edu/cpl/projects/hid/papers/cvpr01_final.pdf.

TECCIENCIA. **Equações de 2o grau no nosso cotidiano**. [Online; acessado em 13 de dezembro de 2016], <http://tecciencia.ufba.br/equacao-do-2o-grau/equacoes-de-2o-grau-no-nosso-cotidiano>.

UPDATE, B. **History of Biometrics**. [Online; acessado em 13 de outubro de 2015]. Disponível em: <<http://www.biometricupdate.com/201501/history-of-biometrics>>.

VAUGHAN, C. L.; DAVID, B. L.; O'CONNOR, J. C. **Dynamics of Human Gait**. [S.l.]: Kiboho Publishers, 1999. v.2nd Edition.

VENKAT, I.; WILDE, P. Robust Gait Recognition by Learning and Exploiting Sub-gait Characteristics. **International Journal of Computer Vision**, v.1, n.91, p.7–23, 2010.

VERGE, T. **Cracking Kinect**: the Xbox One's new sensor could be a hardware hacker's dream. [Online; acessado em 25 de outubro de 2015], <http://www.theverge.com/2013/5/31/4383114/hacking-the-new-microsoft-kinect>.

WANG, J.; SHE, M.; NAHAVANDI, S.; KOUZANI, A. A Review of Vision-based Gait Recognition Methods for Human Identification. **IEEE**, p.320–327, 2010.

WANG, L.; NING, H.; TAN, T.; HU, W. Fusion of Static and Dynamic Body Biometrics for Gait Recognition. **Transcations on Circuits and System for Video Technology**, v.14, p.149–158, 2005.

WEKA. **Waikato Environment for Knowledge Analysis**. [Online; acessado em 07 de março de 2017], <http://www.cs.waikato.ac.nz/ml/>.

XBOX. **Kinect para Xbox One**. [Online; acessado em 17 de março de 2017], <http://www.xbox.com/pt-BR/xbox-one/accessories/kinect>.

YAM, C.; NIXON, M. S. **Model-based Gait Recognition**. [Online; acessado em 22 de novembro de 2015], http://eprints.soton.ac.uk/268238/1/Model-based_Gait_Recognition_final.pdf.

YAM C. NIXON, M. S.; CARTER, J. N. Automatic extraction and description of human gait models for recognition purposes. **Computer Vision and Image Understanding**, v.90, n.1, p.1–41, 2003.

YAM, C.; NIXON, M. S.; CARTER, J. N. Automated person recognition by walking and running via model-based approaches. **Pattern Recognition**, v.37, n.5, p.1057–1072, 2004.

YEOH, T. et al. Improved Gait Recognition with Automatic Body Joint Identification. **Visual Informatics: Sustaning Research and Innovations**, p.245–256, 2011.

YESHION, T. Anthropometry: The first system of identification. **The Forence Teacher Maganize**, n.23, p.46–51, 2014.

YOO, J.-H.; NIXON, M. S. Automated Markerless Analysis of Human Gait Motion for Recognition and Classification. **ETRI Journal**, v.33, n.2, p.259–266, 2011.

YOO JANG-HEE; MOON, K.-Y.; HWANG DOOSUNG; NIXON, M. S. Automated Human Recognition by Gait using Neural Network. **IEEE**, p.1–6, 2008.

APÊNDICE A ORGANIZAÇÃO DOS DADOS

Os dados resultantes das capturas estão disponíveis em <https://github.com/lw-dias/kinect-data>. A estrutura dos arquivos neste endereço está organizada de forma que cada pessoa capturada possui uma pasta contendo três arquivos, um deles descreve as características da pessoa, outro contém a delimitação dos ciclos e por fim o terceiro possui os dados resultantes da captura.

Nas próximas sub-seções será abordado cada tipo de arquivo individualmente, objetivando um melhor esclarecimento de cada um.

