

VIRGINIA ORTIZ ANDERSSON

**Identificação Biométrica com Antropometria e Caminhar Humano  
Utilizando o Kinect**

Dissertação de Mestrado apresentada  
ao Programa de Pós-Graduação em  
Computação da Universidade Federal de  
Pelotas, como requisito parcial para a  
obtenção do grau de Mestre em Ciência da  
Computação

Orientador: Prof. Dr. Ricardo Matsumura de Araújo

Pelotas, 2014

Universidade Federal de Pelotas / Sistema de Bibliotecas  
Catalogação na Publicação

A544i Andersson, Virginia Ortiz

Identificação biométrica com antropometria e caminhar humano utilizando o kinect / Virginia Ortiz Andersson ; Ricardo Matsumura de Araújo, orientador. — Pelotas, 2014.

102 f. : il.

Dissertação (Mestrado) — Programa de Pós-Graduação em Computação, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas, 2014.

1. Aprendizado de máquina. 2. Antropometria. 3. Caminhar humano. 4. K-nearest neighbor. 5. Support vector machines. I. Araújo, Ricardo Matsumura de, orient. II. Título.

CDD : 005

## **AGRADECIMENTOS**

Agradeço, primeiramente, ao Prof. Ricardo Matsumura de Araújo, meu orientador, pela oportunidade de desenvolver este interessante e desafiador trabalho durante o mestrado, pela paciência com minhas eternas dúvidas e principalmente por seu espírito motivador durante todo o processo. Agradeço também aos professores do PPGC, pela dedicação e comprometimento com o curso, que tornaram essa experiência extremamente gratificante e enriquecedora. Também, o trabalho do aluno da Engenharia da Computação desta mesma instituição, Rafael Dutra, pela disciplina, empenho e disposição no auxílio à captura de indivíduos para popular nosso banco de dados do caminhar humano.

Agradeço principalmente meu noivo Henrique Avila Vianna, responsável por minha motivação em ingressar e terminar o mestrado, com seu amor e apoio incondicional. Graças a sua disciplina em também terminar o mestrado você é uma grande fonte de inspiração. Muito obrigado, te amo!

Não posso também deixar de agradecer meus pais, por toda a educação e proteção que me deram durante toda a minha vida, valores que no mundo e na sociedade de hoje são preciosos, como ouro. É o que vocês são para mim.

Finalmente agradeço a UFPel, como instituição, pois foi nela que me graduei em Ciência da Computação, tornei-me servidora e aluna do mestrado no PPGC.

*I've missed more than 9000 shots in my career.  
I've lost almost 300 games.  
26 times, I've been trusted to take the game winning shot and missed.  
I've failed over and over and over again in my life.  
And that is why I succeed.*  
— MICHAEL JORDAN

## RESUMO

ANDERSSON, Virginia Ortiz. **Identificação Biométrica com Antropometria e Caminhar Humano Utilizando o Kinect**. 2014. 102 f. Dissertação (Mestrado) – Programa de Pós-Graduação em Computação. Universidade Federal de Pelotas, Pelotas.

O presente trabalho mostra o resultado da aplicação de algoritmos de aprendizado de máquina para classificar indivíduos através de atributos antropométricos e do caminhar humano, obtidos com o auxílio do sensor Microsoft Kinect como alternativa para obtenção de um modelo do corpo humano durante o caminhar. Para investigar o uso do Kinect nessa tarefa, é apresentada a metodologia aplicada na utilização das posições em 3D das principais articulações rastreadas pelo sensor, ou o “esqueleto”, de 140 indivíduos capturados para este trabalho caminhando em frente ao Kinect. A partir desses pontos foram obtidos atributos antropométricos e do caminhar humano. É explorada a classificação levando em conta o número de indivíduos, o número de quadros necessários para a classificação por antropometria, e o número de ciclos de caminhar capturados para uma acurácia satisfatória utilizando atributos do caminhar humano. Os resultados obtidos são comparados com o número de exemplos classificados corretamente fornecidos em trabalhos estado-da-arte, e mostram que a metodologia proposta apresenta resultados superiores, ou semelhantes, aos fornecidos nos trabalhos referenciados. Finalmente, alguns experimentos adicionais também são apresentados, como a classificação por gênero (feminino e masculino) e por índice de massa corporal.

**Palavras-chave:** Aprendizado de Máquina, Antropometria, Caminhar Humano, k-Nearest Neighbor, Support Vector Machines.

## ABSTRACT

ANDERSSON, Virginia Ortiz. **Biometric Identification with Anthropometry and Human Gait Using Kinect**. 2014. 102 f. Dissertação (Mestrado) – Programa de Pós-Graduação em Computação. Universidade Federal de Pelotas, Pelotas.

The present work shows the result of applying machine learning algorithms to classify individuals by anthropometric attributes and human gait obtained with the help of Microsoft Kinect sensor as an alternative to obtaining a model of the human body during walking. To investigate the use of Kinect in this task, the methodology applied in the use of 3D positions of the major joints tracked by the sensor, or the “skeleton” of 140 individuals captured for this work, walking in front of the Kinect is presented. From these 3D points, human gait and anthropometric parameters were obtained. It explored the classification taking into account the number of individuals, the number of frames required for classification by anthropometry, and the number of gait cycles captured towards a satisfactory accuracy using attributes of human gait. The results obtained are compared with the number of correctly classified instances provided in works state-of-art, and show that the proposed methodology provides superior or similar results to those provided in the referenced work. Finally, additional experiments are also presented, such as sorting by gender (male and female) and body mass index.

**Keywords:** Biometrics, Gait, Anthropometrics, Machine Learning, k-Nearest Neighbor, Support Vector Machines.

## LISTA DE FIGURAS

Figura 1	<i>Pipeline</i> da análise do caminhar humano. . . . .	18
Figura 2	Ilustração das medições realizadas no sistema biométrico de Bertillon (DCJS-NY). . . . .	20
Figura 3	O primeiro cartão de identificação preenchido no <i>New York State Bertillon Bureau</i> (NYDCJS, 1997). . . . .	20
Figura 4	O <i>New York State Bertillon Bureau</i> , em 1902, com funcionárias trabalhando na busca de cartões de medições antropométricas categorizadas (NYDCJS, 1997). . . . .	22
Figura 5	Analogia de movimento de uma roda demonstrando a natureza cíclica de um movimento à frente. Adaptado de Vaughan, Davis and O'connor (1992). . . . .	22
Figura 6	Ciclo de caminhar, adaptado de Cunado, Nixon e Carter (2003) e Vaughan, Davis and O'connor (1992). . . . .	23
Figura 7	Fases do ciclo de caminhar, adaptado de Yam, Nixon e Carter (2004) e Cunado, Nixon e Carter (2003). . . . .	24
Figura 8	Planos para extrair componentes cinemáticos usados por Murray, Drought e Kory (1964) (VAUGHAN; DAVIS; O'CONNOR, 1992). . . . .	25
Figura 9	Medida do passo e da passada a partir de um ponto do centro ao calcanhar do pé (MURRAY; DROUGHT; KORY, 1964). . . . .	26
Figura 10	Ângulos em graus descritos pelo quadril durante o ciclo de caminhar. As diferentes curvas são referentes à variação de idade dos indivíduos (MURRAY; DROUGHT; KORY, 1964). . . . .	28
Figura 11	Ângulos, em graus, descritos pelo joelho dos indivíduos durante o caminhar (MURRAY; DROUGHT; KORY, 1964). . . . .	29
Figura 12	Ângulos em graus do tornozelo descritos durante o ciclo de caminhar (MURRAY; DROUGHT; KORY, 1964). . . . .	29
Figura 13	Quadros de vídeo do banco de dados do caminhar SOTON (SHUTLER et al., 2004). . . . .	31
Figura 14	CASIA <i>dataset</i> A apresentado quadros de sequências de vídeos de indivíduos caminhando ao ar livre com diversos ângulos de câmera (ZHENG et al., 2011). . . . .	31
Figura 15	Quadros capturados de sequências de vídeo presentes no CASIA <i>dataset</i> B, em ambiente controlado incluindo indivíduos utilizando mochilas e outros acessórios (ZHENG et al., 2011). . . . .	32
Figura 16	CASIA <i>dataset</i> C, contendo quadros capturados de sequências de vídeo em infravermelho a noite (ZHENG et al., 2011). . . . .	32

Figura 17	Quadros de <i>streams</i> de profundidade capturadas com o Kinect e armazenadas no banco TUM-GAID (HOFMANN; BACHMANN; RIGOLL, 2012). . . . .	33
Figura 18	Modelo de pêndulo para obtenção dos ângulos (CUNADO; NIXON; CARTER, 2003). . . . .	34
Figura 19	Os pêndulos superior e inferior modelam o fêmur e a perna inferior respectivamente, conectados pela articulação do joelho (k) (YAM; NIXON; CARTER, 2004). . . . .	35
Figura 20	(a) esqueleto inicial estimado, (b) remoção de ruído e (c) figura final (YOO; NIXON, 2011). . . . .	35
Figura 21	Silhuetas binárias normalizadas do ciclo de caminhar e seus respectivos GEIs (HAN; BHANU, 2006). . . . .	38
Figura 22	(a) Reconstrução do modelo de voxel, (b) GEV e (c) secção do GEV mostrando as diferentes densidades internas (SIVAPALAN et al., 2011). . . . .	38
Figura 23	Exemplo de imagem capturada do banco de dados criado: (a) imagem de profundidade não preparada, (b) silhueta 3D, (c) silhueta binária (SIVAPALAN et al., 2011). . . . .	39
Figura 24	Silhueta com caixa delimitadora: O segmento em que liga os vértices da direita representa a altura estimada (GUILLEN et al., 2009). . . . .	40
Figura 25	Divisão do caminhar em subpartes (VENKAT; WILDE, 2010). . . . .	41
Figura 26	Divisão da silhueta em segmentos (YEOH et al., 2011). . . . .	42
Figura 27	(a) Ângulos da forquilha e joelho e atributos coletados da silhueta (YEOH et al., 2011). . . . .	43
Figura 28	Representação visual dos DGHEI utilizando diferentes configurações (HOFMANN; BACHMANN; RIGOLL, 2012). . . . .	44
Figura 29	Representação visual de 19 vetores rotulados com classe A e B e uma classe desconhecida. O vetor com classe desconhecida é portanto, rotulado com seus 4 ( $k=4$ ) vizinhos mais próximos de acordo com a distância euclidiana entre eles (PETERSON, 2009). . . . .	46
Figura 30	Representação de uma <i>Support Vector Machine</i> de classificação binária (SUDHA; BHAVANI, 2012). . . . .	49
Figura 31	Microsoft Kinect (MICROSOFT, 2012). . . . .	50
Figura 32	Profundidade no Kinect (MICROSOFT, 2011). . . . .	51
Figura 33	Arquitetura simplificada do PrimeSense PS1080 SoC, principal componente do Microsoft Kinect (PRIMESENSE, 2012). . . . .	52
Figura 34	Pontos do esqueleto relativos as articulações do corpo humano (MICROSOFT, 2011). . . . .	53
Figura 35	Figuras de esqueleto construídas a partir dos pontos retornados pela NUI <i>Skeleton</i> API (MICROSOFT, 2011). . . . .	53
Figura 36	Espaço tridimensional do esqueleto (MICROSOFT, 2011). . . . .	54
Figura 37	Pipeline da análise do caminhar humano com o uso do Kinect. . . . .	56
Figura 38	Arquitetura da metodologia proposta. . . . .	56
Figura 39	Distribuição, em percentagem, do gênero dos indivíduos . . . . .	57
Figura 40	Distribuição de alturas por indivíduo . . . . .	58
Figura 41	Distribuição de peso por indivíduo . . . . .	58

Figura 42	Distribuição de idade por indivíduo . . . . .	58
Figura 43	Distribuição de IMC por indivíduo . . . . .	59
Figura 44	Microsoft Kinect com um disco giratório (prato rotativo) adicionando ao sensor o movimento panorâmico não presente no sensor. . . . .	60
Figura 45	Trajétoria semi-circular mostrando os sentidos horário e anti-horário percorridos pelo indivíduo. . . . .	60
Figura 46	<i>Setup</i> para captura do caminhar apresentando o sensor Kinect e marcas no chão para guiar as pessoas. . . . .	61
Figura 47	<i>Snapshot</i> da base de dados: Todos os indivíduos capturados estão em um arquivo contendo gênero, altura, peso e cada diretório do indivíduo contém arquivos de caminhadas, contendo pontos 3D de todas as articulações humanas rastreadas. Os quadros são arranjados sequencialmente com um rótulo e coordenadas x, y e z em cada linha. . . . .	62
Figura 48	Aplicação do filtro ARMA nos ângulos obtidos do joelho direito de um indivíduo durante 1 caminhada capturada, realizando a trajetória semi-circular nos sentidos de ida e de volta. . . . .	64
Figura 49	Medidas dos segmentos do corpo calculadas a partir das articulações, ilustradas como círculos, fornecidas pela API do Kinect <i>NUI Skeleton</i> . . . . .	64
Figura 50	Diferença entre alturas fornecidas durante o experimento e alturas calculadas pelo Kinect. . . . .	65
Figura 51	Em (a) é mostrado os segmentos inferiores do corpo utilizados para extrair os parâmetros cinemáticos, (b) os ângulos $\omega$ , $\gamma$ e $\alpha$ formados entre os segmentos segundo o modelo de pêndulo de (CUNADO; NIXON; CARTER, 2003) e (c) o ângulo $\beta$ de abertura do pé. . . . .	66
Figura 52	Fases de extensão (vales) e flexão (picos) em uma sequência de valores de ângulos do joelho direito descritos durante uma caminhada capturada. . . . .	67
Figura 53	Diferença entre as posições dos calcanhares direito e esquerdo: A média aritmética entre os valores mais altos (picos) resulta no tamanho do passo médio durante toda a caminhada desenvolvida. . . . .	67
Figura 54	Resultado dos ICCs por algoritmos de aprendizado de máquina utilizando todos os atributos e o maior número de indivíduos do conjunto. . . . .	82
Figura 55	Resultado dos ICCs utilizando k-NN (k=4) por <i>datasets</i> com diferentes atributos e o método proposto (todos os atributos). . . . .	82
Figura 56	Mediana dos ICCs de cada grupo de indivíduos. . . . .	84
Figura 57	Mediana dos ICCs de cada grupo de indivíduos para <i>datasets</i> contendo atributos antropométricos. . . . .	84
Figura 58	Histogramas com número de exemplos de caminhar distribuídos por número de vales altos e por ciclos. . . . .	85
Figura 59	Histograma com número de exemplos de caminhar distribuídos por número de quadros capturados. . . . .	86
Figura 60	Resultados das instâncias classificadas corretamente (ICCs) utilizando atributos de antropometria por grupos de números de quadros capturados. . . . .	86

Figura 61	Número de instâncias classificadas corretamente (ICCs) para diferentes números de ciclos. . . . .	87
Figura 62	Resultados obtidos por Yoo e Nixon (2011) e resultados obtidos pelo método proposto: ICCs apresenta o número de instâncias classificadas corretamente para o mesmo número de indivíduos. . . . .	89
Figura 63	Número de ICCs em pontos percentuais para a classificação utilizando o método proposto, com 140 indivíduos, e o método DGHEIs, com 176 indivíduos. . . . .	90
Figura 64	Aplicação de uma função logarítmica próxima a curva de ICCs obtida para todos os atributos do método proposto possibilita acompanhar um possível resultado com até 176 indivíduos. . . . .	90
Figura 65	Percentual de ICCs para a classificação de 10 indivíduos, comparando o método proposto com o resultado obtido em Munsell et al.(2012) e Preis et al. (2012). . . . .	91

## LISTA DE TABELAS

Tabela 1	Grupos de Atributos . . . . .	69
Tabela 2	Composição dos <i>Datasets</i> . . . . .	69
Tabela 3	Implementações dos algoritmos de Aprendizado de Máquina no WEKA . . . . .	70
Tabela 4	Comparação dos resultados obtidos utilizando a Caixa delimitadora	79
Tabela 5	Comparação dos resultados obtidos utilizando a Caixa delimitadora	79
Tabela 6	Comparação de 30 Indivíduos . . . . .	80
Tabela 7	Comparação de 60 Indivíduos . . . . .	80
Tabela 8	Comparação de 100 Indivíduos . . . . .	80
Tabela 9	Configuração dos algoritmos e melhores resultados obtidos . . . . .	81
Tabela 10	Atributos selecionados através da seleção de subconjunto baseado em correlação e o quanto remover cada um destes atributos afeta a taxa de ICCs, em pontos percentuais. . . . .	83
Tabela 11	Comparação entre as metodologias . . . . .	88
Tabela 12	Comparação entre as metodologias . . . . .	90
Tabela 13	Comparação entre as metodologias . . . . .	91
Tabela 14	Classificação por Gênero . . . . .	92
Tabela 15	Matriz de confusão g1-g2 para todos os atributos . . . . .	92
Tabela 16	Matriz de confusão g2-g1 para todos os atributos . . . . .	92
Tabela 17	Classificação por Índice de Massa Corporal . . . . .	93
Tabela 18	Matriz de confusão imc1-imc2 para atributos espaço-temporais . . . . .	93
Tabela 19	Matriz de confusão imc2-imc1 para atributos espaço-temporais . . . . .	94

## LISTA DE ABREVIATURAS E SIGLAS

3D	Três Dimensões
API	<i>Application Programming Interface</i>
ARMA	<i>Autoregressive Central Moving Average</i>
ASCII	<i>American Standard Code for Information Interchange</i>
CASIA	<i>The Chinese Academy of Science</i>
CPU	<i>Central Processing Unit</i>
DGHEI	<i>Depth Gradient Histogram Energy Image</i>
EWMA	<i>Exponential Weighted Moving Average</i>
Ex	<i>Extension</i>
FFT	<i>Fast Fourier Transform</i>
FI	<i>Flexion</i>
FP	<i>Falso Positivo</i>
GAID	<i>Gait from Audio, Image and Depth</i>
GEI	<i>Gait Energy Images</i>
GEV	<i>Gait Energy Volumes</i>
HOG	<i>Histograms of Oriented Gradients</i>
IBk	<i>Instance-Based k-Nearest Neighbor</i>
ICC	Instâncias Classificadas Corretamente
IMC	Índice de Massa Corporal
K-NN	<i>k-Nearest Neighbor</i>
LibSVM	<i>Library for Support Vector Machines</i>
MDA	<i>Multiple Discriminat Analysis</i>
MLP	<i>Multilayer Perceptron</i>
NUI	<i>Natural User Interface</i>
OAA	<i>One Against All</i>
OAO	<i>One Against One</i>

PS1080 *PrimeSense Chip 1080*  
RGB *Red, Green, Blue*  
SMO *Sequential Minimal Optimization*  
SoC *System on a Chip*  
SOTON *Southampton Database*  
SVM *Support Vector Machines*  
TUM *Technische Universitat Munchen*  
USB *Universal Serial Bus*  
VN *Verdadeiro Negativo*  
WEKA *Waikato Environment for Knowledge Analysis*  
XBOX *Microsoft X-Box 360 VideoGame*

# SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	15
<b>2</b>	<b>REVISÃO TEÓRICA E ESTADO DA ARTE</b>	19
<b>2.1</b>	<b>Conceitos de Antropometria e a Técnica de <i>Bertillonage</i></b>	19
<b>2.2</b>	<b>Teoria e Conceitos do Caminhar Humano</b>	21
2.2.1	Características cíclicas do caminhar humano	22
2.2.2	Descrição do ciclo de caminhar	23
2.2.3	Componentes do ciclo de caminhar	26
2.2.4	Aquisição do caminhar humano	30
2.2.5	Bancos de dados do caminhar humano	30
<b>2.3</b>	<b>Análise do Caminhar Humano</b>	33
2.3.1	Análise baseada em modelo	33
2.3.2	Modelo de Pêndulo	33
2.3.3	Figuras de movimento	35
2.3.4	Análise livre de modelo	36
2.3.5	GEIs e GEVs	37
2.3.6	Caixa delimitadora	39
2.3.7	Divisão do caminhar humano em subpartes	41
2.3.8	Divisão do corpo em segmentos	42
2.3.9	DGHEIs	43
<b>2.4</b>	<b>Reconhecendo e classificando o padrão de caminhar</b>	45
2.4.1	<i>k</i> -Nearest Neighbor	45
2.4.2	Aplicação do <i>k</i> -Nearest Neighbor na classificação de indivíduos através do caminhar	46
2.4.3	Redes neurais artificiais	47
2.4.4	Redes neurais na classificação de indivíduos através do caminhar	48
2.4.5	SVM - <i>Support Vector Machines</i>	48
2.4.6	SVM multiclasse	49
<b>2.5</b>	<b>O Sensor Microsoft Kinect</b>	50
2.5.1	Hardware e funcionamento do Kinect	50
2.5.2	PrimeSense PS1080 System on a Chip	51
2.5.3	<i>Natural User Interface Application Programmable Interface</i> - NUI API	51
2.5.4	<i>NUI Skeleton Tracking</i> - API para rastreamento do esqueleto	52
2.5.5	Espaço tridimensional do esqueleto	54

<b>3</b>	<b>OBJETIVOS E METODOLOGIA</b>	55
<b>3.1</b>	<b>Objetivos</b>	55
<b>3.2</b>	<b>Metodologia</b>	55
3.2.1	Uso do sensor Kinect na análise do caminhar humano	55
3.2.2	Aquisição de Dados de Indivíduos	57
3.2.3	Ambiente de Captura e Trajetória	59
3.2.4	Base de Dados do Caminhar	61
3.2.5	Filtragem de Indivíduos	61
3.2.6	Suavização das Articulações	62
<b>3.3</b>	<b>Análise Antropométrica</b>	63
<b>3.4</b>	<b>Análise de Parâmetros Espaço-temporal e Cinemáticos do Caminhar Humano</b>	65
<b>3.5</b>	<b>Composição dos <i>Datasets</i></b>	68
<b>3.6</b>	<b>Escolha do Algoritmo de Aprendizado de Máquina</b>	70
3.6.1	Execução e Teste	72
<b>3.7</b>	<b>Comparação com Trabalhos Relacionados</b>	73
3.7.1	Reprodução de metodologias utilizando a base de dados obtida com o Kinect	74
3.7.2	Comparação da metodologia proposta com trabalhos relacionados que utilizam o Kinect na obtenção da caminhada	77
<b>4</b>	<b>RESULTADOS</b>	78
<b>4.1</b>	<b>Utilização da Base de Dados do Esqueleto com outras Metodologias</b>	78
4.1.1	Reprodução da técnica de caixa delimitadora	78
4.1.2	Reprodução da técnica de figuras de movimento	80
<b>4.2</b>	<b>Ajuste dos Algoritmos de Aprendizado de Máquina</b>	81
<b>4.3</b>	<b>Resultados Utilizando os Atributos do Método Proposto</b>	82
4.3.1	Efeito da Divisão por Número de Indivíduos	84
<b>4.4</b>	<b>Exploração do Número de Quadros e Ciclos de Caminhar</b>	85
<b>4.5</b>	<b>Comparação com Resultados de Métodos Estado-da-Arte</b>	88
4.5.1	Figuras de movimento de Yoo e Nixon	88
4.5.2	DGHEIs	89
4.5.3	Outras metodologias utilizando o Sensor Kinect	90
<b>4.6</b>	<b>Resultados Acessórios</b>	91
4.6.1	Classificação por gênero	91
4.6.2	Classificação por índice de massa corporal (IMC)	92
<b>4.7</b>	<b>Publicações</b>	94
<b>5</b>	<b>CONCLUSÃO</b>	95
	<b>REFERÊNCIAS</b>	99

# 1 INTRODUÇÃO

A capacidade de identificar pessoas é um recurso inerente aos seres humanos. As pessoas utilizam, desde que nascem, características físicas e comportamentais presentes na fisiologia do corpo humano para reconhecer uns aos outros. Elas identificam parentes, amigos e colegas ao longo da vida aprendendo padrões no tom de voz, no jeito de caminhar, nas características da face e proporções corporais dessas pessoas. Realizam inconscientemente a biometria, que é o uso de características biológicas físicas e comportamentais para a distinção e identificação de indivíduos.