A.1 Arquivo contendo a delimitação dos trajetos

De cada caminhada realizada pelo indivíduo capturado, procurou-se extrair 5 ciclos do caminhar, ou seja, 5 ciclos do trajeto percorrido pela pessoa caminhando para a direita e 5 ciclos do trajeto que a pessoa percorreu indo para a esquerda. Para que fosse extraído estes ciclos do caminho percorrido pelas pessoas capturadas, foi desenvolvido na linguagem de programação Javascript uma página para a visualização dos dados gravados, a Figura 58 foi extraída desta página criada e mostra um dado momento de uma das capturas:

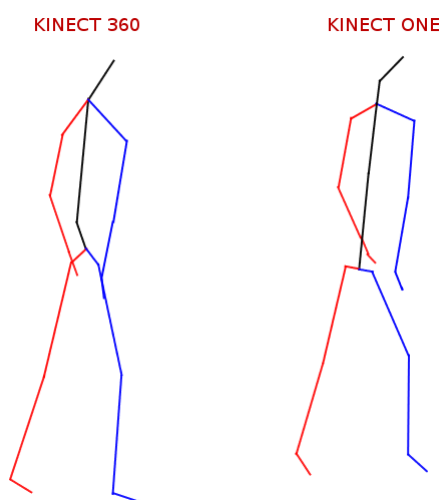


Figura 58 – Dados dos sensores Kinect exibidos na aplicação desenvolvida.

A delimitação dos ciclos foi feita de forma manual pois através da observação de cada captura procurou-se delimitar os quadros que correspondiam a 5 ciclos de cada trajeto percorrido. Estas informações de delimitação dos ciclos foram salvas em um arquivo que posteriormente foi lido pelo aplicativo que foi desenvolvido para extrair os atributos dos indivíduos.

O arquivo de delimitação dos trajetos está organizado de forma que em cada linha está presente o identificador do indivíduo, seguido da delimitação da trajetória para a direita (quadro inicial, quadro final) e após a delimitação da trajetória para a esquerda (quadro inicial, quadro final). A Figura 59 ilustra a organização do arquivo:

	Trajeto para a direita	Trajeto para a esquerda
Pessoa 1,	Quadro inicial,Quadro final,	Quadro inicial,Quadro final
Pessoa 2,	Quadro inicial,Quadro final,	Quadro inicial,Quadro final
Pessoa 3,	Quadro inicial,Quadro final,	Quadro inicial,Quadro final
...		

Figura 59 – Ilustração do arquivo contendo a delimitação dos ciclos.

A.2 Arquivo contendo os dados das capturas

O aplicativo gera para cada captura realizada dois arquivos diferentes, um com os dados do Kinect One e o outro com os dados do Kinect 360. Cada arquivo possui um cabeçalho que descrevem os dados. Neste cabeçalho aparece o nome de cada junta capturada e a posição no espaço que aquela coluna representa (x, y ou z). Como o Kinect 360 captura 20 juntas do esqueleto humano, o cabeçalho de uma captura deste sensor possui separado por vírgula 60 palavras ($20 * 3$), já o cabeçalho do Kinect One possui 75 palavras, pois o sensor captura 25 pontos do esqueleto humano.

Um exemplo do início do cabeçalho do arquivo do Kinect One seria Head_X,Head_Y,Head_Z,Neck_X,Neck_Y,Neck_Z,SpineShoulder_X,SpineShoulder_Y,SpineShoulder_Z...

Logo abaixo do cabeçalho, são gravados os dados, onde cada linha representa um quadro gravado. Seguindo a ordem descrita no cabeçalho, são separadas por virgulas os números que representam as posições de cada junta.

A Figura 60 ilustra como são organizados os dados nos arquivos:

Kinect 360

```
Head_x,Head_y,Head_z,Shoulder_center_x,Shoulder_center_y,Shoulder_center_z...
0.819202,0.663560,2.838875,0.865644,0.501523,2.829401...
0.819457,0.663489,2.838968,0.865873,0.501408,2.829834...
0.819543,0.663151,2.839160,0.866060,0.501274,2.830229...
```

Kinect One

```
Head_X,Head_Y,Head_Z,Neck_X,Neck_Y,Neck_Z,SpineShoulder_X,SpineShoulder_Y,SpineShoulder_Z
0.870586,0.663364,2.775473,0.900859,0.559911,2.762263...
0.871332,0.662706,2.775583,0.901121,0.559768,2.762396...
0.871649,0.661704,2.776140,0.901240,0.559678,2.762670...
```

Figura 60 – Ilustração da organização dos arquivos resultantes das capturas.

APÊNDICE B PROTOCOLO DE CAPTURA

Neste capítulo é descrito o protocolo de captura e uma ilustração do formulário utilizado para anotar as informações dos indivíduos.