Existem evidências de que a identificação de indivíduos por digitais era um método comum na China, no século III, durante a dinastia Qin (SUTROP; LAAS-MIKKO, 2012). Na França, no século 19, Alphonse Bertillon, do departamento de identificação de criminosos da polícia de Paris, desenvolveu e usou um método de identificação de transgressores através das medidas dos segmentos do corpo, técnica conhecida como *Bertillonage*. Logo esse método foi substituído pela análise e armazenamento das impressões digitais, que poderiam deixar rastros nas cenas de crime, facilitando a investigação (JAIN; ROSS; PRABHAKAR, 2004). A biometria evoluiu, então, para a utilização em um grande número de aplicações civis e sua potencial aplicação deu-se na segurança e controle de acesso a locais e sistemas computacionais (CUNADO; NIXON; CARTER, 2003). Atualmente a biometria é usada também para controle de fronteiras regionais, sistemas de saúde, bancos e instituições financeiras, gerenciamento de eleições, seguros e comércio. Em países que se preocupam com segurança internacional, a biometria auxilia no monitoramento e gerenciamento de ambientes públicos ou de grande importância, na prevenção de possíveis ataques (SUTROP; LAAS-MIKKO, 2012).

As características físicas e comportamentais usadas para biometria devem qualificar-se de acordo com os seguintes critérios (i) universalidade: todas as pessoas devem possuir essa característica, (ii) peculiaridade: duas pessoas quaisquer devem possuir essa característica suficientemente diferente e (iii) mensurável: essa característica pode ser medida de alguma forma (JAIN; ROSS; PRABHAKAR, 2004). Um sistema biométrico envolve biometria para reconhecimento de indivíduos. Para que

essas características sejam eficientemente usadas por um sistema biométrico devem ser consideradas também: (i) performance: que refere-se à acurácia e velocidade de aquisição da característica, bem como os recursos para obtê-la e fatores ambientais que a afetam, (ii) aceitabilidade: indica o quanto as pessoas estão dispostas a aceitar o uso de um sistema com uma determinada biometria no seu dia a dia e (iii) engano: mostra o quão fácil é lograr o sistema utilizando métodos fraudulentos.

Sistemas biométricos tem se popularizado devido, principalmente, à vigilância e segurança. Outras funcionalidades comuns são autenticação de sistemas computacionais e produção de conteúdo personalizado, como *e-commerce*, serviços de entretenimento e customização de serviços por perfil de usuário (WANG, 2012). De acordo com a maneira que as características biológicas são coletadas, os sistemas biométricos podem ser classificados em: (i) sistemas ativos, que necessitam da interação do usuário com a interface ou meio coletor, por exemplo, impressões digitais e íris e (ii) sistemas passivos, que não requerem contato com o indivíduo para extração da informação biométrica ou conhecimento do mesmo de que a coleta está sendo feita, como o reconhecimento por voz e face.

Sistemas biométricos de autenticação comumente usados utilizam face, íris, digitais e amostras de voz, enquanto que a vigilância é feita principalmente por reconhecimento de face (SINGH; SINGH, 2013). Entretanto, na prática, esse sistema biométrico tem se mostrado insuficiente, a exemplo do atentado na maratona de Boston em 2013 (GALLAGHER, 2013) ou controverso, como no caso das câmeras de vigilância de Chicago (SCHWARTZ, 2013), levando a alternativas como a campanha ao uso de máscaras em espaços públicos para evitar e protestar contra a vigilância em excesso na mesma cidade (KATZ, 2014). Como a face, íris, voz e impressões digitais de uma pessoa não são secretas, elas podem ser copiadas e utilizadas para burlar sistemas de autenticação, segurança e vigilância, fazendo com que uma pessoa se passe por outro indivíduo, ou atribua suas ações a esse indivíduo (SINGH; SINGH, 2013).

Alternativas ao reconhecimento de face e outras biometrias fáceis de simulação ou fraude são a antropometria e o caminhar humano (*gait*). A antropometria é o conjunto de medidas do corpo humano que, quando combinadas, fornecem uma assinatura do indivíduo, que permite distingui-lo dentre as demais pessoas. Assim como a antropometria, o caminhar humano também pode ser caracterizado com diversos componentes que, juntos, fornecem uma assinatura do indivíduo. A palavra em inglês *gait* significa marcha, andadura, e pode ser aplicada não somente ao caminhar, mas também a ação de correr (NANDINI et al., 2011). O caminhar humano é uma característica biológica física e comportamental e pode sofrer alterações durante um longo período de tempo. O indivíduo pode perder peso, sofrer lesões nas articulações ou cérebro, alterando a maneira que ele caminha (JAIN; ROSS; PRABHAKAR, 2004). Da mesma forma, condições como gravidez, uso de sapatos diferenciados, roupas ou

embriaguez também podem afetar o padrão do caminhar humano (CUNADO; NIXON; CARTER, 2003).

Segundo Cunado, Nixon e Carter (2003) e Murray, Drought e Kory (1964), existem em torno de 20 componentes distintos no caminhar humano. Esses componentes podem ser, segundo Yoo e Nixon (2011), divididos em parâmetros espaço-temporais, que envolvem: medidas de duração, velocidades e distâncias e parâmetros cinemáticos; rotações desenvolvidas pelas articulações do corpo humano durante o ciclo de caminhar ou ciclo total.

Rotações e deslocamentos de membros e articulações, como pélvis e tórax, possuem variações significativas de um indivíduo para outro, mas não apresentam um padrão consistente para um mesmo indivíduo, em caminhadas distintas (MURRAY; DROUGHT; KORY, 1964) (CUNADO; NIXON; CARTER, 1997). Já a rotação das canelas e o deslocamento espacial do tronco possuem características individuais e formam um padrão se forem observados em diferentes experimentos para um mesmo indivíduo. O trabalho de Cunado, Nixon e Carter (2003) se restringiu à análise dos ângulos descritos pelos quadris e joelhos durante o ciclo de caminhada, que apresentaram o padrão mais distintivo em relação a outros componentes cinemáticos do caminhar.

As abordagens envolvidas na extração e análise do caminhar humano podem ser classificadas como (i) baseada em modelo: onde o caminhar é descrito pela teoria, e é reconstruído através de um modelo, que é encaixado à pessoa em cada quadro da sequência de caminhada. Os parâmetros espaço-temporais e cinemáticos, então, são extraídos desse modelo durante a caminhada; e (ii) livre de modelo, onde são usadas características baseadas no comportamento do movimento, forma e silhueta do indivíduo, sem considerar as estruturas teóricas presentes no caminhar (YOO; NIXON, 2011) (NG et al., 2011).

O caminhar humano como biometria não exige proximidade e contato com o indivíduo analisado, mas para a captura dos movimentos humanos, sistemas biométricos podem usar marcadores, o que não é ideal, visto que a biometria através do caminhar é motivada por ser um sistema passivo. Como alternativa, sistemas com esse tipo de biometria utilizam extensiva análise de vídeos e imagens utilizando técnicas de visão computacional.

Apesar da complexidade envolvida na análise do movimento humano, a tecnologia atual de sensores, que capturam e pré-processam imagens, pode simplificar etapas de processamento na análise de vídeos. Uma dessas tecnologias é o Microsoft Kinect (MICROSOFT, 2011), um periférico composto por uma câmera RGB (*Red, Green, Blue*) e um sensor de profundidade, cuja combinação pode capturar os movimentos de corpo inteiro em 3D, reconhecer gestos, entre outras funções Chang2011, além de fornecer um modelo de esqueleto composto por pontos que representam as articulações

humanas.

A Figura 1 mostra o processo padrão para a extração de características do caminhar humano. As etapas apresentadas neste *pipeline* são utilizadas em ambas as abordagens: baseada em modelo e livre de modelo.



Figura 1: *Pipeline* da análise do caminhar humano.

O uso do Kinect simplifica as etapas de captura de vídeo através de uma câmera digital, executa a extração da silhueta da pessoa e cria um modelo de esqueleto correspondente ao indivíduo, retornando pontos que representam a posição das articulações durante a captura dos movimentos, em tempo real. Isso favorece a concentração de esforços computacionais no aprendizado e interpretação dos valores pertencentes ao caminhar humano. Fundamentado nessa afirmação, este trabalho apresenta aplicações de algoritmos de aprendizado de máquina em diversos *datasets* constituídos de atributos antropométricos e do caminhar humano, extraídos de figuras de movimentos obtidas com o sensor Kinect, durante o caminhar de diferentes indivíduos em frente ao sensor. A finalidade desses experimentos é investigar a viabilidade do uso do mesmo sensor em sistemas biométricos que utilizam dados antropométricos e do caminhar humano.

O presente trabalho está dividido em cinco Capítulos. No segundo Capítulo é realizada uma revisão teórica, introduzindo conceitos de antropometria e caminhar humano necessários para o acompanhamento da metodologia e objetivo do trabalho. Além disso são apresentadas algumas técnicas de análise do caminhar humano pertencentes ao estado da arte. O terceiro Capítulo apresenta o objetivo do presente trabalho e a metodologia utilizada para obter resultados deste objetivo, apresentados no quarto Capítulo. Por fim, o quinto Capítulo apresenta as conclusões obtidas neste trabalho.

## 2 REVISÃO TEÓRICA E ESTADO DA ARTE

As Seções seguintes presentes neste Capítulo apresentam uma introdução à Antropometria e Caminhar Humano como formas de biometria, expondo os principais conceitos e abordagens para a modelagem do caminhar humano que permitem extrair as características que compõem o padrão único do caminhar de cada pessoa. São mostrados também conceitos de algoritmos de aprendizado de máquina e sua utilização no reconhecimento biométrico através do caminhar humano e, finalmente, conceitos do sensor Microsoft Kinect úteis para o entendimento deste trabalho.

### 2.1 Conceitos de Antropometria e a Técnica de *Bertillonage*

Segundo Angelo Albrizio em “*Biometry and Anthropometry : from Galton to Constitutional Medicine*” (2007) o emprego do termo “antropometria” foi primeiramente utilizado pelo naturalista Johann Sigismund Elsholtz (1623-1688) em seu manual “*De mutua membrorum proportione*” publicado em Pádua, na Itália. Em seu método, Elsholtz propôs o uso de instrumentos para determinar as medidas externas do corpo humano. Na segunda metade do século 18, a antropometria era utilizada por naturalistas e antropólogos para investigar o homem e suas características morfológicas.

É atribuída a Francis Galton (1822-1911) grande parte dos trabalhos relacionados a Antropometria, datados do final do século 19. A ideia de antropometria de Galton consistia na “arte de medir as faculdades físicas e mentais dos seres humanos”. Estava interessado não somente em medir segmentos lineares do corpo, mas também em medir aptidões, qualidades e características mentais do ser humano (ALBRIZIO, 2007).

A definição de antropometria que se baseia este presente trabalho é “o conjunto de medidas e proporções de todo o corpo humano”, dada por Paul Topinard (1830-1911) e utilizada como sistema biométrico por Alphonse Bertillon (1853-1914), criador da “*Bertillonage*”, que consistia em utilizar diversos segmentos do corpo para identificar indivíduos, técnica que tornou-se popular em meados do século 19 (ALBRIZIO, 2007).

Bertillon agrupou medidas de segmentos do corpo humano que, informalmente,

possuem ossos, entre eles a largura do crânio, comprimento do pé, cômado (comprimento do ante-braço até o dedo médio), tronco e dedo médio esquerdo. Essas medidas, juntamente com cor dos olhos, cor do cabelo e uma *mug shot*, (fotos de frente e de perfil), eram gravados em cartões medindo 16,5cm de altura por 13,9cm de largura (NYDCJS, 1997). A Figura 2 mostra a maneira como as medições eram executadas em um indivíduo, com contato direto com o mesmo, e manualmente. A Figura 3 exibe o primeiro cartão de *Bertillonage* preenchido no *New York State Bertillon Bureau*.



Figura 2: Ilustração das medições realizadas no sistema biométrico de Bertillon (DCJS-NY).

				(M) (L)																																													
Height	75"	Head lgn.	18.5"	L Foot	26.5"	Class	100-10																																										
Stretch	75"	Head width	16"	L Mid F	16.5"	Arrest	6																																										
Trunk	50.5"	Chest width	34.5"	L Lof F	5.5"	Apprentice	Apr 1910																																										
Forearm	15.5"	R Forearm	15.5"	L Forearm	16.5"	By	W. J. Sullivan																																										
Forearm	15.5"	L Forearm	16.5"	L Forearm	16.5"	Occupation	Scholar																																										
Remarks or notes				217 182 18 100-10																																													
<table border="1"> <tr> <td>Age</td> <td>24</td> <td>Build</td> <td>Slender</td> <td>Complexion</td> <td>Fair</td> </tr> <tr> <td>Color of eyes</td> <td>Blue</td> <td>Color of hair</td> <td>Black</td> <td>Color of skin</td> <td>Fair</td> </tr> <tr> <td>Color of nose</td> <td>Black</td> <td>Color of lips</td> <td>Black</td> <td>Color of teeth</td> <td>Black</td> </tr> <tr> <td>Color of hands</td> <td>Black</td> <td>Color of feet</td> <td>Black</td> <td>Color of fingers</td> <td>Black</td> </tr> <tr> <td>Color of toes</td> <td>Black</td> <td>Color of nails</td> <td>Black</td> <td>Color of hair on head</td> <td>Black</td> </tr> <tr> <td>Color of hair on face</td> <td>Black</td> <td>Color of hair on neck</td> <td>Black</td> <td>Color of hair on chest</td> <td>Black</td> </tr> <tr> <td>Color of hair on arms</td> <td>Black</td> <td>Color of hair on legs</td> <td>Black</td> <td>Color of hair on feet</td> <td>Black</td> </tr> </table>								Age	24	Build	Slender	Complexion	Fair	Color of eyes	Blue	Color of hair	Black	Color of skin	Fair	Color of nose	Black	Color of lips	Black	Color of teeth	Black	Color of hands	Black	Color of feet	Black	Color of fingers	Black	Color of toes	Black	Color of nails	Black	Color of hair on head	Black	Color of hair on face	Black	Color of hair on neck	Black	Color of hair on chest	Black	Color of hair on arms	Black	Color of hair on legs	Black	Color of hair on feet	Black
Age	24	Build	Slender	Complexion	Fair																																												
Color of eyes	Blue	Color of hair	Black	Color of skin	Fair																																												
Color of nose	Black	Color of lips	Black	Color of teeth	Black																																												
Color of hands	Black	Color of feet	Black	Color of fingers	Black																																												
Color of toes	Black	Color of nails	Black	Color of hair on head	Black																																												
Color of hair on face	Black	Color of hair on neck	Black	Color of hair on chest	Black																																												
Color of hair on arms	Black	Color of hair on legs	Black	Color of hair on feet	Black																																												
STATE OF NEW YORK, Office of Superintendent of State Prisons, BUREAU OF IDENTIFICATION, Albany, N.Y.																																																	

Figura 3: O primeiro cartão de identificação preenchido no *New York State Bertillon Bureau* (NYDCJS, 1997).

Para classificar um indivíduo, o método de Bertillon consistia em posicionar o in-

divíduo em uma árvore começando com o tamanho do crânio: (i) curto, (ii) médio, ou (iii) largo. Cada uma dessas três divisões é adicionalmente dividida na amplitude do mesmo crânio: (iv) estreito, (v) médio, (vi) amplo. Depois dessa classificação, os dedos médios são classificados em (vii) curtos, (viii) medios, (viii) compridos. Seguindo na árvore, o próximo a ser classificado em curto, médio e largo é o pé esquerdo. O próximo item é o antebraço esquerdo, seguido do dedo mínimo, orelha direita e finalizando com a altura. Ainda é possível classificar nas ramificações restantes, a cor dos olhos e o cabelo do indivíduo (FRAZER, 1875).

Dividindo e agrupando cada uma das medidas dos segmentos do corpo em pequena, média ou grande, Bertillon podia classificar as medidas de qualquer pessoa em uma das 243 categorias distintas. Quando um indivíduo era preso, o mesmo era descrito, medido e fotografado, tendo seus cartões preenchidos, indexados e posicionados nas categorias apropriadas, onde podiam ser comparados com outros cartões nas mesmas categorias (NYDCJS, 1997).

O método de Bertillon foi oficialmente implantado na Polícia de Paris em 1882 e rapidamente se espalhou por toda a França, Europa e no restante do mundo. Em 1887 o método foi implantado nos Estados Unidos. A padronização do sistema biométrico de Bertillon mostrou pela primeira vez que qualquer indivíduo, uma vez que classificado corretamente no sistema, poderia ser posteriormente identificado com acurácia (NYDCJS, 1997).

Problemas com a *Bertillonage* começam a aparecer quando o número de registros armazenado atinge a marca de 50.000, a exemplo do *New York State Bertillon Bureau*, ilustrado na Figura 4. O tempo requerido para a tarefa de pesquisar medidas nas 243 categorias básicas de antropometria, que continham 200 ou mais cartões, aumenta para horas. Outro problema é a não acurácia das medidas dos segmentos do corpo, causada pelo erro ou desconhecimento técnico da pessoa responsável pelas medições, bem como medições feitas em adolescentes infratores que acabam por crescer, tornando os registros incompatíveis. Essas e outras causas contribuíram para a substituição do sistema biométrico de Bertillon, gradativamente, pelo uso de impressões digitais (NYDCJS, 1997) (JAIN; ROSS; PRABHAKAR, 2004).

## 2.2 Teoria e Conceitos do Caminhar Humano

Segundo o livro *Dynamics of Human Gait*, de Vaughan, Davis and O'connor (1992), o ato de caminhar possui dois pré-requisitos básicos: (i) movimentos periódicos de cada pé de uma posição de sustentação até a outra e (ii) Forças de reação do solo, aplicadas ao pé, para sustentar o corpo. Segundo os autores, esses dois pré-requisitos são suficientes para que qualquer forma de caminhar bípede aconteça.



Figura 4: O *New York State Bertillon Bureau*, em 1902, com funcionárias trabalhando na busca de cartões de medições antropométricas categorizadas (NYDCJS, 1997).

### 2.2.1 Características cíclicas do caminhar humano

O caminhar humano possui características periódicas. Possui, inclusive, um espectro de frequência associado a ele que atinge no máximo, em um caminhar normal, a frequência de 6Hz (CUNADO; NIXON; CARTER, 2003). O movimento periódico das pernas é a essência da natureza cíclica do caminhar humano. A ideia de um padrão cíclico que se repete passo a passo a medida que se caminha remete a analogia de uma roda (VAUGHAN; DAVIS; O'CONNOR, 1992). Essa roda demonstra a natureza cíclica da progressão à frente e pode ser visualizada na Figura 5.

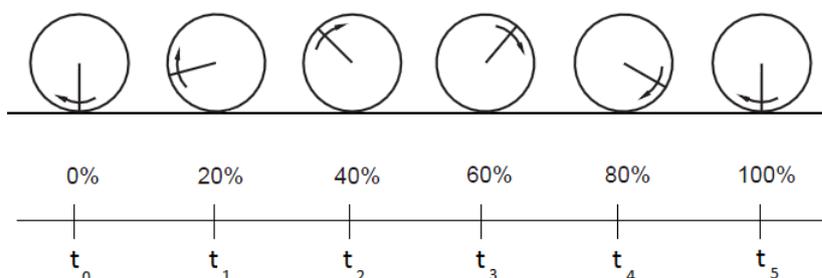


Figura 5: Analogia de movimento de uma roda demonstrando a natureza cíclica de um movimento à frente. Adaptado de Vaughan, Davis and O'connor (1992).

Nesse exemplo, a roda está movimentando-se da esquerda para a direita e no instante de visualização  $t_0$ , o raio aponta verticalmente para baixo. Por convenção dos autores, o início do ciclo é referenciado como 0%. Continuando o movimento da roda, o raio movimenta-se em sentido horário. Em  $t_1$  o movimento da roda atinge 20%, e percorreu 20% de 360° (72°) e a cada 20% de avanço, percorre mais 72° e em  $t_5$  o raio retorna ao mesmo estado que  $t_0$  e o ciclo se completa em 100%.

A analogia da roda pode ser aplicada ao movimento do caminhar humano. Apenas um ciclo da roda foi analisado, considerando que os próximos ciclos sucessivos, (de  $t_5$  a  $t_{10}$  se fossem capturados, por exemplo), seriam iguais. Normalmente descreve-se o ciclo do caminhar humano da mesma maneira, considerando que todos os ciclos

sucessivos são iguais. Segundo Vaughan, Davis e O'Connor (1992) esse pressuposto não é rigorosamente correto, mas é uma aproximação razoável para a maioria das pessoas.

### 2.2.2 Descrição do ciclo de caminhar

Um ciclo de caminhar é o intervalo entre sucessivas instâncias de batidas do calcanhar de um mesmo pé com o chão (CUNADO; NIXON; CARTER, 2003). Considerando o exemplo na Figura 5, o calcanhar do pé direito em contato com o chão é o raio alinhado verticalmente em  $t_0$  e o calcanhar do mesmo pé, novamente em contato com o chão, é o raio alinhado verticalmente em  $t_5$ , finalizando o ciclo.

Segundo Cunado, Nixon e Carter (2003), cada perna passa por dois períodos distintos: (i) postura ou *stance*, quando o pé está em contato com o chão e (ii) balanço ou *swing*, quando o pé está suspenso, adiante da próxima batida do calcanhar com o chão. O intervalo de tempo entre a batida de calcanhares opostos é chamado de passo e um ciclo de caminhar contém dois passos.

Na Figura 6, o tempo  $t_0$  marca o início da fase de postura, onde o tornozelo direito flexiona-se para posicionar o pé plano ao chão (pé-plano), e o peso do corpo é transferido para o mesmo, em  $t_1$ . A perna direita entra na fase de balanço em  $t_2$  e  $t_3$  com o desligamento do calcanhar do chão (calcanhar livre) e ultrapassa a perna esquerda em  $t_4$  terminando o ciclo com o polegar fora do chão (polegar livre).

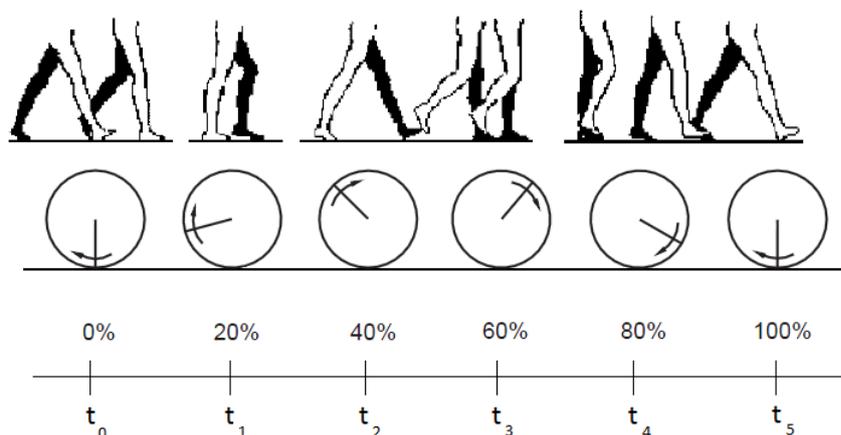


Figura 6: Ciclo de caminhar, adaptado de Cunado, Nixon e Carter (2003) e Vaughan, Davis and O'Connor (1992).

A Figura 7 mostra os termos usados para descrever o ciclo do caminhar humano. Cada fase do caminhar humano e seus tempos de duração foram apresentados no estudo realizado por Murray, Drought e Kory (1964), em *Walking Patterns of Normal Man*. Utilizaram uma técnica fotográfica simples e pouco dispendiosa para medir e temporizar componentes cinemáticos nos planos sagital, frontal e transversal dos indivíduos,

como ilustra a Figura 8. Para maiores detalhes do trabalho de Murray, Drought e Kory (1964) ver referência.

Os itens seguintes expõem algumas observações feitas por Murray, Drought e Kory (1964) em seu experimento e conceitua cada fase existente no ciclo de caminhar humano, apresentadas na Figura 7.

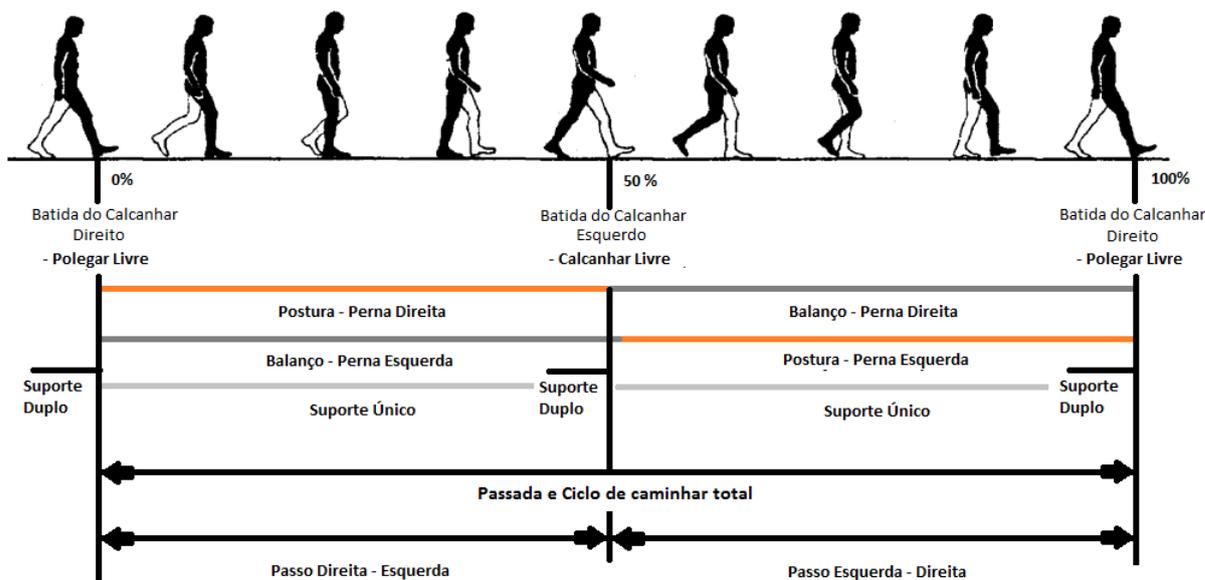


Figura 7: Fases do ciclo de caminhar, adaptado de Yam, Nixon e Carter (2004) e Cunado, Nixon e Carter (2003).

**Ciclo de caminhar total** O ciclo de caminhar total é o intervalo de tempo entre sucessivas batidas de um mesmo calcanhar no chão. Murray, Drought e Kory (1964) separaram em seu experimento o tempo médio do ciclo de caminhada de indivíduos por idade e peso e verificaram que não existia mudança significativa no tempo do ciclo de caminhar para testes repetidos, assim como verificaram não haver relação do tempo de ciclo de caminhar com a idade e a altura dos indivíduos (MURRAY; DROUGHT; KORY, 1964).

**Postura** Durante o caminhar, tanto o pé esquerdo como o direito entram em contato com o solo. Esta é a fase de postura e ocupa um percentual do tempo total do ciclo. No resto desse tempo, o pé está suspenso na fase de balanço. Murray, Drought e Kory (1964) constataram em seu experimento que não existem diferenças significativas entre a fase de postura direita ou esquerda no ciclo de caminhar de um mesmo teste e também entre testes diferentes para um mesmo indivíduo.

**Balanço** A fase em que o pé está suspenso no ar, movimentando-se para frente enquanto que o pé oposto está em contato com o solo é chamada de balanço.

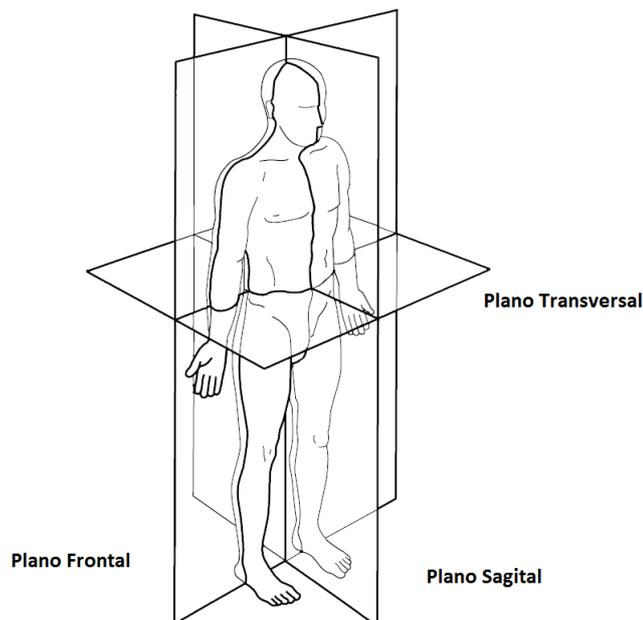


Figura 8: Planos para extrair componentes cinemáticos usados por Murray, Drought e Kory (1964) (VAUGHAN; DAVIS; O'CONNOR, 1992).

Existem, segundo os autores, dois períodos completos de balanço em um ciclo de caminhar total. Como nas outras fases, não existem diferenças significativas entre um mesmo teste e testes diferentes para um mesmo indivíduo (MURRAY; DROUGHT; KORY, 1964).

**Suporte duplo e suporte único** A fase de postura possui uma etapa de suporte único, quando um pé está em contato com o chão e o outro está no balanço, e duas etapas de suporte duplo, quando ambos os pés estão em contato com o chão, ao mesmo tempo. Em um ciclo de caminhada, existem ao todo duas etapas de suporte duplo, como mostra a Figura 5. Não foram detectadas diferenças entre testes para um mesmo indivíduo e também com relação a altura e a idade (MURRAY; DROUGHT; KORY, 1964).

**Comprimento do passo e da passada** Segundo Murray, Drought e Kory (1964), o comprimento da passada é a distância, no plano de progressão do ciclo de caminhar, entre sucessivos contatos de um mesmo pé com o chão (direito-a-direito e esquerdo-a-esquerdo). O comprimento do passo é a distância entre sucessivos pontos de contato de um pé com o chão e seu pé oposto (direito-a-esquerdo e esquerdo-a-direito). Ambos são medidos de um ponto central no eixo longitudinal do pé, que coincide com o final do calcanhar, internamente, como mostra a Figura 9.

Como nas outras fases do ciclo de caminhar humano, o comprimento do passo e da passada não mostraram diferenças significantes entre os testes feitos para

um mesmo indivíduo. Porém, os autores constataram que o comprimento médio da passada e do passo para o grupo de indivíduos de 60 a 65 anos eram menores que o comprimento médio dos indivíduos mais novos. Também, indivíduos mais baixos apresentaram os menores comprimentos de passos e passadas, e os mais altos, os maiores (MURRAY; DROUGHT; KORY, 1964).



Figura 9: Medida do passo e da passada a partir de um ponto do centro ao calcanhar do pé (MURRAY; DROUGHT; KORY, 1964).

### 2.2.3 Componentes do ciclo de caminhar

Os trabalhos de Murray, Drought e Kory (1964) e Cunado, Nixon e Carter (2003) apresentam evidências de que cada indivíduo possui um padrão de caminhar único, e que existem aproximadamente 20 componentes de movimento distintos no caminhar humano que fazem parte desse padrão. Alguns destes componentes são mensuráveis apenas no plano transversal, superior cabeça, e outros no plano sagital e frontal. Entre esses componentes, os autores consideraram em sua análise, as rotações sagitais da pélvis, quadris, joelhos e tornozelos, bem como os deslocamentos espaciais do tronco (oscilação vertical, oscilação lateral, e deslocamento para a frente).

No trabalho de Cunado, Nixon e Carter(2003), os autores limitaram-se à análise das rotações sagitais do joelho e quadris, devido as limitações dos sistemas de visão computacional na época de publicação do trabalho. Posteriormente, trabalhos relacionados apresentaram evidências de que essas articulações concentram as principais características que permitem a singularidade do caminhar humano para cada indivíduo.

Nos trabalhos subsequentes, Yoo e Nixon (2011) em “*Automated Markerless Analysis of Human Gait Motion for Recognition and Classification*”, apresentam a classificação dos componentes do ciclo de caminhar de forma mais elucidativa, em parâmetros espaço-temporais e parâmetros cinemáticos.

Com a evolução das pesquisas em biometria através do caminhar e o avanço das técnicas em visão computacional, outros trabalhos, como o de Yeoh et al. (2011), Nandini et al. (2011) e Ng et al. (2011), abordam o caminhar humano de forma holística, sem depender do comportamento dos parâmetros de movimento e sem considerar características do corpo humano na extração de parâmetros do caminhar. Esse tipo de abordagem é chamada de livre de modelo e será apresentada na Seção 2.3.4. Os parâmetros de movimento espaço-temporais e cinemáticos serão descritos nos itens

subsequentes, pois são parte da teoria do caminhar humano. A abordagem que leva em conta a teoria do caminhar é a abordagem baseada em modelo e será apresentada na Seção 2.3.1

**Parâmetros de movimento espaço-temporais** Parâmetros de movimento espaço-temporais são aquelas características do ciclo de caminhar que são medidas em função do tempo, como o tempo que o indivíduo leva para executar um ciclo de caminhar, a velocidade com que se desloca ou a duração de fases intermediárias do ciclo de caminhar. Como parâmetros espaço-temporais(ou temporais e espaciais), Yoo e Nixon (2011) definem o comprimento da passada, período do caminhar, tempo ou duração do ciclo de caminhar e velocidade do indivíduo, extraídos das figuras de movimento.

O período do caminhar é obtido pelo número de quadros de um ciclo de caminhar na sequência de imagens. A duração do ciclo de caminhar (ou tempo de ciclo) é obtida pelo período em razão da taxa de quadros de captura dos vídeos, como mostra a Equação 1. O comprimento da passada é a distância entre sucessivos pontos de contato de um mesmo pé durante o ciclo de caminhar, e é determinada pelos autores pelas coordenadas de deslocamento à frente das figuras de imagem. A velocidade do indivíduo é obtida utilizando o comprimento da passada em razão do tempo de ciclo, como mostra a Equação 2.

$$\text{Tempo de Ciclo}(s) = \frac{\text{Periodo de Caminhar (quadros)}}{\text{Taxa de Quadros (quadros/s)}} \quad (1)$$

$$\text{Velocidade (m/s)} = \frac{\text{Comprimento da Passada (m)}}{\text{Tempo de Ciclo (s)}} \quad (2)$$

**Parâmetros de movimento cinemáticos** Os parâmetros de movimento cinemáticos, segundo Yoo e Nixon (2011) são geralmente associados a ângulos formados entre segmentos do corpo e também pela maneira que se relacionam com as fases do ciclo de caminhar. Esses ângulos são medidos de uma articulação em relação a outra. O trabalho de Murray, Drought e Kory (1964) explica com detalhes a rotação do quadril, rotação dos joelhos e rotação dos tornozelos. Maiores detalhes sobre o comportamento de outros parâmetros cinemáticos consulte Murray, Drought e Kory (1964).

- Rotação dos quadril - Murray, Drought e Kory (1964) definiram a rotação dos quadril, como pode ser visto na curva apresentada na Figura 10, caracterizada por uma fase de extensão (Ex) e uma fase de flexão (Fl), no ciclo de caminhar. Segundo os autores, quando o ciclo inicia, com a batida do calcanhar, o quadril está em flexão. Quando o tronco começa a se mover à frente, o quadril estende

firmemente para assegurar uma posição ereta sobre a base que suporta o corpo. No final do passo (50% do ciclo de caminhar), coincide com a finalização da extensão. Nessa etapa, o quadril está estendido em média a 10 graus a mais do que quando o corpo está parado. Na próxima etapa, o quadril começa a entrar em flexão, para a fase balanço. Assim que a flexão do quadril se completa (85% do ciclo de caminhar), o quadril se mantém em sua posição relativa até a batida de calcanhar.

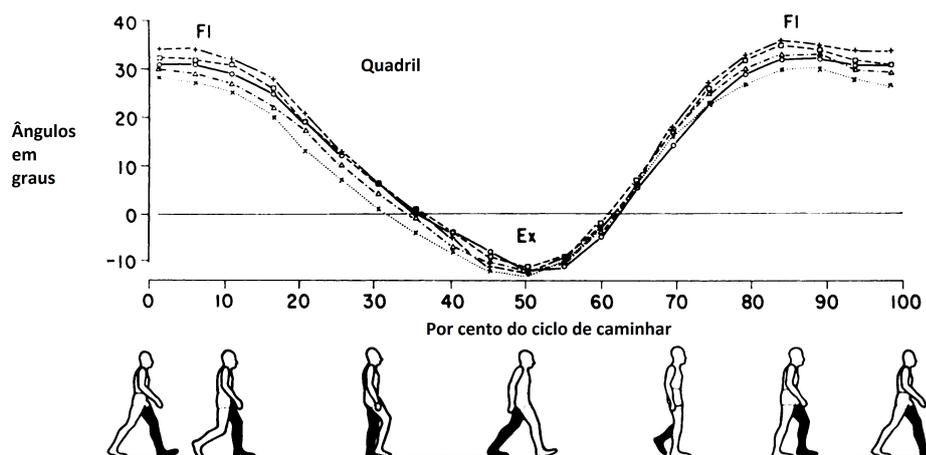


Figura 10: Ângulos em graus descritos pelo quadril durante o ciclo de caminhar. As diferentes curvas são referentes à variação de idade dos indivíduos (MURRAY; DROUGHT; KORY, 1964).

- Rotação dos joelhos - No ciclo de caminhar, a curva de rotação do joelho, exibida na Figura 11, possui duas fases de extensão: Ex1 e Ex2; e duas fases de flexão: FI1 e FI2. Segundo Murray, Drought e Kory (1964), na batida de calcanhar, no início do ciclo, o joelho inicia estendido, e conforme o indivíduo avança, o joelho entra na primeira e menor fase de flexão FI1. Conforme os autores, o pico de FI1 ocorre em torno de 15% do ciclo de caminhar. Conforme o indivíduo avança, o joelho entra na primeira fase de extensão Ex1 e seu pico ocorre em 40% do ciclo de caminhar. Entrando na fase de balanço, o joelho flexiona e entra na fase FI2 atingindo seu pico em 70% do ciclo. O joelho então, inverte sua direção e estende-se rapidamente e atinge sua posição mais reta, pouco antes da próxima batida de calcanhar, finalizando o ciclo.
- Rotação dos tornozelos - Assim como o joelho, as rotações do tornozelo, apresentadas durante o ciclo de caminhar na Figura 12 também possui duas fases de extensão: Ex1 e Ex2; e duas fases de flexão: FI1 e FI2. Após a primeira batida de calcanhar, o tornozelo entra na fase de extensão Ex1 que permite o contato do pé totalmente com o chão. A perna começa a rotacionar para frente sobre o pé fixo, enquanto o tornozelo inicia sua fase de flexão FI1. Aproximadamente em

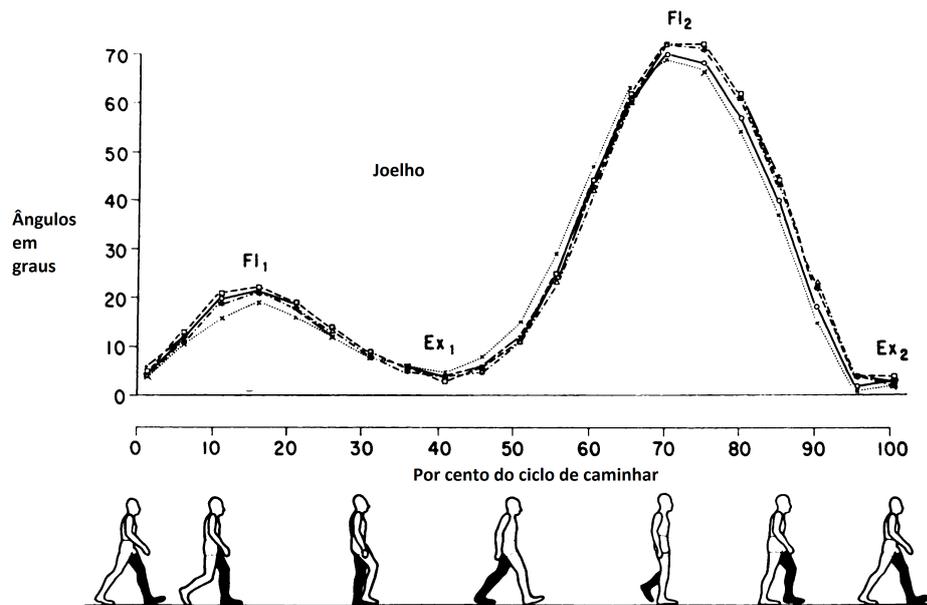


Figura 11: Ângulos, em graus, descritos pelo joelho dos indivíduos durante o caminhar (MURRAY; DROUGHT; KORY, 1964).

40% do ciclo de caminhar, o tornozelo reverte para a segunda fase de extensão Ex<sub>2</sub> que inicia lenta e torna-se rápida, elevando o calcanhar e mudando a base de apoio do pé para a parte dianteira. O tornozelo continua a extensão, enquanto que o joelho e os quadris estão na fase de flexão, até que o dedo do pé tenha se desligado do chão. Então o tornozelo prontamente se flexiona (FI<sub>2</sub>) para permitir que o dedo do pé tenha espaço livre durante a fase de balanço e então estende-se novamente à frente novamente para próxima batida de calcanhar (MURRAY; DROUGHT; KORY, 1964).

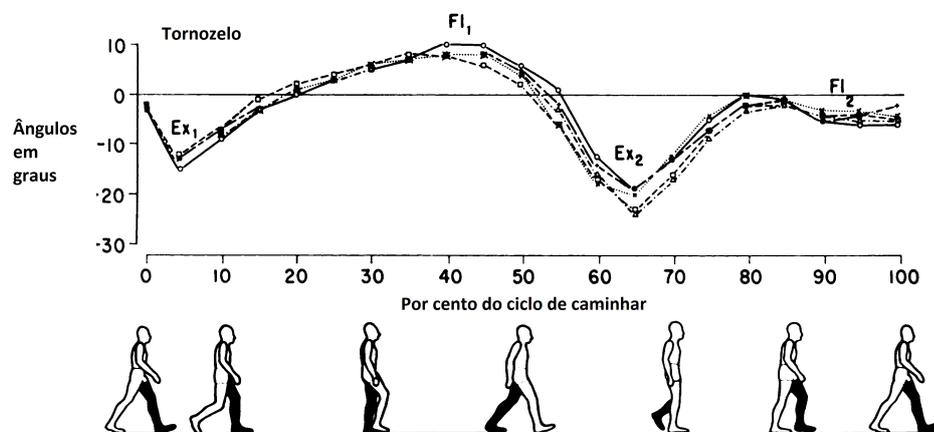


Figura 12: Ângulos em graus do tornozelo descritos durante o ciclo de caminhar (MURRAY; DROUGHT; KORY, 1964).