Antes de realizar as coletas dos dados é necessário encontrar um local onde exista espaço suficiente para demarcar no chão o caminho que as pessoas devem percorrer. O caminho pode ser marcado com tinta, com papéis colados ao longo do trecho ou de outro modo, desde que fique claro para a pessoa onde ela deve andar.

É interessante que o local de captura fique perto do fluxo de pessoas, não o próximo demais a ponto das pessoas ficarem passando na frente do sensor, enquanto outra pessoa é capturada, e não longe demais de forma que fique difícil conseguir pessoas que contribuam com as capturas.

Neste trabalho a distância dos sensores para a pessoa capturada foi de 2.40 metros, pois nesta distância ambos os sensores conseguiam capturar todo o corpo dos indivíduos.

Em relação a iluminação, caso o ambiente seja iluminado por luz natural, o ideal é realizar as coletas sempre no mesmo horário, pois a iluminação do ambiente pode interferir na qualidade das capturas. Vale salientar que nos testes realizados, enquanto se procurava o local ideal, verificou-se que de modo geral ambientes fechados permitem uma melhor qualidade nas gravações.

É importante retirar qualquer objeto que esteja de fundo no ambiente de captura, como cadeira, ventilador, abajur, entre outros, pois notou-se que quando existem objetos ao fundo, em alguns momentos os sensores tem a tendência de confundir objetos com pessoas.

Após definido o local, é importante marcar no chão o caminho que as pessoas devem percorrer durante as capturas. Pode-se inicialmente colocar o disco giratório no centro do local de captura e a partir dele medir a distância do centro do disco para os pontos a serem demarcados no chão. Para que as pessoas tenham uma boa visualização do trajeto, é interessante marcar ao menos cinco pontos no chão, podendo marcar mais, como nove pontos por exemplo, conforme ilustra a Figura 61:



Figura 61 – Demarcação do trajeto no chão.

Após marcar os pontos no chão, pode-se ainda fazer um tracejado entre eles para ficar mais claramente delimitado o caminho.

Feita a marcação do trajeto, pode-se montar os equipamentos. Se for utilizado dois sensores nas capturas, é importante colocar um em cima do outro, para que ambos possam manter a mesma distância das pessoas. Os sensores devem estar em cima de um suporte para que eles fiquem na altura ideal para as capturas. Para uma distância de 2.40 metros entre os sensores e as pessoas, uma altura de 23,5 cm para o Kinect que ficar embaixo e 30,5 cm para o Kinect que ficar em cima permite uma boa qualidade nas gravações. É importante na hora de fixar uma altura que os sensores consigam capturar o corpo inteiro das pessoas.

Após montado todo o equipamento (suporte, Kinects, computador) e estando o software de captura já aberto no computador, pode-se convidar as pessoas para serem gravadas pelos sensores.

Toda vez que for convidar alguém para participar das capturas é importante explicar para a pessoa que ela deve percorrer o caminho delimitado no chão, mostrando onde começar e onde terminar a caminhada. Se for gravar as pessoas caminhando em ambas as direções, é necessário explicar que após o indivíduo percorrer o caminho ele deve retornar pelo mesmo trajeto. A melhor maneira de as pessoas entenderem o que devem fazer é se for dado um exemplo, por isso a pessoa que está realizando as capturas pode caminhar no tracejado para exemplificar o que deve ser feito. É pertinente salientar para a pessoa que embora ela deva seguir o caminho marcado no chão, é importante que ela caminhe o mais naturalmente possível.

Neste trabalho, cada pessoa foi instruída a caminhar partindo do ponto a esquerda do sensor até o ponto a direita e depois retornar pelo mesmo caminho. Elas foram orientadas a fazer isto 5 vezes para que para cada indivíduo fosse possível ter 5 exemplos em cada direção.

Após convidar a pessoa para ser coletada pelos sensores Kinect, é importante anotar os dados dela para que os resultados possam ser melhor estudados.

A tabela 49 ilustra o formulário utilizado para anotar as informações dos indivíduos coletados:

Tabela 49 – Ilustração do formulário utilizado para anotar as informações dos indivíduos coletados

Data:						
	Nome	Idade	Peso	Altura	Roupa	Curso
1						
2						
3						