#### 2.2.4 Aquisição do caminhar humano

Em 1964, Murray, Drought e Kory utilizaram uma técnica de fotografia com *interrupted-light* e pessoas utilizando marcadores luminosos no corpo. Um espelho, localizado na parte superior do local de trajeto permitiu que características transversais pudessem ser analisadas. Atualmente, pesquisadores contam com diferentes bancos de dados para obter imagens de indivíduos caminhando normalmente, em ambiente controlado em laboratório, com *chroma-key*, caminhando em uma esteira e ao ar livre, além de diversos ângulos de visão, roupas diferenciadas, sapatos e acessórios para carregamento, como bolsas e mochilas. A Seção 2.2.5 apresenta alguns bancos de dados do caminhar humano que são comumente utilizados para obtenção de imagens de indivíduos caminhando.

#### 2.2.5 Bancos de dados do caminhar humano

**SOTON *Southamphthon database*:** O banco de dados SOTON contém vídeos de mais de 100 indivíduos diferentes caminhando, em sua maioria universitários jovens e saudáveis durante o verão. O banco possui vídeos em ambiente controlado e ao ar livre. No ambiente controlado (interno), o indivíduo caminha com um fundo *chroma-key* verde para facilitar a segmentação das imagens (YOO; NIXON, 2011).

A Figura 13 apresenta alguns quadros de vídeos obtidos do banco de dados SOTON. Na Figura 13(a), uma pessoa caminha no ambiente interno, com iluminação controlada e a câmera na posição normal, na Figura 13(b), ambiente interno controlado com a câmera oblíqua, na Figura 13(c), um indivíduo caminha no ambiente externo, ao ar livre. Na Figura 13(d) o indivíduo caminha em uma esteira, e na Figura 13(e) é apresentada uma imagem em alta resolução do indivíduo parado para fins de identificação facial e estimativa de altura (SHUTLER et al., 2004).

**CASIA *gait dataset*:** O instituto de automação da Academia Chinesa de Ciências (CASIA) proporciona aos pesquisadores da biometria através do caminhar um conjunto de banco de dados chamado *CASIA Gait Dataset*, que é dividido em *dataset* A, B, C e D. Cada um desses bancos de dados possui vídeos e imagens de indivíduos caminhando e diferenciam-se tanto pelo número de ângulos que envolvem, quanto pelo ambiente (externo ou interno) e o tipo de câmera utilizada (ZHENG et al., 2011).

O *dataset* A (Figura 14) foi criado em dezembro de 2001, e possui dados de 20 indivíduos durante o caminhar, ao ar livre, em forma de quadros de vídeo ou sequência de imagens. Cada pessoa possui 12 sequências de imagens: 4 sequências para cada uma das direções: paralela ao plano da câmera, 45° e

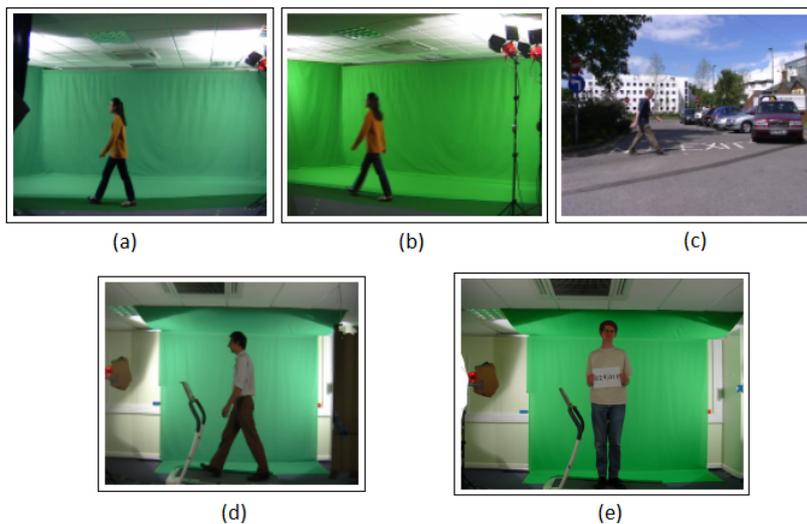


Figura 13: Quadros de vídeo do banco de dados do caminhar SOTON (SHUTLER et al., 2004).

90° ao plano da câmera. O tamanho da sequência é variável, entre 37 e 127 quadros, visto que as pessoas possuem velocidade de caminhar distintas. A Figura 14 mostra algumas imagens do *CASIA dataset A*.



Figura 14: *CASIA dataset A* apresentado quadros de sequências de vídeos de indivíduos caminhando ao ar livre com diversos ângulos de câmera (ZHENG et al., 2011).

O *dataset B* (Figura 15), de janeiro de 2005, é um grande banco de dados com 124 indivíduos, cujos dados do caminhar foram capturados em 11 ângulos de visão. O *dataset B* possui vídeos dos indivíduos caminhando e foram abordados aspectos como roupas e transporte de objetos. O banco de dados também fornece as silhuetas dos indivíduos segmentadas.

O *dataset C* (Figura 16) possui vídeos de 153 indivíduos caminhando, capturados por uma câmera infra-vermelho. Apresenta para cada pessoa um vídeo com o caminhar normal, devagar, rápido e normal com transporte de objeto. Segundo os autores, todos os vídeos foram capturados à noite, como apresentado na Figura 16.

Em 2009, foi preparado o *dataset D*, que além dos vídeos da caminhada, si-



Figura 15: Quadros capturados de seqüências de vídeo presentes no CASIA *dataset B*, em ambiente controlado incluindo indivíduos utilizando mochilas e outros acessórios (ZHENG et al., 2011).

multaneamente, foram recuperadas informações sobre a pressão do pé no chão de cada indivíduo, através de um *footscan*. Mais detalhes sobre o CASIA ver referência (ZHENG et al., 2011).

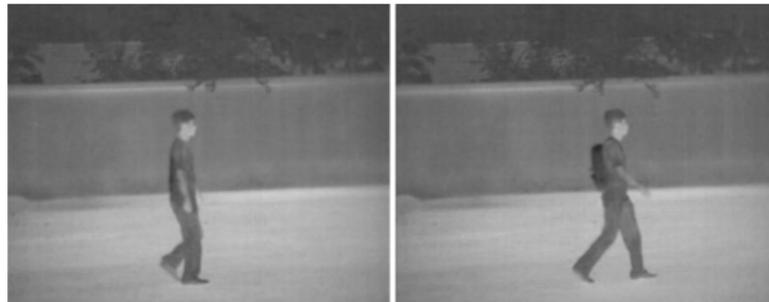


Figura 16: CASIA *dataset C*, contendo quadros capturados de seqüências de vídeo em infravermelho a noite (ZHENG et al., 2011).

**TUM-GAID *database*:** O *Technische Universität München* (TUM) *Gait from Audio, Image and Depth* (GAID) é um banco de dados do caminhar humano adquirido com o sensor Kinect com informações de vídeo (RGB), profundidade e também em áudio em quatro canais, apresentadas na Figura 17. As *streams* de vídeo e profundidade foram capturadas na resolução padrão do Kinect,  $640 \times 480$  em uma taxa de 30 quadros por segundo (HOFMANN; BACHMANN; RIGOLL, 2012).

O TUM-GAID possui um total de 176 indivíduos capturados, 103 homens e 73 mulheres. Cada indivíduo armazenado caminha 10 seqüências com as seguintes configurações: 6 caminhadas normais, 2 caminhadas carregando alguma bagagem (bolsa, mochila) para causar ruído visual e 2 caminhadas com um tipo de sapato especial que causa ruído no áudio. As caminhadas foram realizadas perpendicularmente ao sensor Kinect, a uma distância de aproximadamente 4 metros (HOFMANN; BACHMANN; RIGOLL, 2012).

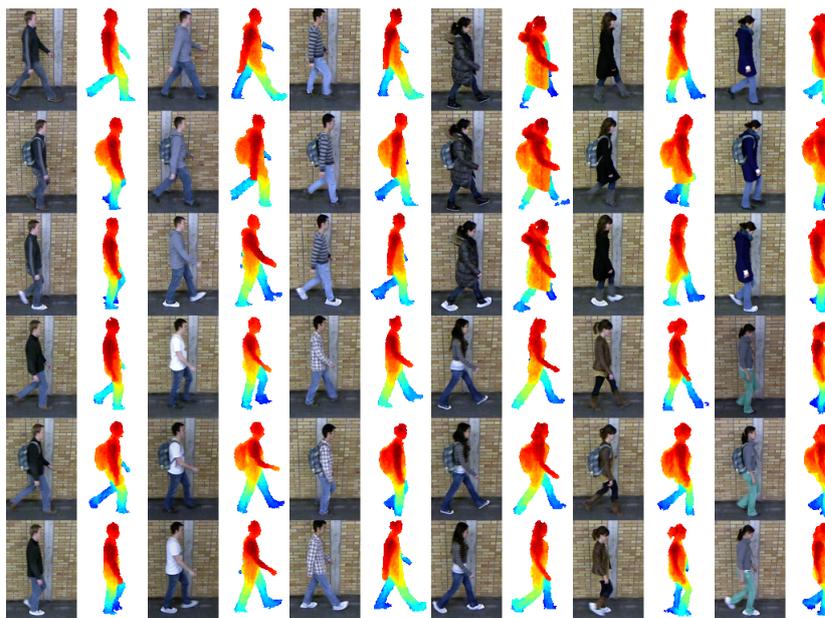


Figura 17: Quadros de *streams* de profundidade capturadas com o Kinect e armazenadas no banco TUM-GAID (HOFMANN; BACHMANN; RIGOLL, 2012).

## 2.3 Análise do Caminhar Humano

Para analisar o comportamento do caminhar humano, é necessário utilizar-se de uma representação. A análise e modelagem do caminhar humano pode ser dividida em duas categorias distintas, classificadas como (i) baseada em modelo (*Model-based*): onde a estrutura e os movimentos do corpo humano são modelados e os parâmetros do caminhar são extraídos do modelo, e (ii) livre de modelo (*Model-Free*): o movimento do corpo humano é representado de maneira concisa e independente da estrutura básica do corpo humano (YEOH et al., 2011).

### 2.3.1 Análise baseada em modelo

Considerando as características do ciclo de caminhar e seus parâmetros espaciais, temporais e cinemáticos, a análise baseada em modelo permite extrair diretamente esses parâmetros da posição das articulações do modelo. Aproxima-se, então, uma função matemática que descreve melhor o comportamento das articulações durante o ciclo de caminhar.

### 2.3.2 Modelo de Pêndulo

Para modelar o comportamento dos parâmetros cinemáticos (ângulos descritos pelas articulações), Cunado, Nixon e Carter (2003) propuseram a utilização do modelo de pêndulo. Nesse modelo, as pernas estão ligadas como dois pêndulos em série. Esse modelo é baseado na ideia de Murray, Drought e Kory (1964) que o caminhar apresenta características rítmicas (CUNADO; NIXON; CARTER, 2003). A Figura 18

apresenta o modelo de pêndulo, onde a perna superior é o segmento do quadril até o joelho e está suspensa pelo ponto que representa a articulação do quadril e a perna inferior é o segmento do joelho até o tornozelo, e está suspensa pelo joelho.

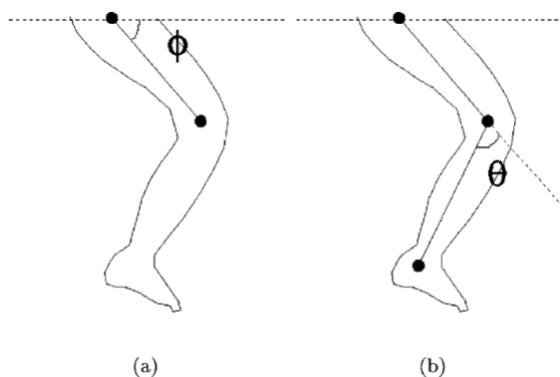


Figura 18: Modelo de pêndulo para obtenção dos ângulos (CUNADO; NIXON; CARTER, 2003).

Segundo Cunado, Nixon e Carter (1997) a teoria de Fourier permite que sinais periódicos sejam representados como uma fundamental harmônica, ou seja, o movimento dos membros inferiores no caminhar podem ser descritos dessa maneira.

O modelo de pêndulo permite que os padrões de rotação dos ângulos entre as articulações dos quadris e joelhos sejam tratados como sinais periódicos. Então, as técnicas de transformadas de Fourier podem ser usadas para obter a frequência do espectro do sinal periódico, e o espectro de diferentes indivíduos pode ser usado como característica distintiva ou única.

Cunado, Nixon e Carter (2003) aplicaram a transformada de Fourier nos padrões de rotação e o espectro de magnitude e de fase foram examinados. As parcelas de magnitude mostraram certa variação para cada pessoa, enquanto que as parcelas de espectro de fase mostraram grande variação entre os indivíduos. Segundo os autores, a maneira que a inclinação do quadril muda é de maior interesse do que o ângulo formado por ele. Para mais detalhes das transformadas utilizadas juntamente com esse modelo, consulte Cunado, Nixon e Carter (2003).

Yam, Nixon e Carter (2003) apresentaram em “*Automated person recognition by walking and running via model-based approaches*” o modelo de movimento do quadril e os modelos para descrever o movimento do fêmur (*thigh*) e do joelho (*knee*): (i) modelo simétrico bilateral, como mostra a Figura 19 e (ii) modelo do oscilador acoplado. Mais detalhes e especificações para ambos os modelos, consulte Yam, Nixon e Carter (2003).

O modelo bilateral simétrico foi proposto por Yam, Nixon e Carter (2003) baseado na observação do movimento aparente das pernas humanas. Elas podem ser modeladas, segundo os autores, como dois pêndulos acoplados em série, como mostra a

Figura 19. Seguindo a convenção biomecânica, a rotação do ângulo do joelho desse modelo é relativo a rotação do fêmur.

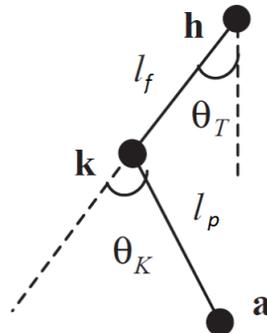


Figura 19: Os pêndulos superior e inferior modelam o fêmur e a perna inferior respectivamente, conectados pela articulação do joelho (k) (YAM; NIXON; CARTER, 2004).

### 2.3.3 Figuras de movimento

O modelo de pêndulo foi usado em trabalhos como o de Yam, Nixon e Carter (2004), e recentemente, figuras de movimento (*stick figures*), também conhecidas como “esqueletização” (*skeletonization*) foram usadas no trabalho de Yoo e Nixon (2011) para extração dos parâmetros de movimento. A ideia é representar o corpo humano com pontos que representam suas articulações principais e segmentos que representam os membros do corpo, como pode ser visto na Figura 20.

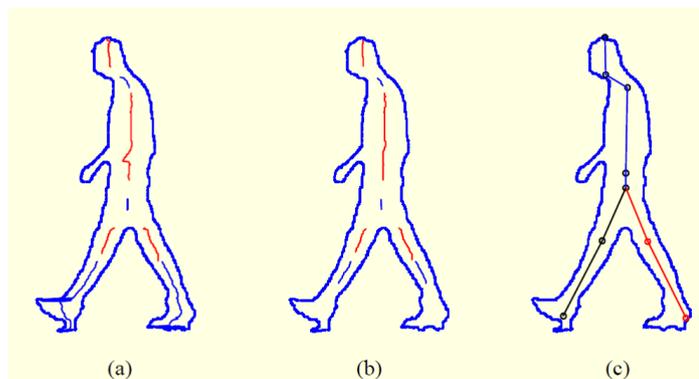


Figura 20: (a) esqueleto inicial estimado, (b) remoção de ruído e (c) figura final (YOO; NIXON, 2011).

Yoo e Nixon (2011), a partir de um estudo com medidas de cadáveres, estimaram um esqueleto inicial para silhuetas de pessoas caminhando provenientes do banco de dados SOTON. Para um corpo com altura  $H$ , a posição vertical estimada para o pescoço, ombros, coluna, pélvis, joelho e tornozelo foram setadas como sendo  $0.870H$ ,  $0.818H$ ,  $0.530H$ ,  $0.480H$ ,  $0.285H$  e  $0.039H$  respectivamente. A figura de esqueleto então é obtida através da Equação 3, para cada segmento do corpo  $p$

$$x_{s,p} = (x_{b,p} + x_{e,p}) \quad (3)$$

onde  $x_b$  e  $x_e$  e representam a posição horizontal do primeiro e último pixel da linha horizontal. Cada ponto do esqueleto é calculado utilizando a Equação 4

$$x_p, y_p = [x_i + L_p \cos(\phi + \theta_p), y_i + L_p \sin(\phi + \theta_p)] \quad (4)$$

onde  $\phi$  é o deslocamento de fase,  $x_i$  e  $y_i$  são as coordenadas das posições previamente estabelecidas e  $L_p$  é o tamanho dos segmentos do corpo, aproximados por dados de segmentos anatômicos (YOO; NIXON, 2011). Os ângulos  $\theta_p$ , onde  $p$  é o segmento dado, por sua vez, são aproximados usando regressão linear, como mostra a Equação 5

$$\theta_p = \tan^{-1} \left( \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})} \right) \quad (5)$$

onde  $n$  é o número de pontos do esqueleto no segmento do corpo. Para cada frame, os pontos do corpo são calculados utilizando as Equações 4 e 5. A trajetória das figuras de movimento contém os parâmetros espaço-temporais e descreve o movimento do caminhar, como apresentado na Seção 2.2.3.

Segundo Yoo e Nixon (2011), as trajetórias são uma função vetorial em cada quadro  $t$  de uma sequência de imagens, e apresentam uma periodicidade como no movimento do caminhar humano. Da mesma forma, os ângulos das articulações podem ser determinados pelas coordenadas dos pontos do corpo, e são medidos de uma articulação em relação a outra. Logo, o ângulo relativo em cada articulação é derivado do ângulo extraído (YOO; NIXON, 2011).

#### 2.3.4 Análise livre de modelo

A análise livre de modelo diferencia todo o padrão de movimento do caminhar através de uma representação concisa que não leva em conta a estrutura básica e fundamental do caminhar humano, apresentada no presente Capítulo. Nessa abordagem, são utilizadas como parâmetros do caminhar características aparentes extraídas durante o ciclo (YEOH et al., 2011), que não possuem conexão com o funcionamento fundamental do caminhar de uma pessoa.

Como exemplos de abordagens livre de modelo tem-se o trabalho de Sivapalan et al. (2011) que utilizaram imagens volumétricas obtidas da silhueta do indivíduo para reproduzir a energia do caminhar (GEI - *Gait Energy Images*) e propuseram uma extensão ao modelo GEI, o GEV (*Gait Energy Volume*) (SIVAPALAN et al., 2011). Também apresenta-se como abordagem livre de modelo a Caixa delimitadora (*bounding box*) que fornece características geométricas do caminhar de uma pessoa, como

altura, largura, área e um ângulo formado pela diagonal principal da caixa, utilizadas como parâmetros do caminhar humano (NANDINI et al., 2011).

Venkat e De Wilde (2010) propuseram uma modelagem baseada na exploração de características das subpartes do caminhar, segmentando a silhueta do indivíduo em partes independentes. Cada parte segmentada contém uma subparte do caminhar, que pode ser relacionada em um grafo acíclico direcionado, criando uma topologia de Rede Bayesiana.

Outros exemplos de análise livre de modelo são os trabalhos de Yeoh et al. (2011) e Ng et al. (2011), os quais a partir da silhueta segmentada dos indivíduos, inferiram matematicamente as posições dos joelhos e da forquilha (altura do chão até o osso púbis). Como parâmetros, os autores utilizam os ângulos descritos pelo joelho, pela forquilha, altura da forquilha, tamanho do passo, altura e largura da silhueta do indivíduo. Os trabalhos de Yeoh et al. (2011) e Ng et al. (2011) são classificados como livres de modelo, embora utilizem algumas características dos fundamentos do caminhar, tais como ângulos descritos pelos joelhos e tamanho do passo. As Seções 2.3.5 a 2.3.9 descrevem as abordagens livre de modelo utilizadas por alguns autores citados.

### 2.3.5 GEIs e GEVs

GEI (*Gait Energy Images*), ou Imagens da energia do caminhar, representam um padrão de movimento temporal dentro do ciclo de caminhar, em uma única imagem, que incorpora diversos parâmetros do caminhar humano, como frequência de movimento, mudanças temporais e espaciais do corpo assim como o padrão estatístico do formato do mesmo. O modelo GEI foi proposto por Han e Bhanu (2006) em "*Individual Recognition using Gait Energy Images*". Segundo Han e Bhanu (2006), os parâmetros espaço-temporais do ciclo de caminhar humano podem ser representados em uma única imagem 2D, pois de acordo com o autor, a ordem das poses no ciclo de caminhar humano é sempre a mesma e essas poses apresentam diferenças durante o ciclo. A Figura 21 mostra uma sequência de silhuetas normalizadas de duas pessoas que corresponde a um ciclo de caminhar e seus respectivos GEIs, na extrema direita.

Dada uma sequência de quadros de silhuetas binárias representadas por  $B_t(x, y)$ , que correspondem a um quadro do ciclo de caminhar em um tempo  $t$ , o GEI da sequência de imagens é definido pela Equação 6

$$G(x, y) = \frac{1}{N} \sum_{t=1}^N B_t(x, y) \quad (6)$$

onde  $N$  é o número de quadros no ciclo completo de uma sequência de silhuetas,  $t$  é o número do quadro na sequência (momento do tempo), e  $x$  e  $y$  são valores nas coordenadas da imagem 2D. Os GEIs, segundo o autor, são referenciados como

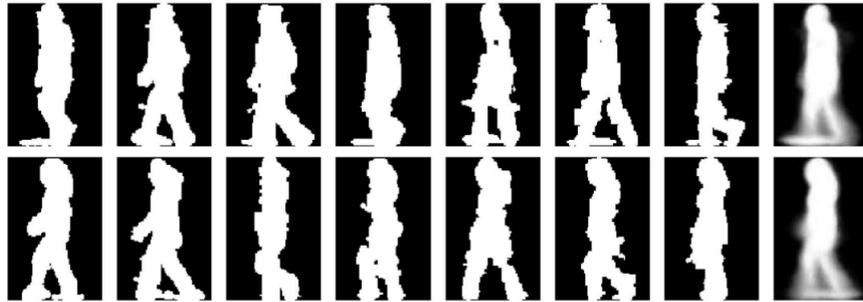


Figura 21: Silhuetas binárias normalizadas do ciclo de caminhar e seus respectivos GEIs (HAN; BHANU, 2006).

imagens de energia do caminhar pois (i) cada imagem de silhueta é a imagem espaço-normalizada da energia do caminhar humano, naquele momento, (ii) o GEI é a imagem tempo-normalizada da energia acumulada durante o ciclo completo de caminhar e (iii) um pixel com valor de intensidade elevada no GEI significa que o caminhar humano ocorre mais frequentemente naquela posição, com maior energia. A abordagem de representação de GEI economiza espaço de armazenamento e tempo computacional para classificação ao passo que é menos sensível a ruídos na silhueta em quadros individuais.

Os GEVs (*Gait Energy Volumes*) são uma extensão do conceito de GEI para 3D, apresentado por Sivapalan et al. (2011) em “*Gait Energy Volumes and Frontal Gait Recognition using Depth Images*”. Segundo Sivapalan et al. (2011), o GEI possui a inconveniência de ser dependente do ângulo de visão, ou seja, os GEIs funcionam melhor quando o caminhar do indivíduo é capturado de lado.

Ainda, segundo os autores, usar três dimensões contorna o problema da dependência de ângulo de visão. O método do GEV executa, como nos GEIs, uma média temporal nos volumes de voxel (pixels volumétricos) reconstruídos, ao invés de silhuetas segmentadas, como mostra a Figura 22 que apresenta o volume binário e seu correspondente GEV (SIVAPALAN et al., 2011).

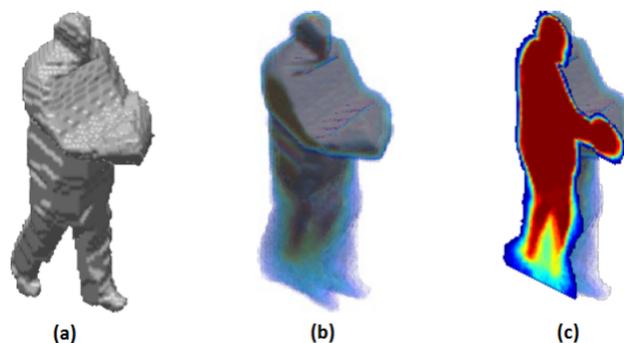


Figura 22: (a) Reconstrução do modelo de voxel, (b) GEV e (c) secção do GEV mostrando as diferentes densidades internas (SIVAPALAN et al., 2011).

A construção dos GEVs segue a mesma linha dos GEIs, como mostra a Equação 7

$$GEV(k) = \frac{1}{n} \sum_{t=1}^n V(t) \quad (7)$$

onde  $n$  é o número de quadros,  $t$  é o quadro atual e  $V$  os volumes de voxels alinhados. Para construir os GEVs de um ângulo de visão frontal, os autores criaram um pequeno banco de dados com imagens de profundidade de pessoas caminhando com o sensor Microsoft Kinect. A Figura 23 mostra um exemplo de quadro do banco criado pelos autores.

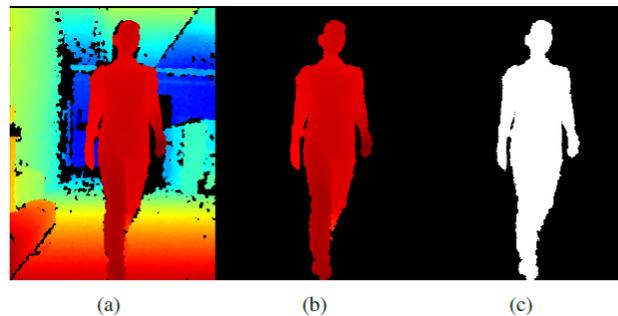


Figura 23: Exemplo de imagem capturada do banco de dados criado: (a) imagem de profundidade não preparada, (b) silhueta 3D, (c) silhueta binária (SIVAPALAN et al., 2011).

Como as abordagens livre de modelo, os GEI e GEV não procuram parâmetros do caminhar definidos pela teoria do caminhar humano, como ângulos entre articulações e comprimento do passo. A extração das características das imagens de energia, que podem ser interpretadas como atributos, tanto no GEI quanto GEVs, foram obtidas utilizando *Multiple Discriminant Analysis* (MDA) e submetidas a um filtro Bayesiano. Para mais detalhes sobre a técnica de MDA ver Sivapalan et al. (2011).

### 2.3.6 Caixa delimitadora

A Caixa delimitadora (*bundle rectangle*) é uma técnica livre de modelo utilizada por Guillen, Padilla e Hernandez (2009) em "*Gait Recognition System: Bundle Rectangle Approach*", onde os autores propõem a extração da assinatura do caminhar de uma sequência de imagens com a silhueta segmentada, através dos atributos da área, diagonal principal (e sua orientação), altura, largura, e o espectro de potência total da caixa delimitadora. Para cada quadro do vídeo de caminhar disponível no banco de dados utilizado pelos autores, é feita uma segmentação da silhueta do indivíduo. A Caixa delimitadora é então gerada em torno da silhueta segmentada possibilitando o fornecimento dos parâmetros utilizados, como mostra a Figura 24.

A Caixa delimitadora, posteriormente, foi utilizada por Nandini et al. (2011), com o

nome de “*Bounding box*” onde utilizam os mesmos atributos que os autores originais juntamente com o pico da Transformada de Hough. A Transformada de Hough é um método utilizado para encontrar linhas, círculos ou outras formas geométricas simples em uma imagem. Originalmente, a transformada foi proposta para encontrar linhas e posteriormente generalizada para as outras formas (BRADSKI; KAEHLER, 2008).

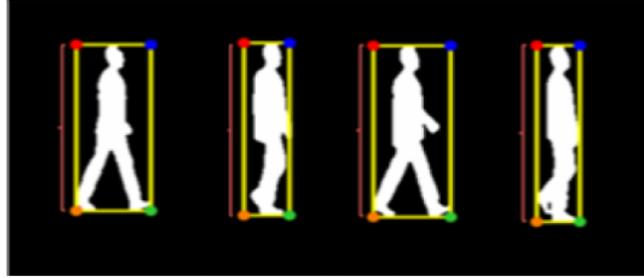


Figura 24: Silhueta com caixa delimitadora: O segmento em que liga os vértices da direita representa a altura estimada (GUILLEN et al., 2009).

A altura média extraída da Caixa delimitadora é dada pela Equação 8

$$H = \frac{\sum_{i=1}^{N_H} [y_{max}(t) - y_{min}(t)]_i}{N_H} \quad (8)$$

onde  $N_H$  é o total de quadros capturados e  $y_{max}$  e  $y_{min}$  são as coordenadas  $y$  dos pontos superior e inferior da caixa delimitadora, respectivamente.

Assim como a altura da caixa, a largura média é extraída pela Equação 9

$$W = \frac{\sum_{i=1}^{N_W} [x_{max}(t) - x_{min}(t)]_i}{N_W} \quad (9)$$

onde  $N_W$  é o total de quadros capturados e  $x_{min}$  e  $x_{max}$  são os pontos mais a esquerda e mais a direita, respectivamente, da caixa delimitadora, no eixo  $x$ . A partir da altura e da largura da caixa delimitadora os autores obtiveram a valores da área da caixa em cada quadro, dada pela Equação 10

$$A = \frac{\sum [H(t)W(t)]_i}{N_A} \quad (10)$$

onde  $N_A$  é a quantidade de quadros capturados e  $H(t)$  e  $W(t)$  os valores calculados para altura e largura da caixa no quadro atual.

Outro parâmetro extraído é o ângulo da diagonal principal da caixa, dado pela Equação 11 onde a média do arco-tangente da altura  $H(t)$  e da largura  $W(t)$  em cada quadro é dada por

$$\alpha = \frac{\sum_{i=1}^{N_\alpha} \left[ \tan^{-1} \left( \frac{H(t)}{W(t)} \right) \right]_i}{N_\alpha} \quad (11)$$

onde  $N_\alpha$  é a quantidade de quadros capturados.

O último parâmetro extraído da Caixa delimitadora é o potência espectral total, dada pelo somatório de cada aplicação da Transformada Rápida de Fourier (FFT) nas áreas calculadas em cada quadro  $A(f)$ , como mostra a Equação 12

$$P = \sum_{f=1}^{N_f} A(f) \quad (12)$$

onde  $N_f$  representa o número de quadros capturados.

### 2.3.7 Divisão do caminhar humano em subpartes

Venkat e De Wilde em “*Robust gait recognition by learning and exploiting sub-gait characteristics*” (2010) segmentaram silhuetas de indivíduos em imagens do caminhar, em subpartes (*subgaits*), cada parte contendo um pedaço do caminhar humano. A Figura 25 mostra um quadro típico de caminhar com suas subpartes identificadas como  $U$  (caminhar superior),  $M$  (caminhar médio),  $L$  (caminhar inferior),  $LF$  (caminhar esquerdo) e  $R$  (caminhar direito), e agrupadas no conjunto  $S = \{U, M, L, LF, R\}$ .

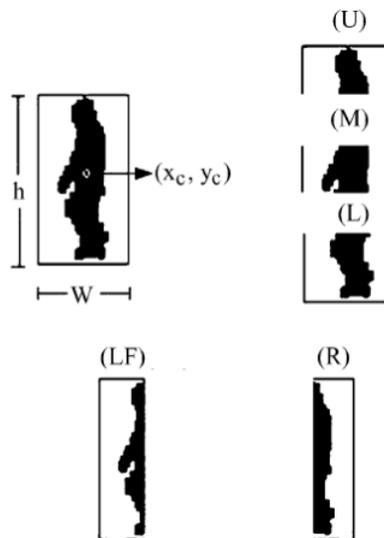


Figura 25: Divisão do caminhar em subpartes (VENKAT; WILDE, 2010).

Manipulando os arquivos binários que representam as silhuetas, os autores computaram uma caixa delimitadora, como descrito na Seção 2.3.6, e a redimensionaram para 64 x 44 pixels. Para definir as subpartes do caminhar, os autores utilizaram morfologia matemática. Para maiores detalhes das Equações utilizadas na obtenção do modelo consulte Venkat e De Wilde (2010).

### 2.3.8 Divisão do corpo em segmentos

Yeoh et al. (2011) em “*Improved Gait Recognition with Automatic Body Joint Identification*” propuseram uma abordagem livre de modelo em que a silhueta do corpo do indivíduo caminhando, em cada quadro, é dividida em oito segmentos. Cada articulação é automaticamente identificada e suas trajetórias computadas. Utilizaram imagens de silhuetas segmentadas do banco de dados SOTON com variações de uso de roupas e acessórios de carregar.

Aplicaram nas silhuetas técnicas de erosão seguidas de dilatação, para remover ruídos causados pela sombra no pé da pessoa. A altura e largura da silhueta foram medidas, e posteriormente, foram utilizadas como atributos na classificação do caminhar. Baseado em um estudo da proporção humana, os autores dividiram a silhueta em oito partes, como mostra a Figura 26,

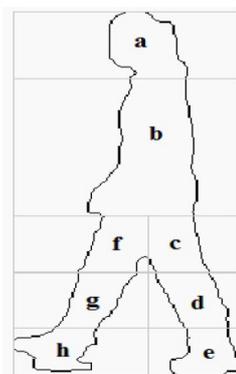


Figura 26: Divisão da silhueta em segmentos (YEOH et al., 2011).

onde *a* representa a cabeça e o pescoço, *b* representa o tronco, *c* representa o quadril e o fêmur superior direito, *d* representa a parte inferior do fêmur direito, próximo ao joelho, *e* representa a panturrilha inferior direita e o pé. Do lado esquerdo, *f* representa o quadril e a parte superior do fêmur, *g* representa a parte inferior do fêmur próximo ao joelho e *h* a panturrilha inferior juntamente com o pé.

Segundo Yeoh et al. (2011), para extrair as articulações que definem o caminhar humano, as posições verticais do quadril, dos joelhos e tornozelos em relação a altura do corpo foram estimadas pelas proporções do corpo humano definidas a priori. Uma linha horizontal foi desenhada de lado a lado das articulações e a posição horizontal das duas bordas foi calculada pela Equação 13

$$x_c = x_1 + \frac{x_2 - x_1}{2} \quad (13)$$

onde  $x_c$  é a posição central,  $x_1$  e  $x_2$  são a posição horizontal da primeira borda e da segunda borda na imagem, respectivamente, ao longo da linha horizontal que passa através da articulação.

Yeoh et al. (2011) também calcularam, a partir das posições das articulações encontradas, a variação dos ângulos o qual chamaram de trajetória das articulações. Esses ângulos e todos os atributos coletados da divisão da silhueta do caminhar em subpartes podem ser observados na Figura 27. A Equação 14 mostra como é calculado a variação angular de todas as articulações  $\theta$ ,

$$\theta = \tan^{-1} \left( \frac{p2_x - p1_x}{p2_y - p1_y} \right) \quad (14)$$

e para obter o tamanho do passo, é calculada da distância euclidiana entre os dois tornozelos em cada frame da sequência de caminhada (YEOH et al., 2011).

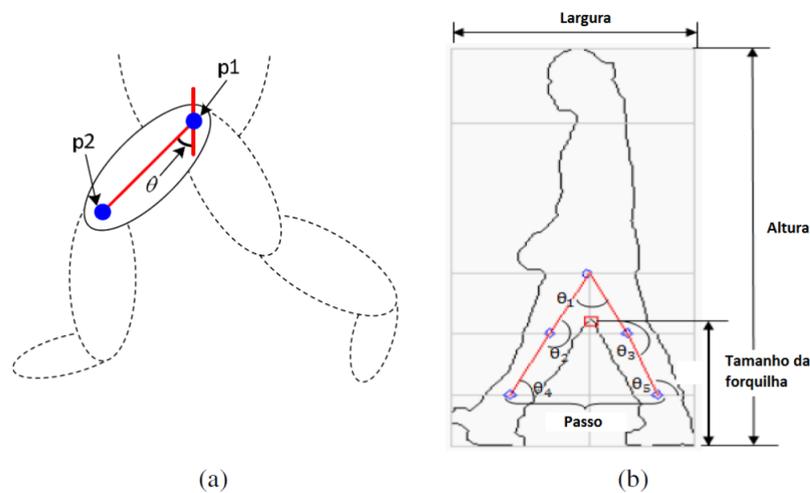


Figura 27: (a) Ângulos da forquilha e joelho e atributos coletados da silhueta (YEOH et al., 2011).

### 2.3.9 DGHEIs

Em “2.5D Gait Biometrics using the Depth Gradient Histogram Energy Image”, Hofmann, Bachmann e Rigoll (2012) apresentam uma técnica livre de modelo, para extração de parâmetros do caminhar, intitulada “Imagens de energia do histograma do gradiente de profundidade” ou *Depth Gradient Histogram Energy Image* - DGHEI.

A técnica de DGHEI consiste em utilizar o armazenamento de histogramas, motivado pelo conceito de “*histograms of oriented gradients*” (HOG), frequentemente utilizados para detecção de objetos em imagens. Para extrair o DGHEI primeiramente é necessário calcular histogramas de gradientes orientados em cada quadro  $t$ , como mostra a Equação 15 e a Equação 16:

$$r(x, y) = \sqrt{u(x, y)^2 + v(x, y)^2} \quad (15)$$

$$\theta(x, y) = \tan^{-1} 2(u(x, y), v(x, y)) + \pi \quad (16)$$

que obtém a magnitude  $r$  e a orientação  $\theta$  dos dados de profundidade e  $v(x, y)$  e  $u(x, y)$  definidos pelas Equações 17 e 18.

$$u(x, y) = I(x - 1, y) - I(x + 1, y) \quad (17)$$

$$v(x, y) = I(x, y - 1) - I(x, y + 1) \quad (18)$$

As orientações dos gradientes em cada pixel são discretizadas em 9 orientações, segundo a Equação 19.

$$\theta(x, y) = \left\lfloor \frac{9 \times \theta(x, y)}{2\pi} \right\rfloor \quad (19)$$

Esses gradientes discretizados são então agregados em células, contendo tipicamente  $8 \times 8$  pixels, formando uma grade densa destas células sem sobreposições. Cada célula é formada por histogramas de gradientes orientados com 9 bandas, e posteriormente, são normalizadas 4 vezes por blocos de células vizinhas, levando a 36 valores por cada célula. Finalmente, os histogramas de gradiente são agrupados em uma média sobre todo o ciclo de caminhar consistindo de  $T$  quadros, segundo a Equação 20

$$H(i, f, j) = \frac{1}{T} \sum_{t=1}^T h_t(i, j, f) \quad (20)$$

que resulta no DGHEI, apresentado na Figura 28. Na Equação 20,  $i$  e  $j$  apontam para o histograma localizado na célula  $(i, j)$  e  $f = 1 \dots 36$  é índice da banda do histograma. Cada ciclo de caminhar é então representado por um vetor multidimensional  $H(i, j, f)$ . Segundo Hofmann, Bachmann e Rigoll (2012) os DGHEIs acrescentam profundidade aos GEIs (HAN; BHANU, 2006), assim como os GEVs (SIVAPALAN et al., 2011) ambos apresentados na Seção 3.2.1.

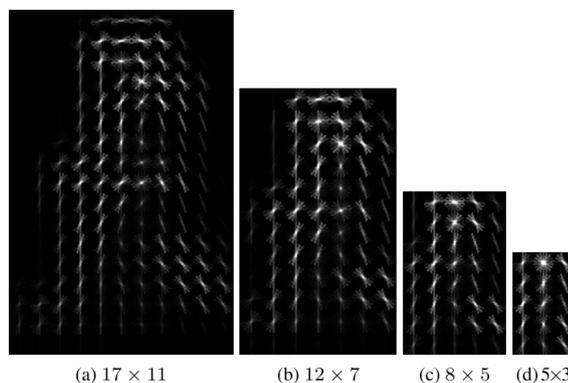


Figura 28: Representação visual dos DGHEI utilizando diferentes configurações (HOFMANN; BACHMANN; RIGOLL, 2012).

## 2.4 Reconhecendo e classificando o padrão de caminhar

A característica mais distintiva do caminhar humano é o fato de que é individualista (YOO; NIXON, 2011), ou seja, cada pessoa possui seu padrão de caminhar, como já foi abordado anteriormente. Devido a essa característica, é possível classificar uma pessoa pela maneira que ela caminha; reconhecê-la entre outros indivíduos, o que torna o caminhar humano uma forma de biometria.

Os parâmetros do caminhar humano, explicados na Seção 2.2, temporais, espaciais e cinemáticos, são utilizados como atributos na classificação e no reconhecimento de pessoas. Alguns parâmetros podem ser utilizados diretamente, como o tempo que o indivíduo leva para executar um ciclo de caminhar, o tamanho da passada e do passo e a variação dos ângulos entre articulações, e outros podem ser derivados de “indicadores estatísticos”, como média, mediana e desvio padrão. Cada indivíduo possui então um vetor de atributos extraídos do ciclo de caminhar. Outras características como altura e tamanho dos membros podem ser utilizadas em conjunto, dentro dos parâmetros temporais e espaciais.

Com esse vetor de atributos, é possível utilizar um algoritmo de aprendizagem, que classifique os padrões de cada indivíduo, dado um conjunto de exemplos. As Seções 2.4.1 a 2.4.6 discutem alguns algoritmos que foram utilizados em trabalhos de investigação do caminhar humano como forma de biometria, que serão utilizados nos experimentos descritos na metodologia do presente trabalho.

### 2.4.1 *k-Nearest Neighbor*

Para a maioria dos experimentos de reconhecimento do caminhar humano, o algoritmo *k-Nearest Neighbor*, ou K-NN, é a primeira escolha entre os algoritmos de aprendizagem de máquina existentes para classificação do conjunto de exemplos (YOO; NIXON, 2011), sendo também uma escolha recomendada quando existe pouca ou nenhuma informação sobre a distribuição dos dados a serem classificados (PETERSON, 2009).

O algoritmo de aprendizado supervisionado K-NN é utilizado para classificar um objeto não rotulado, baseado no rótulo de seus vizinhos mais próximos em um espaço de exemplos. Essa proximidade é, frequentemente, baseada em uma métrica de distância entre dois pontos, por exemplo, a distância Euclidiana. De maneira simplificada, a regra de classificação do k-NN é associar a uma amostra de teste, o rótulo da maioria das categorias de seus “k” vizinhos mais próximos (PETERSON, 2009).

Segundo Cover e Hart (1967), um conjunto de treino  $X$  consiste em  $n$  pares de vetores e rótulos, dispersos em um espaço de classes. Dado um novo par  $(x, \theta)$ , onde apenas a medida  $x$  é observável, o valor de  $\theta$  é estimado pela utilização dos dados contidos no conjunto  $X$  com os vetores e rótulos já conhecidos (supervisionado). Um

vetor  $x'$ , segundo os autores, é um vizinho mais próximo de  $x$ , se segundo a Equação 21

$$d(x'_n, x) = \min d(x_i, x) \quad i = 1, 2, \dots, n \quad (21)$$

a distância mínima entre um vetor vizinho e o vetor testado são iguais ao conjunto de distâncias mínimas entre os outros vetores e o vetor testado. Normalmente, o “voto” do vizinho possui um peso, de acordo com a distância do vetor testado e seus “k” vizinhos mais próximos. Assim,  $\theta$  é estimado com o rótulo desses “k” vizinhos mais próximos, como mostra a Figura 29.

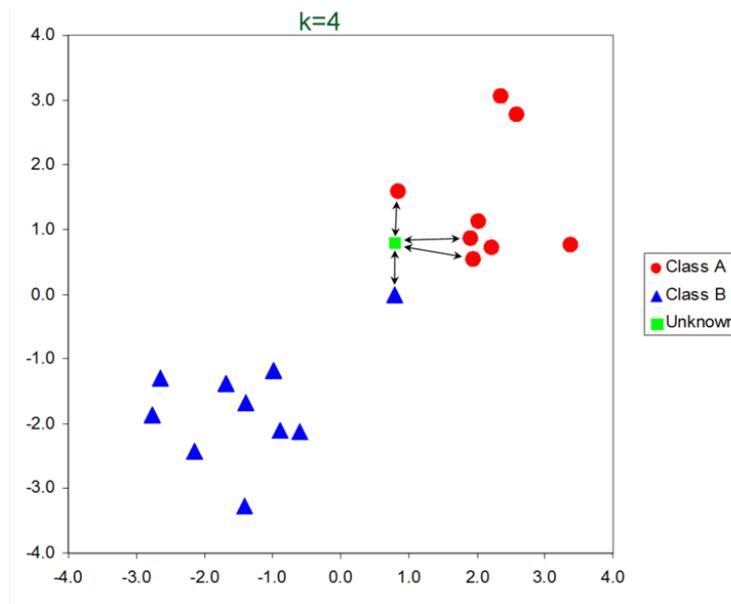


Figura 29: Representação visual de 19 vetores rotulados com classe A e B e uma classe desconhecida. O vetor com classe desconhecida é portanto, rotulado com seus 4 ( $k=4$ ) vizinhos mais próximos de acordo com a distância euclidiana entre eles (PETERSON, 2009).

#### 2.4.2 Aplicação do k-Nearest Neighbor na classificação de indivíduos através do caminhar

Para classificar um indivíduo através de seu caminhar, Yoo e Nixon (2011) utilizaram em sua abordagem baseada em modelo um vetor com 10 atributos: altura, tempo de ciclo de caminhar, tamanho da passada, velocidade, média dos ângulos entre as articulações (aquelas determinadas nos parâmetros cinemáticos do caminhar), variação do ângulo dos quadris e as coordenadas centrais do ciclograma quadril Joelho.

O conjunto de dados que os autores utilizaram é proveniente do banco de dados SOTON com ambiente controlado (*indoor*), descrito na Subseção 3.3.1. De 100

indivíduos, foram utilizadas sete sequências de imagens para cada um desses indivíduos. Um total de 500 vetores de atributos extraídos de 4 dessas sequências foram usados como conjunto de treino, e 100 vetores de atributos extraídos das 3 sequências restantes dos indivíduos foram usados como conjunto de teste (YOO; NIXON, 2011).

Os resultados obtidos por Yoo e Nixon (2011) com o uso do algoritmo k-NN variam de 82% a 96.7% de classificações corretas, com grupos de 30, 60 e 100 indivíduos e mostram que os atributos utilizados possuem contexto no reconhecimento de indivíduos através do caminhar.

Nos trabalhos de Yeoh et al. (2011) e Ng et al. (2011), com uma abordagem livre de modelo, foram utilizados os atributos: tamanho do passo, largura da silhueta do indivíduo, altura, trajetória do joelho esquerdo, trajetória do joelho direito e altura da forquilha. Os autores também adicionaram ao vetor outros atributos para melhorar a classificação dos indivíduos, tais como: média aritmética dos máximos locais para largura da silhueta, para altura da forquilha, trajetória dos quadris, trajetória do tornozelo direito, e para o tamanho do passo. Ao todo, os 15 atributos foram utilizados no vetor para cada indivíduo.

Em Yeoh et al. (2011), foram utilizadas sequências de imagens do banco de dados SOTON, de 11 indivíduos caminhando em paralelo a uma câmera estática. Dessas sequências, foram utilizados 13 fatores covariantes: 10 tipos de bolsas e 3 tipos de velocidades de caminhar. Além disso, cada indivíduo utilizou diferentes sapatos em diferentes sequências. Desse conjunto de dados, 1360 sequências de caminhadas foram utilizadas como conjunto de treino e 1362 sequências de caminhadas foram utilizadas como conjunto de teste.

Os autores utilizaram uma extensão do algoritmo k-NN com técnicas fuzzy, conhecido como *fuzzy k-NN*. A equação fuzzy no algoritmo k-NN computa o valor associação de um indivíduo não classificado (sem rótulo) pelo valor da associação de um indivíduo classificado e a distância entre o indivíduo não classificado e seus k vizinhos mais próximos. Para as 1362 sequências, e k variando entre 1 e 8, Yeoh et al (2011) obtiveram 86.56% a 87.44% e Ng et al. (2011), de 66.62% a 80.58% de acertos na classificação dos indivíduos.

### **2.4.3 Redes neurais artificiais**

O campo das redes neurais foi proposto originalmente por psiquiatras e neurobiólogos na tentativa de sintetizar o comportamento de um neurônio humano. Segundo Han e Kamber em "*Data Mining - Concepts and Techniques*" (2006), a grosso modo, uma rede neural é um conjunto de unidades de entrada e saída conectadas em que cada conexão possui um peso associado. Na fase de aprendizagem, a rede neural aprende ajustando seus pesos, para poder predizer um rótulo correto de classe das tuplas de entrada.

Redes neurais, segundo os autores, envolvem longos períodos de treinamento, e são mais adequadas a aplicações onde isso é viável. Além disso, as redes neurais requerem um número de parâmetros referentes a sua topologia ou estrutura que são melhor determinados empiricamente. Também não é fácil para os humanos interpretar o significado simbólico por trás dos pesos resultantes na rede. Em contrapartida, redes neurais possuem um alto grau de tolerância a ruídos nos dados, podem ser usadas quando não se tem conhecimento das dependências e relações entre os atributos e classes e são adequadas quando existem valores contínuos nas entradas e saídas da rede (HAN; KAMBER, 2006). Ainda segundo os autores, as redes neurais artificiais foram aplicadas com sucesso no reconhecimento de caracteres escritos à mão, patologias e remédios laboratoriais, no treinamento de pronúncia de um texto em inglês e no reconhecimento de padrões.

#### 2.4.4 Redes neurais na classificação de indivíduos através do caminhar

Redes neurais foram igualmente utilizadas para classificação de indivíduos no trabalho de Yoo e Nixon (2011). Após uma verificação da validade dos atributos utilizados no experimento pelo algoritmo k-NN, os autores utilizaram uma rede neural com *back-propagation* para identificar pessoas através do caminhar, utilizando sequências de imagens do banco de dados SOTON. Os 10 atributos do caminhar humano selecionados para cada indivíduo foram utilizados pelos autores como dados de entrada e o número de neurônios nas camadas ocultas foi setado para 28 e 13 respectivamente. A rede neural atingiu 90% de acertos para 30 indivíduos. Baseado nesse experimento, Yoo e Nixon (2011) concluíram que, apesar da rede neural não atingir 100% de acertos, os resultados mostraram que as pessoas são únicas no seu padrão de caminhar, corroborando com outros resultados similares.

#### 2.4.5 SVM - *Support Vector Machines*

As *Support Vector Machines* foram propostas por Cortes e Vapnik, em 1995 como ferramenta de aprendizado supervisionado para uso geral em reconhecimento de padrões (SUDHA; BHAVANI, 2012). É comumente aplicada em reconhecimento de escrita manual, de objetos em imagens, identificação por voz e reconhecimento de face (BURGES, 1998).

Na SVM, a entrada original é mapeada em um espaço de características, onde o hiperplano ótimo é determinado para maximizar a habilidade de generalização do classificador, como mostra a Figura 30. A SVM padrão é um classificador binário que pode identificar se uma instância pertence a uma classe ou não, definida pela Equação 22

$$\theta = \{(x_1, y_1), (x_2, y_2) \dots (x_i, y_i)\}, x_i \in R^n, y_i = \{1, -1\} \quad (22)$$

onde  $x_i$  são vetores de características (exemplos) e  $y_i$  são os rótulos de classe correspondentes. Seu funcionamento define fronteiras separando duas classes na forma de um hiperplano linear no espaço de exemplos onde a distância entre as fronteiras das duas classes é definida como margem do hiperplano (SUDHA; BHAVANI, 2012).

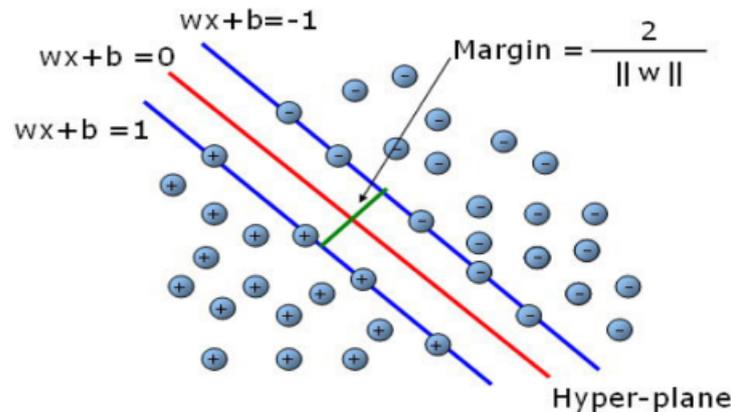


Figura 30: Representação de uma *Support Vector Machine* de classificação binária (SUDHA; BHAVANI, 2012).

#### 2.4.6 SVM multiclasse

Existem dois tipos de SVMs multiclasse: (i) uma classe contra todas (OAA - *one against all*) e (ii) uma classe contra uma classe (OAO - *one against one*) (SUDHA; BHAVANI, 2012). A SVM OAA treina  $k$  SVMs binárias onde  $k$  é o número de classes. A SVM  $k^i$  é treinada com todos os exemplos da classe  $i$  como exemplos “positivos”, e o restante dos exemplos de outras diferentes classes como exemplos “negativos”. Todas as  $k$  SVMs treinadas geram  $k$  funções de decisão. A sequência de teste é então rotulada de acordo com a maior saída entre os  $k$ , segundo a Equação 23:

$$c = \max(w_i x + b) \quad (23)$$

onde  $c$  é a classe da entrada  $x$  e  $i = [1, k]$  o número de classes (SUDHA; BHAVANI, 2012).

Na SVM OAO cada classe é combinada com todas as demais classes, gerando pares  $(i, j)$  correspondentes a cada  $SVM_{ij}$  treinada. Assim, serão treinados  $k(k-1)/2$  SVMs e obtidas  $k(k-1)/2$  funções de decisão. Para cada exemplo de entrada, é decidido a que classe ele pertence, computando um voto para a classe que ele foi votada, como mostra a Equação 24:

$$c = \text{sign}(w_{ij} x + b) \quad (24)$$

onde, se  $x$  pertence à classe  $i$ , então o voto para a classe  $i$ , caso contrário o voto

para classe  $j$  é incrementado. Finalmente,  $x$  pertence à classe mais votada (SUDHA; BHAVANI, 2012).

## 2.5 O Sensor Microsoft Kinect

Em Novembro de 2010, a Microsoft apresentou pela primeira vez o Microsoft Kinect para video-games XBOX 360. Era um dispositivo que, ligado ao XBOX, possibilitaria o usuário jogar sem o uso de controles, apenas com os movimentos do corpo e comandos de voz. A ideia do dispositivo Kinect foi concebida em 2009 pelo brasileiro Alex Kipman, que também batizou o nome-código do projeto Kinect como “*Project Natal*” homenageando também o seu país de origem (MICROSOFT, 2009).

### 2.5.1 Hardware e funcionamento do Kinect

O Kinect é um sensor equipado de uma câmera RGB, um sensor de profundidade composto de um emissor de luz infra-vermelho e uma câmera sensível à profundidade, como mostra a Figura 31. Esse hardware está associado a uma biblioteca de software chamada NUI API (*Natural User Application Programming Interface*) que recupera informações dos sensores e controla o dispositivo Kinect (MICROSOFT, 2011).



Figura 31: Microsoft Kinect (MICROSOFT, 2012).

Andersen et al. (2012) apresentaram em seu trabalho “*Kinect Depth Sensor Evaluation for Computer Vision Applications*” uma visão geral do funcionamento do sensor: O princípio básico, segundo os autores, por trás do sensor de profundidade do Kinect é a emissão de um padrão de infra-vermelho e a captura simultânea da imagem desse infra-vermelho com uma câmera tradicional equipada com um filtro, que permite capturar o infra-vermelho e bloquear outras formas de onda. O processador de imagens do Kinect usa as posições relativas dos pontos no padrão do infra-vermelho para calcular o deslocamento da profundidade em cada posição de pixel na imagem (ANDERSEN et al., 2012). Segundo Microsoft (2011), cada pixel do mapa de profundidade representa a distância cartesiana, em milímetros, do plano da câmera até o objeto mais próximo, naquela coordenada  $(x,y)$  em particular, como mostra a Figura 32. Se o valor do pixel é 0, isso indica que o sensor não encontrou nenhum objeto no seu espaço de alcance naquela localização  $(x,y)$ . Essa projeção, segundo os autores, é mencionado

como espaço de profundidade.

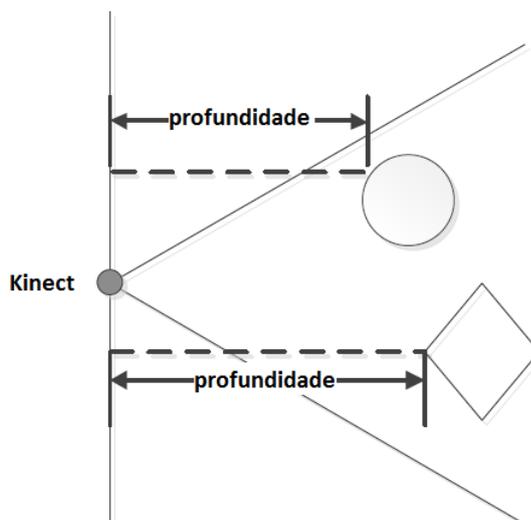


Figura 32: Profundidade no Kinect (MICROSOFT, 2011).

Segundo Andersen et al. (2012), os valores correntes de profundidade são as distâncias do plano da câmera, ao invés das do sensor propriamente dito, como mostra a Figura 32. A Subseção 2.5.2 apresenta uma visão geral das especificações de hardware do Kinect e a Subseção 2.5.3 descreve a NUI API, que proporciona utilizar os recursos do sensor em aplicações que utilizam informações das coordenadas  $x$ ,  $y$  e  $z$  de objetos 3D.

### 2.5.2 PrimeSense PS1080 System on a Chip

O PrimeSense PS1080 SoC é um sistema multi-sensor, que provê, de maneira síncrona, imagem de profundidade, imagem colorida e *stream* de áudio (PRIMESENSE, 2012). Possui, também, uma interface USB 2.0, responsável pela transmissão de dados para o hardware hospedeiro (computador ou video-game XBOX 360). Contudo, o PS1080 não depende da CPU do hardware hospedeiro, já que todos os algoritmos de aquisição de profundidade são executados no próprio PS1080, como mostra a Figura 33.

### 2.5.3 Natural User Interface Application Programmable Interface - NUI API

A NUI API é a principal componente da API do Kinect. É a NUI API que proporciona acesso aos sensores Kinect conectados ao computador, ao fluxo de dados (*streams*) de imagem e profundidade provenientes desses sensores e o processamento de informação de imagem e profundidade, resultado em um conjunto de pontos que compõem um modelo de esqueleto (MICROSOFT, 2011). O *stream* de dados fornecido pela NUI é entregue em forma de sucessivos frames de imagens estáticas.

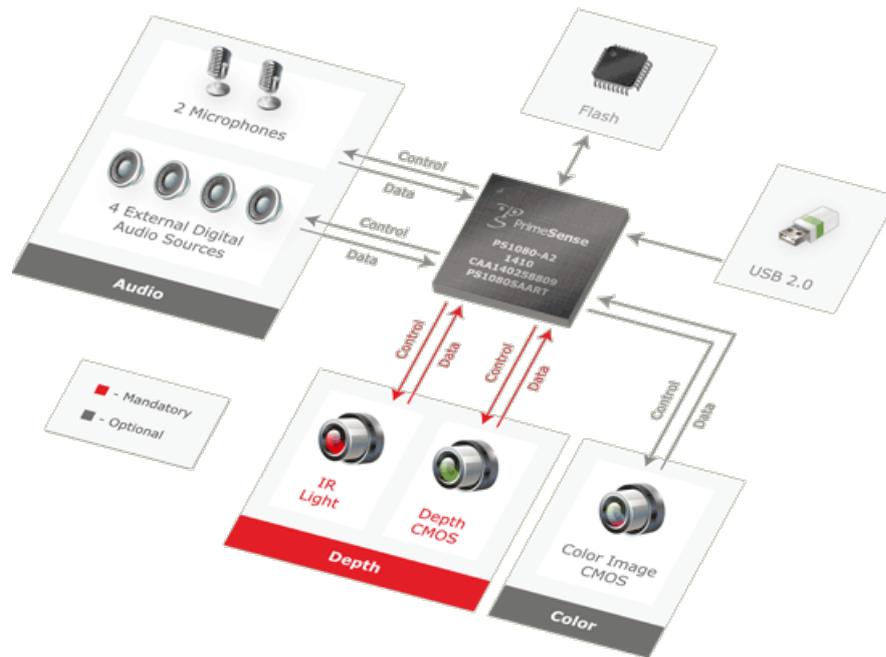


Figura 33: Arquitetura simplificada do PrimeSense PS1080 SoC, principal componente do Microsoft Kinect (PRIMESENSE, 2012).

Quando a NUI API é inicializada pela aplicação, são identificados os *streams* que serão utilizados pela mesma, além de identificar a resolução, o tipo de imagem e o número de buffers que serão utilizados (MICROSOFT, 2011).

#### 2.5.4 NUI Skeleton Tracking - API para rastreamento do esqueleto

A NUI *Skeleton* API é responsável por fornecer a localização, posição detalhada e orientação de até dois indivíduos em frente ao sensor Kinect. Essa informação é fornecida à aplicação como um conjunto de pontos apresentados na Figura 34, chamados de pontos do esqueleto (*skeleton points*). Esses pontos compõem o esqueleto apresentado na Figura 35, e representa a posição corrente da pessoa e sua pose em frente ao sensor (MICROSOFT, 2011).

A aplicação obtém o quadro mais recente de pontos do esqueleto, ou o esqueleto propriamente dito, da mesma maneira que obtém o último quadro de vídeo: a partir de um método que o recupera, e retorna um buffer de dados. Esse buffer contém um conjunto de estruturas de dados de esqueletos, um para cada indivíduo (esqueleto) que está sendo rastreado pelo sistema (MICROSOFT, 2011). Nem todos os indivíduos rastreados possuem dados do esqueleto, alguns são passivamente rastreados, contendo apenas o valor da posição para fins de rastreamento.

Para todos os esqueletos rastreados retornados no quadro, são fornecidos (i) o estado atual do rastreamento: *skeleton-tracked* para esqueletos ativos e *position-only*, para esqueletos que só contém informações sobre rastreamento; (ii) uma ID única de rastreamento, que se mantém atribuída ao indivíduo rastreado, enquanto o mesmo

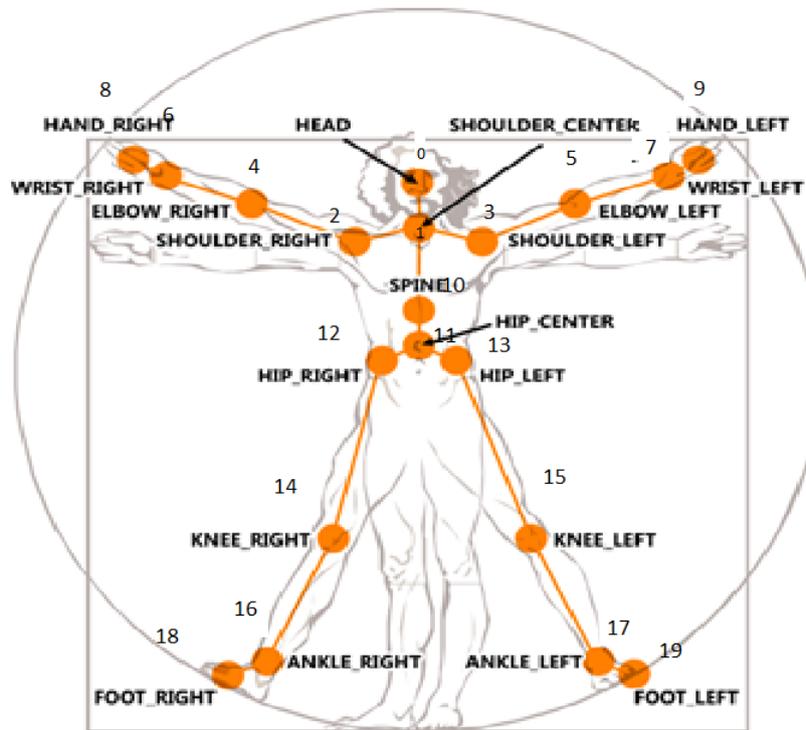


Figura 34: Pontos do esqueleto relativos as articulações do corpo humano (MICROSOFT, 2011).

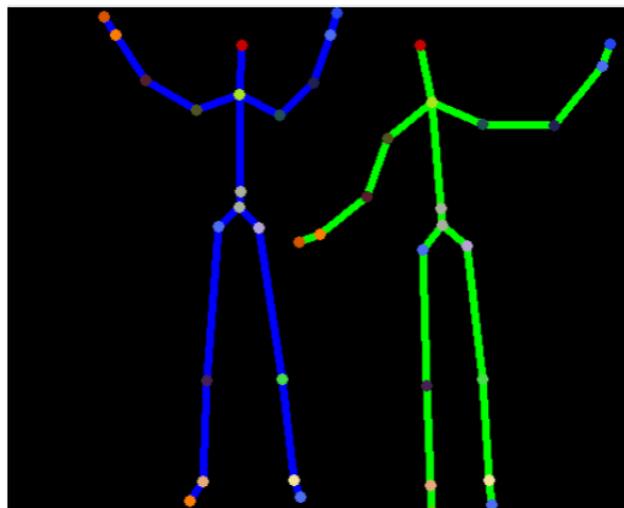


Figura 35: Figuras de esqueleto construídas a partir dos pontos retornados pela NUI Skeleton API (MICROSOFT, 2011).

se locomove em frente ao campo de visão do sensor Kinect; (iii) uma posição que indica o centro de massa do indivíduo rastreado, que é formada por um vetor com 4 dimensões. Esse vetor é a única informação de posicionamento disponível para indivíduos passivamente rastreados. E (iv), para indivíduos rastreados ativamente, um conjunto com todos os dados do esqueleto obtido (MICROSOFT, 2011).

### 2.5.5 Espaço tridimensional do esqueleto

Segundo Microsoft, (2011), as posições do esqueleto são expressas em coordenadas x, y e z, em metros. O eixos X, Y e Z pertencem ao sensor de profundidade do Kinect, e valem-se da regra da mão direita, como mostra a Figura 36. Considerando como referência a direção do sensor, os valores positivos do eixo Z estendem-se na direção que o sensor aponta, os valores positivos do eixo Y estendem-se acima do sensor e os valores positivos do eixo X estendem-se à esquerda do sensor. Os autores utilizam o termo *Skeleton space*, ou espaço do esqueleto, para referenciar esse conjunto de coordenadas.

A posição de localização do sensor Kinect afeta as imagens que as câmeras capturam. Por exemplo, o sensor pode estar localizado em uma superfície não paralela ao solo, ou estar verticalmente rotacionado a fim de otimizar o campo de visão. Quando isto acontece, o eixo Y do espaço tridimensional do esqueleto não é paralelo ao eixo de gravidade (perpendicular ao solo). Os indivíduos, quando rastreados nessas condições, parecem estar inclinados (MICROSOFT, 2011).

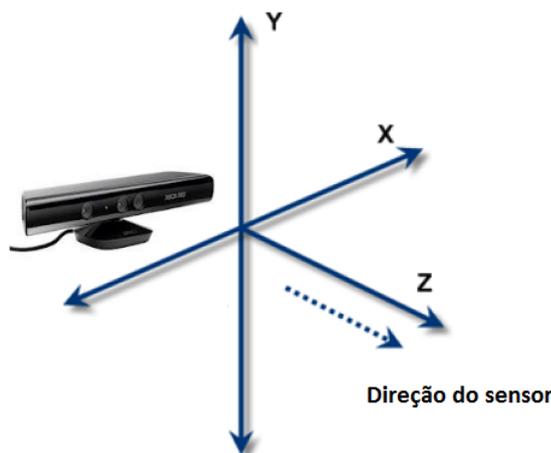


Figura 36: Espaço tridimensional do esqueleto (MICROSOFT, 2011).

## 3 OBJETIVOS E METODOLOGIA

### 3.1 Objetivos

O objetivo do presente trabalho é fornecer resultados sobre a viabilidade do uso do sensor Microsoft Kinect no fornecimento de um modelo ou figura de movimento para a biometria por antropometria e caminhar humano, utilizando um banco de dados de indivíduos caminhando capturado com o sensor Kinect.

Como objetivos específicos é verificado (i) o uso da base de dados com metodologias de outros autores baseada em modelo e livre de modelo; (ii) o impacto da aplicação do método proposto em grupos com diferentes tamanhos; (iii) a influência do número de quadros e ciclos de caminhar capturados no resultado de instâncias classificadas corretamente e (iv) comparação dos resultados do método proposto com resultados obtidos em trabalhos estado-da-arte.

### 3.2 Metodologia

#### 3.2.1 Uso do sensor Kinect na análise do caminhar humano

No Capítulo 2 foram vistas as principais especificações técnicas do Microsoft Kinect, uma visão geral da NUI API, e a NUI *Skeleton Tracking*, a API para rastreamento das principais articulações humanas. Também no Capítulo 2 foram apresentadas as abordagens baseada em modelo e livre de modelo para extração de características do caminhar a partir de silhuetas segmentadas em vídeos e sequências de imagens.

O *pipeline* da análise do caminhar humano, independente da abordagem utilizada para extração de características, apresenta as etapas de captura de vídeo do indivíduo caminhando, processamento dos frames do vídeo, ou sequências, para extração da silhueta do corpo do indivíduo em movimento do resto da imagem. A partir dessa etapa, dependendo da abordagem escolhida, pode-se estimar posições das articulações, ou “esqueletizar” a silhueta no frame.

Essas etapas podem ser simplificadas significativamente com a utilização do sensor Kinect, como mostra a Figura 37. Se a abordagem escolhida for baseada em

modelo, pode-se utilizar os pontos fornecidos pela NUI *Skeleton* API, por exemplo. Contudo, se a abordagem escolhida for livre de modelo, pode-se utilizar a stream de profundidade fornecida pelo sensor para gerar os GEVs ou DGHEIs, explicados no Capítulo 2. O trabalho de Andersen et al. (2012) apresenta uma investigação sobre a precisão do sensor Kinect em relação a resolução 3D e profundidade, que não é o foco deste trabalho.



Figura 37: Pipeline da análise do caminhar humano com o uso do Kinect.

A metodologia proposta é descrita na arquitetura apresentada na Figura 38. Ela descreve as etapas necessárias para captura e armazenamento de indivíduos, filtragem e suavização de dados do esqueleto, extração de atributos antropométricos e do caminhar humano - espaço-temporais e cinemáticos - e finalmente o construção de vetores de exemplos que irão compor os *datasets*, que serão submetidos a algoritmos de aprendizado de máquina tentando a classificação dos indivíduos.

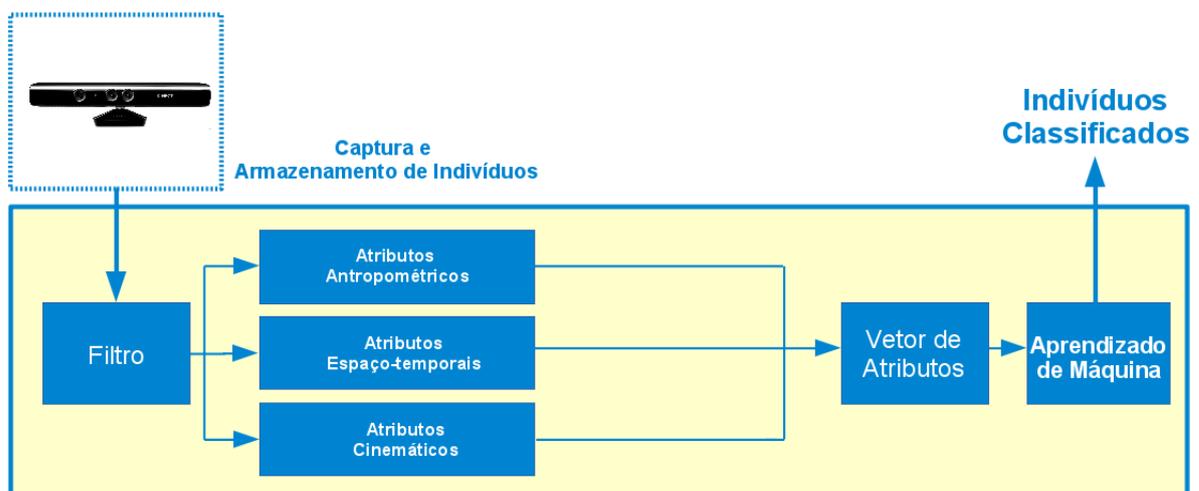


Figura 38: Arquitetura da metodologia proposta.

### 3.2.2 Aquisição de Dados de Indivíduos

Com o intuito de capturar o movimento do caminhar humano, voluntários caminharam em frente ao sensor Kinect enquanto a posição de suas principais articulações no espaço eram armazenadas. Estudantes, servidores e professores da Universidade Federal de Pelotas, a maioria desses dos cursos de Ciência e Engenharia da Computação, foram convidados a participar do experimento, totalizando 164 indivíduos capturados. Durante o experimento, todas as pessoas forneceram altura e gênero, totalizando 113 homens e 51 mulheres, como mostra a Figura 39. Desse total, 112 indivíduos também forneceram idade e peso.

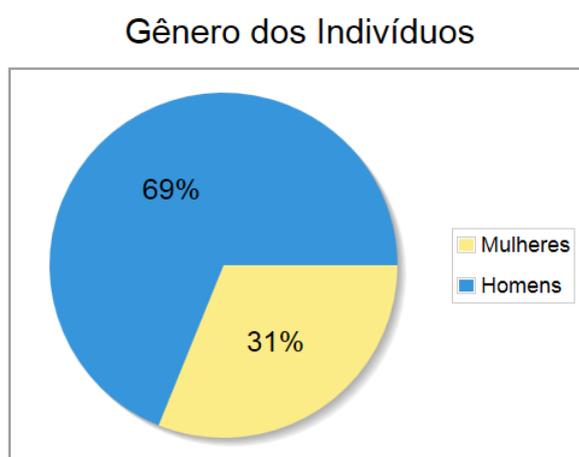


Figura 39: Distribuição, em porcentagem, do gênero dos indivíduos

A distribuição dos indivíduos capturados por altura, peso e idade pode ser visualizada nas Figuras 40, 41 e 42 respectivamente. Na Figura 40, percebe-se uma maior concentração de pessoas com alturas entre 1.7m e 1.85m, com a menor altura registrada igual 1.5m e a maior 2m. Na Figura 41, dos 112 indivíduos que forneceram o valor de massa corporal (peso), a maior concentração de pessoas apresenta pesos de 60kg até 90kg, com o menor peso registrado igual a 45kg e o maior 120Kg. O índice de massa corporal (IMC) dessas pessoas foi calculado através da Equação 25 (WHO, 2014), onde a massa corporal é dividida pelo quadrado da altura, e sua distribuição por indivíduo é apresentada na Figura 43. Como é ilustrado, 6 pessoas apresentam magreza, 73 são saudáveis, 26 apresentam sobrepeso e 7 são obesos.

$$IMC = \frac{massa(kg)}{(altura(m))^2} \quad (25)$$

Na Figura 42, a concentração se dá com idades entre 20 e 23 anos, sendo a menor registrada igual a 17 anos e a maior idade igual a 45 anos. Todos os indivíduos capturados eram aparentemente saudáveis e usavam roupas leves de verão. Saias longas não foram permitidas por problemas de segmentação do sensor Kinect. A maioria dos

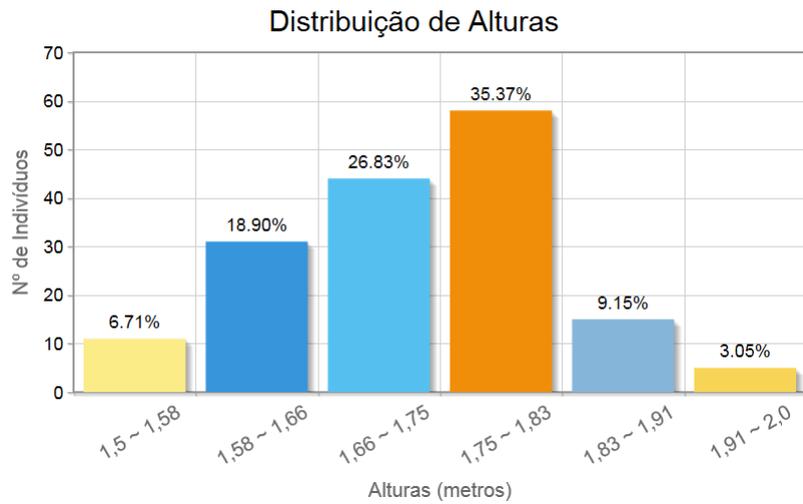


Figura 40: Distribuição de alturas por indivíduo

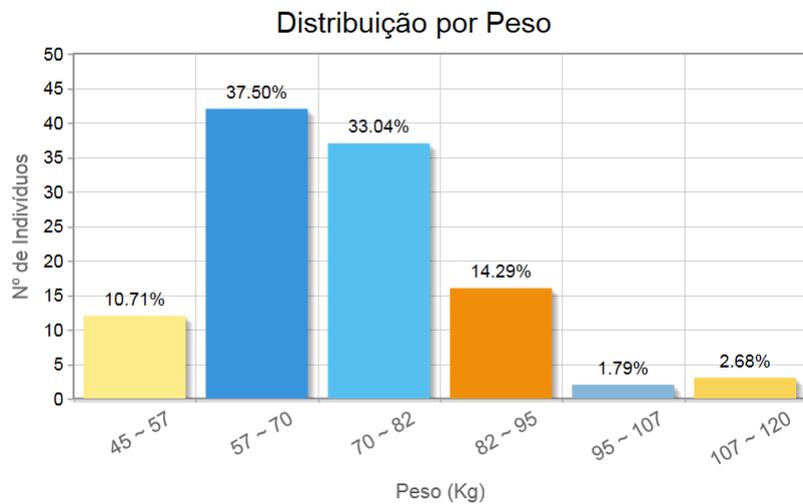


Figura 41: Distribuição de peso por indivíduo

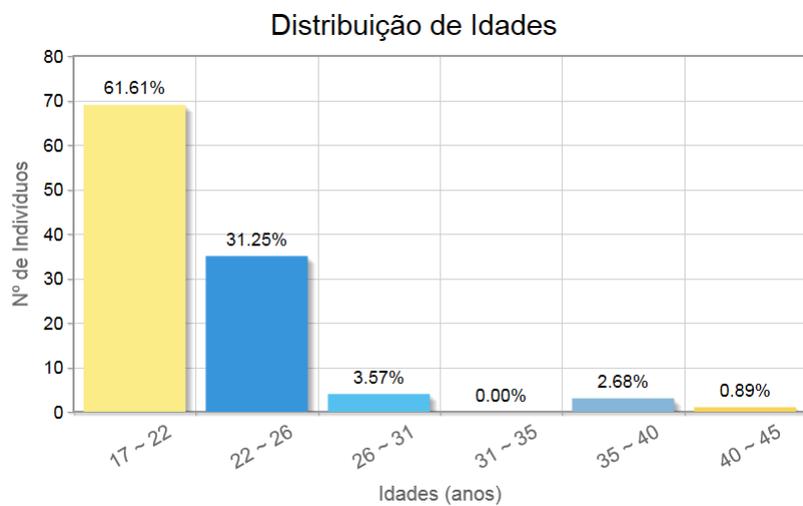


Figura 42: Distribuição de idade por indivíduo

sapatos usados eram esportivos ou sociais sem salto. Das mulheres, apenas 2 usaram sapato com salto alto. Do total de indivíduos, 11 caminharam com acessórios do tipo mochila ou bolsa. O nome fornecido pelo indivíduo durante o experimento foi utilizado apenas para fins de rotulagem. Não foi requisitado qualquer tipo de identificação real da pessoa, permanecendo esta no anonimato.

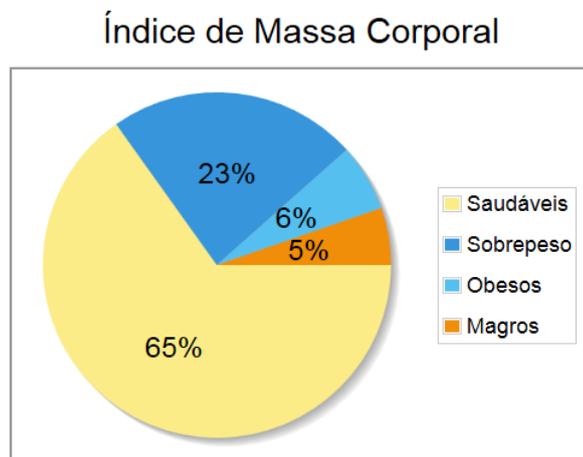


Figura 43: Distribuição de IMC por indivíduo

### 3.2.3 Ambiente de Captura e Trajetória

As principais dificuldades encontradas ao se utilizar o sensor Kinect na captura de indivíduos caminhando, já que o mesmo dispositivo não foi feito para tal propósito, foram: (i) encontrar uma trajetória que possibilitasse capturar o maior número de ciclos de caminhar por caminhada sem interrupção; (ii) ajustar a iluminação, onde houvesse o mínimo de interferência ao infra-vermelho e (iv) encontrar um espaço suficiente para locomoção sem oclusão de objetos.

Observando essas limitações, as capturas foram executadas em salas de aula vazias com em média 3,5 x 11 metros, com iluminação natural indireta reduzida (cortinas nas janelas) e quando necessário, com iluminação artificial (lâmpadas fluorescentes de iluminação local). As capturas foram realizadas durante o dia e no verão. Em geral, as roupas dos indivíduos não atrapalharam as capturas, mas algumas mulheres que utilizavam saia longa não puderam participar do experimento, devido à problemas na obtenção dos pontos das articulações inferiores.

Com o auxílio de um disco giratório, exibido na Figura 44, foi possível capturar os indivíduos em uma trajetória semi-circular, nos sentidos horário e anti-horário, como mostra a Figura 45. A adição do movimento panorâmico na câmera do sensor Kinect possibilita capturar mais quadros de um indivíduo durante sua caminhada. Essa trajetória foi escolhida entre outros dois tipos de trajetória, paralela e perpendicular, por permitir capturar mais ciclos de caminhar sem a utilização de dispositivos como

esteiras, mantendo a pessoa por mais tempo no campo de visão do sensor.



Figura 44: Microsoft Kinect com um disco giratório (prato rotativo) adicionando ao sensor o movimento panorâmico não presente no sensor.

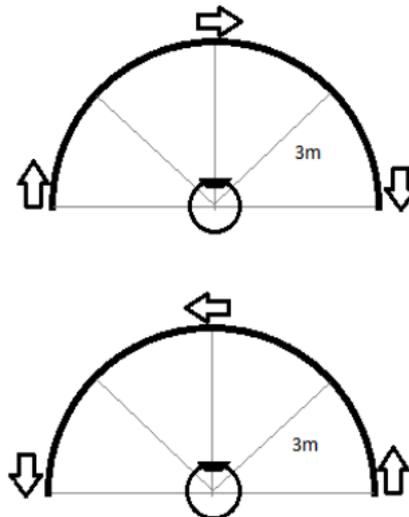


Figura 45: Trajetória semi-circular mostrando os sentidos horário e anti-horário percorridos pelo indivíduo.

A Figura 46 apresenta o *setup* final para a captura da caminhada dos indivíduos. Marcas no chão indicam os limites que os indivíduos deveriam utilizar referente à trajetória estabelecida na Figura 45. Foi solicitado aos mesmos que caminhassem livremente seguindo a trajetória, nos sentidos horário e anti-horário. Esse procedimento foi repetido, em média, 5 vezes para cada pessoa. A maioria dos indivíduos possui 5 caminhadas completas, mas alguns possuem menos ou mais devido a problemas ocasionais durante o processo de captura. A cada caminhada completa, o software

de captura salvava as informações das articulações em arquivos, populando a base de dados utilizada no presente projeto.



Figura 46: *Setup* para captura do caminhar apresentando o sensor Kinect e marcas no chão para guiar as pessoas.

### 3.2.4 Base de Dados do Caminhar

A estrutura da base de dados é ilustrada na Figura 47. Cada caminhada foi armazenada em um arquivo ASCII contendo quadros, que são compostos de pontos do esqueleto capturado naquele instante. Cada articulação é referenciada por seu nome, iniciando com a “cabeça” e terminando com o “pe-esquerdo”. Cada quadro é posicionado sequencialmente no arquivo, como mostra a Figura 47. Além de informações do esqueleto, as informações de gênero, altura, peso e idade também estão presentes em arquivos separados.

A base de dados com as caminhadas capturadas pode ser acessada em: <http://ricardoaraujo.net/kinect/>.

### 3.2.5 Filtragem de Indivíduos

A técnica livre de modelo Caixa delimitadora ou *Bundle Rectangle*, proposta por Guillen, Padilla e Hernandez (2009) para extração de parâmetros do caminhar humano, apresentada no Capítulo 3, possui recursos úteis para a análise dos quadros capturados durante o caminhar dos indivíduos. Como os parâmetros extraídos do *Bundle Rectangle* são média e desvio padrão da altura e largura do retângulo, ao aplicar a técnica nas 164 pessoas em todas as caminhadas, observou-se que algumas caminhadas apresentavam valores de desvio padrão de altura e largura altos (valores maiores ou próximos a 1.0). Algumas dessas caminhadas foram investigadas com a reconstrução do esqueleto em vídeo e observou-se que aquelas caminhadas apre-

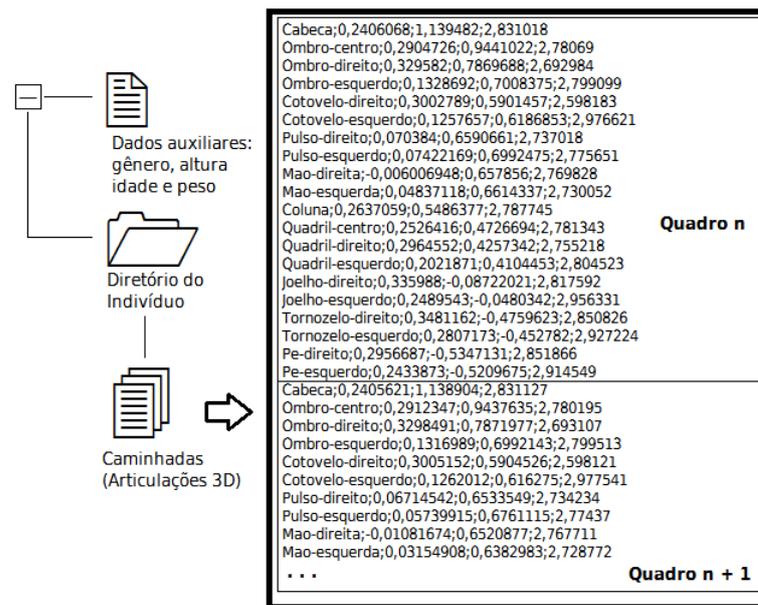


Figura 47: *Snapshot* da base de dados: Todos os indivíduos capturados estão em um arquivo contendo gênero, altura, peso e cada diretório do indivíduo contém arquivos de caminhadas, contendo pontos 3D de todas as articulações humanas rastreadas. Os quadros são arranjados sequencialmente com um rótulo e coordenadas x, y e z em cada linha.

sentavam problemas de captura. Indivíduos que apresentavam 2 ou mais caminhadas nessas condições foram excluídos da lista de análise, resultando em 24 exclusões, restando 140 pessoas para a análise.

### 3.2.6 Suavização das Articulações

Os dados do esqueleto na sua forma original, não tratados, possuem ruído nas posições das articulações retornadas pela NUI *Skeleton API* (AZIMI, 2012). Segundo (AZIMI, 2012), antes do uso das informações retornadas pela API, é necessária a aplicação de um filtro de redução de ruído nas articulações a fim de remover a maior parte desse. Ainda, segundo o autor, esses filtros são chamados filtros de suavização de articulações pois sua aplicação em sequência resulta em posições de articulações mais suaves.

Para filtrar o ruído das articulações, foram implementados dois filtros: (i) Média móvel central auto regressiva ou *Auto regressive central moving average* (ARMA) e (ii) Média móvel exponencialmente ponderada, ou *Exponential weighted moving average* (EWMA) (AZIMI, 2012). Em (i), a articulação ou posição filtrada é resultado da média ponderada das  $M$  entradas passadas e  $N$  entradas futuras segundo a Equação 26:

$$\hat{X}_n = \sum_{i=-M}^N a_i X_{n-i} \quad (26)$$

onde  $\hat{X}_n$  é o resultado da média dos  $M$  valores passados e  $N$  valores futuros de  $X$ , ponderada pelo coeficiente  $a_i$ , como mostra a Equação 27.

$$a_i = \frac{1}{M + N} \quad (27)$$

Em (ii), a saída do filtro exponencial no instante  $n$  é uma média ponderada de todas as entradas passadas dada pela Equação 28:

$$\hat{X}_n = \alpha \sum_{i=0}^n (1 - \alpha)^i X_{n-i} \quad (28)$$

onde os pesos são dados pelo coeficiente  $a_i$  seguindo a Equação 29 que diminui exponencialmente com o tempo segundo o fator de amortecimento (*dampening*)  $\alpha$ .

$$a_i = \alpha(1 - \alpha)^i \quad (29)$$

Segundo (AZIMI, 2012), quanto maior a janela de quadros  $M$  e  $N$ , mais suavizado será o valor resultante para filtros ARMA. Valores entre 5 e 10 foram testados para o filtro (i), e os resultados foram reconstruídos visualmente e comparados com os valores originais. Foram avaliadas na reconstrução das caminhadas, principalmente, aspectos na fluidez do movimento das articulações inferiores. Dos valores testados, foram escolhidos 16 quadros ( $M = 8$  e  $N = 8$ ) com o coeficiente  $a_i = 0,0625$ , por apresentarem melhor resultados visuais na reprodução da caminhada.

No filtro (ii), o valor resultante no instante  $n$  é uma média ponderada de  $n$  entradas passadas. Foram testados de 8 a 30 quadros passados, e escolhidos 30 quadros ( $n = 30$ ) por apresentarem, de maneira semelhante ao filtro ARMA, melhores resultados na reprodução visual da caminhada, com o fator de amortecimento  $\alpha = 0.2$ .

Foi escolhido o filtro (i) ARMA após uma avaliação subjetiva entre os dois filtros utilizando a mesma técnica de visualização das reconstruções das caminhadas. O filtro ARMA com a configuração escolhida foi aplicado, então, em todas as caminhadas para todos os indivíduos. Na Figura 48 são apresentados os ângulos do joelho direito de um indivíduo sem a aplicação do filtro e os valores de ângulos resultantes com a aplicação do filtro ARMA.

### 3.3 Análise Antropométrica

Para cada quadro capturado pelo sensor Kinect, as medidas de todos os segmentos do corpo formados pelas articulações rastreadas, descrito na Figura 49 foram calculadas usando a Equação 30 ou distância Euclidiana, onde  $A$  e  $B$  são articulações que possuem 3 dimensões ( $x, y, z$ ).

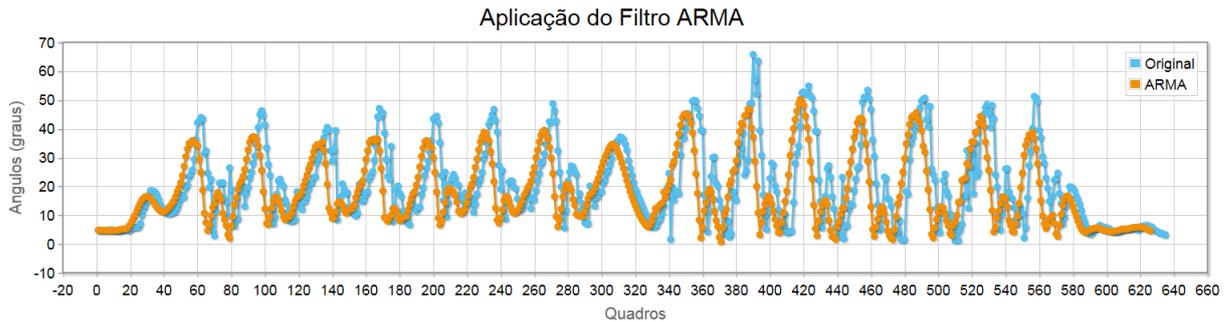


Figura 48: Aplicação do filtro ARMA nos ângulos obtidos do joelho direito de um indivíduo durante 1 caminhada capturada, realizando a trajetória semi-circular nos sentidos de ida e de volta.

$$segmento_{A,B} = \sqrt{\sum_{i=1}^3 (A_i - B_i)^2} \quad (30)$$

A média e desvio padrão de cada medida de segmento entre todos os quadros capturados foram calculadas com o propósito de caracterizar o indivíduo pela antropometria. Medidas aquém e além de dois desvios padrão da média, para menos e para mais, foram descartadas e atribuídas como ruído. Esse tipo de filtro secundário é necessário devido a outros ruídos não naturais aos dados do esqueleto, como interferência no infra-vermelho e auto-occlusão, produzindo informações inexatas.

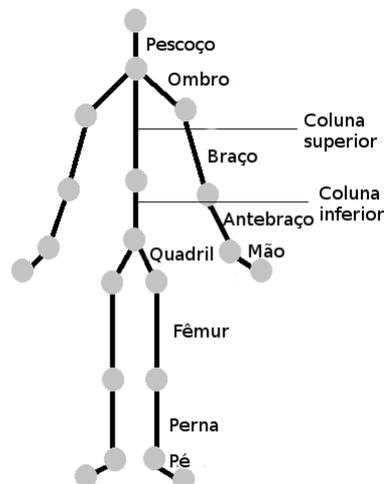


Figura 49: Medidas dos segmentos do corpo calculadas a partir das articulações, ilustradas como círculos, fornecidas pela API do Kinect *NUI Skeleton*

Embora o indivíduo tenha informado a sua altura durante o experimento, a altura obtida pelo sensor Kinect foi definida como a soma dos segmentos do pescoço, coluna superior e inferior, e médias da soma entre quadris, fêmur e perna, dos lados esquerdo e direito. Na Figura 50 é possível observar que a altura calculada usando

as articulações rastreadas pelo sensor Kinect possui em alguns casos uma diferença para menos em relação a altura fornecida, e em outros casos a altura é muito próxima a fornecida.

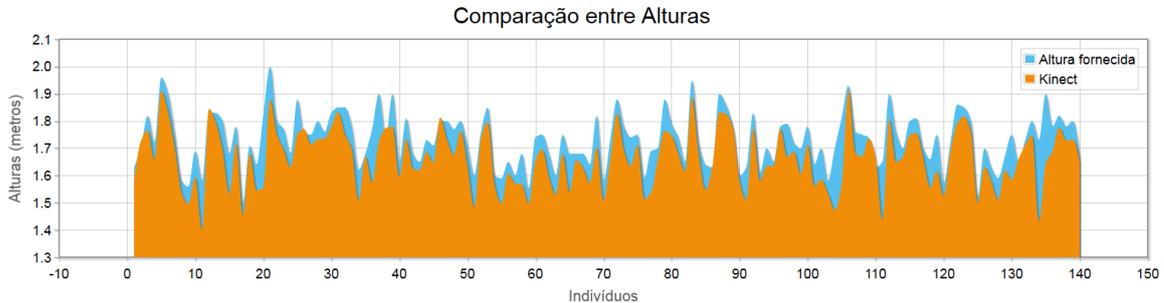


Figura 50: Diferença entre alturas fornecidas durante o experimento e alturas calculadas pelo Kinect.

### 3.4 Análise de Parâmetros Espaço-temporal e Cinemáticos do Caminhar Humano

Dos parâmetros cinemáticos do caminhar humano, a partir da abordagem baseada em modelo, foram extraídos os ângulos descritos pelas articulações inferiores do corpo humano, de cada quadro capturado pelo sensor Kinect, usando o modelo de pêndulo proposto por Cunado, Nixon e Carter (2003). Além das articulações dos quadris, joelhos e tornozelos, foi utilizado também como parâmetro cinemático, o ângulo da abertura do pé, descrito por Murray, Drought e Kory (1964).

O modelo de pêndulo adaptado para este trabalho pode ser visualizado na Figura 51. Os segmentos inferiores do corpo humano estão destacados na Figura 51(a), mostrando a localização da extração dos parâmetros cinemáticos utilizando o modelo. Os ângulos são calculados, em cada quadro capturado, utilizando três pontos: duas articulações que formam o segmento  $A_1$  e  $A_2$ , e um terceiro ponto imaginário  $p_3$ , composto por projeções das coordenadas  $x$ ,  $y$  e  $z$  das duas articulações, formando um triângulo retângulo, como mostra a Figura 51(b).

As coordenadas utilizadas pelos pontos imaginários dependem das articulações que junto com ele formam o triângulo retângulo e a partir desse é possível extrair os ângulos formados pelas articulações com seus segmentos adjacentes utilizando a Equação 31.

$$angulo = sen^{-1} \left( \frac{A_2 - p_3}{A_2 - A_1} \right) \quad (31)$$

Como mostra a Figura 51(b), o ângulo  $\omega$  é formado durante o ciclo de caminhar como uma rotação do quadril, assim como o ângulo  $\gamma$  descreve a rotação do joelho e

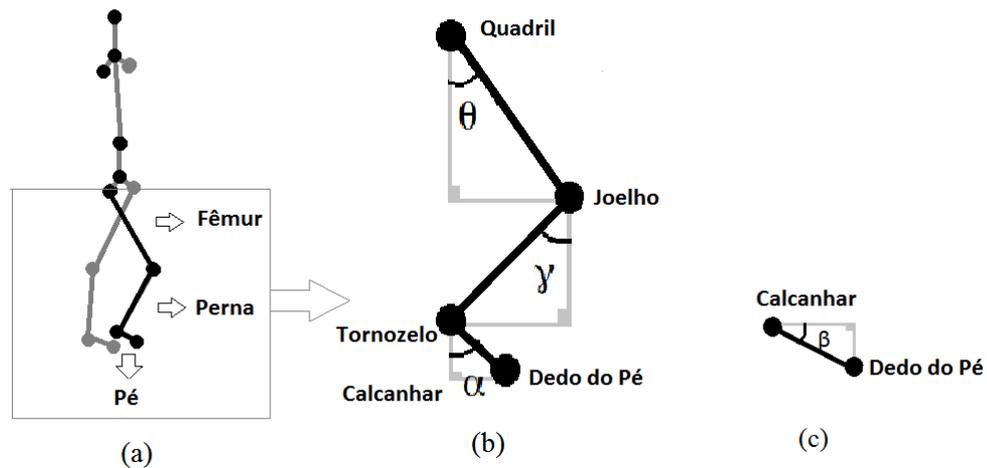


Figura 51: Em (a) é mostrado os segmentos inferiores do corpo utilizados para extrair os parâmetros cinemáticos, (b) os ângulos  $\omega$ ,  $\gamma$  e  $\alpha$  formados entre os segmentos segundo o modelo de pêndulo de (CUNADO; NIXON; CARTER, 2003) e (c) o ângulo  $\beta$  de abertura do pé.

$\alpha$  a rotação do tornozelo. O ângulo  $\beta$  define a rotação de abertura do pé com eixo no calcânhar. Esses ângulos descrevem curvas periódicas durante o caminhar que possuem características úteis para o reconhecimento biométrico (MURRAY; DROUGHT; KORY, 1964).

As curvas periódicas são formadas pelos dois sentidos executados durante a caminhada, com a presença de uma região central composta de picos fora do padrão que indicam a mudança de sentido da trajetória. Essas curvas são compostas por fases de flexão e extensão, visualmente notáveis por picos (flexão) e vales (extensão) (MURRAY; DROUGHT; KORY, 1964), como mostra a Figura 52 e foram usadas neste presente trabalho para caracterizar os parâmetros cinemáticos de cada indivíduo. Cada pico de flexão possui uma fase alta e uma fase baixa, assim como os vales de extensão. Esses altos e baixos foram tratados separadamente, gerando para cada fase alta e baixa uma média aritmética e um desvio padrão, de todo o caminhar capturado.

Os picos e vales no conjunto de ângulos foram encontrados calculando as diferenças de primeira ordem entre os ângulos, como mostra a Equação 32:

$$diff(n) = \text{angulos}(n + 1) - \text{angulos}(n) \quad (32)$$

onde  $diff$  é o conjunto com o resultado de todas as diferenças calculadas. Percorrendo  $diff$ , se  $diff(n)$  é maior que  $\theta$  e a próxima diferença  $diff(n + 1)$  é menor que  $\theta$  e a atual  $diff(n)$  é diferente da próxima  $diff(n + 1)$ , essa diferença atual pode ser um “pico”. A posição será um pico se a diferença entre seus valores originais for maior que 2.

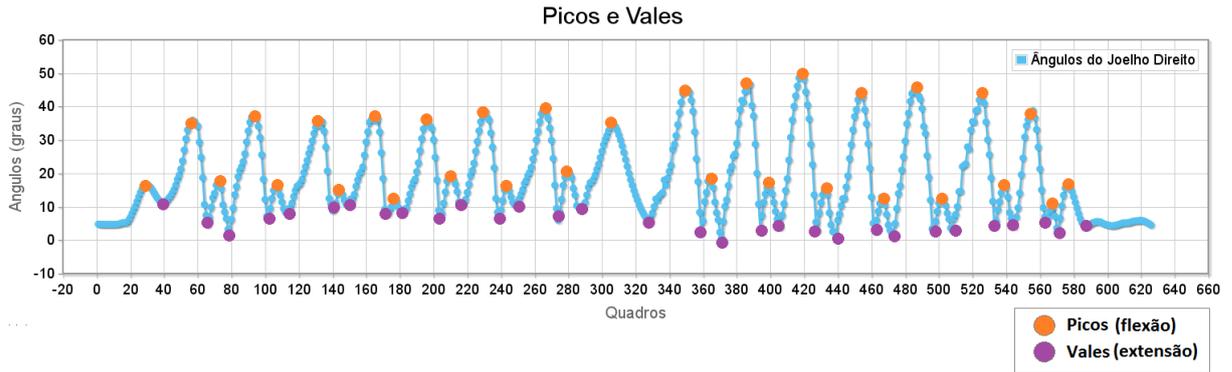


Figura 52: Fases de extensão (vales) e flexão (picos) em uma sequência de valores de ângulos do joelho direito descritos durante uma caminhada capturada.

O mesmo princípio é utilizado para encontrar vales: se  $diff(n)$  é menor que  $\theta$  e a próxima diferença  $diff(n+1)$  é maior que  $\theta$  e a atual  $diff(n)$  é diferente da próxima  $diff(n+1)$ , essa diferença atual pode ser um "vale". Ela, a posição, será um vale se a diferença entre seus valores originais for maior que 2. No final, as posições dos picos e vales são retornadas pelo algoritmo.

Para cada quadro capturado também foram extraídos parâmetros espaço-temporais, descritos em (YOO; NIXON, 2011): tamanho do passo, tamanho da passada (ou tamanho do ciclo de caminhar), duração do ciclo e velocidade. Os parâmetros espaço-temporais foram calculados com base no tamanho do passo do indivíduo e a taxa de quadros por segundo capturados pelo sensor Kinect, igual a 30fps (*frames per second*).

Na análise do passo foi calculada, para cada quadro capturado, a diferença entre as posições dos calcanhares esquerdo e direito do indivíduo durante toda a caminhada armazenada. Essas posições podem ser visualizadas na curva apresentada na Figura 53 que mostra os menores (vales) e maiores (picos) valores entre as diferenças.

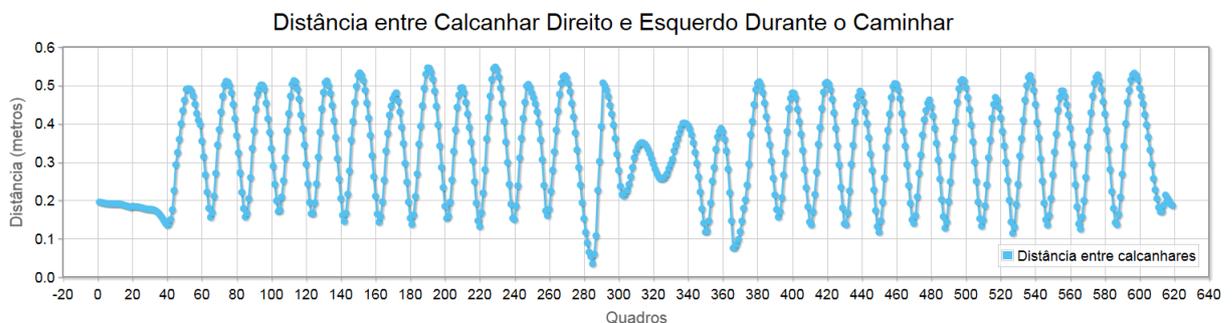


Figura 53: Diferença entre as posições dos calcanhares direito e esquerdo: A média aritmética entre os valores mais altos (picos) resulta no tamanho do passo médio durante toda a caminhada desenvolvida.

Utilizando a mesma técnica de localização de picos e vales para extrair atributos

cinemáticos nas curvas de ângulos, foram obtidos todos os valores de picos durante o caminhar e, em consequência disso, o tamanho do passo do indivíduo foi obtido através da média aritmética entre esses valores, como mostra a Equação 33:

$$tamanhoPasso = \frac{\sum_{i=0}^n picos_i}{n} \quad (33)$$

onde  $n$  são todas as maiores diferenças entre os calcanhares dadas por  $picos$ . Outro atributo espaço-temporal é o tamanho da passada, ou tamanho do ciclo de caminhar, que é dado pela espaço entre toques de um mesmo calcanhar no chão, conforme visto no Capítulo 2. O tamanho da passada  $tamanhoCiclo$  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 então calculado através da média aritmética de cada valor, pela Equação 34:

$$tamanhoCicloMedio = \frac{\sum_{i=0}^n tamanhoCiclo_i}{n} \quad (34)$$

A duração desse ciclo e a velocidade são calculados utilizando a taxa média de captura de quadros do sensor Kinect, conforme elucidado anteriormente, de 30 quadros por segundo. A duração e a velocidade são dados, então, pelas Equações 35 e 36:

$$duracaoCiclo = \frac{numQuadrosCicloMedio(quadros)}{30(quadros/s)} \quad (35)$$

$$velocidade = \frac{tamanhoCicloMedio}{duracaoCiclo} \quad (36)$$

onde  $numQuadrosCicloMedio$  é o número de quadros, em média, que um ciclo de caminhar possui e  $30quadros/s$  a taxa de amostragem.

### 3.5 Composição dos *Datasets*

Cada conjunto de atributo foi agrupado de acordo com o tipo ao qual ele pertence. Na Tabela 1 é apresentado cada grupo, seu nome, os atributos que cada grupo possui e o total de atributos por conjunto. Após a extração de cada atributo das medidas do corpo humano e do caminhar (cinemáticos e espaço-temporais), utilizando o método proposto neste trabalho, foram gerados *Datasets* contendo os atributos escolhidos por tipo: (i) antropometria, (ii) cinemáticos, (iii) espaço-temporais e por combinação: (iv) antropometria e cinemático, (v) antropometria e espaço-temporal e (vi) espaço-temporal e cinemático e (vii) todos os atributos, como mostra a Tabela 2.

Os tipos e as combinações citadas nas Tabelas 1 e 2 originaram, cada um, um *Dataset* composto dos atributos e de um atributo de rótulo identificando os indivíduos.

Tabela 1: Grupos de Atributos

Grupo	Atributos	Total
Antropometria	média aritmética do pescoço, ombro-direito, ombro-esquerdo, braço-direito, braço-esquerdo, antebraço-direito, antebraço-esquerdo, mão-direita, mão-esquerda, coluna-superior, coluna-inferior, quadril-direito, quadril-esquerdo, fêmur-direito, fêmur-esquerdo, perna-direita, perna-esquerda, pé-direito, pé-esquerdo, altura	20
Cinemático	media e desvio padrão dos picos-baixos, picos-altos, vales-baixos e vales-altos dos ângulos do: quadril-direito, quadril-esquerdo, joelho-direito, joelho-esquerdo, tornozelo-direito, tornozelo-esquerdo, abertura-tornozelo-direito e abertura-tornozelo-esquerdo.	56
Espaço-temporal	duração do ciclo, tamanho da passada, tamanho do passo e velocidade.	4

Tabela 2: Composição dos *Datasets*

Combinação	Conjunto de Atributos	Total
(iv)	Antropometria + Cinemático	76
(v)	Antropometria + Espaço-temporal	24
(vi)	Espaço-temporal + Cinemático	60
(vii)	Antropometria + Espaço-temporal + Cinemático	80

Cada caminhada capturada e armazenada gerou um exemplo para cada indivíduo no *Dataset*. Esses *Datasets* são classificáveis por algum algoritmo de aprendizado de máquina escolhido.

No presente trabalho é proposto, como método, o uso do *dataset* com os atributos antropométricos, cinemáticos e espaço-temporais (*dataset* vii) em conjunto. Entretanto, a classificação individual dos mesmos também é explorada, com a finalidade de explorar a capacidade antropométrica e do caminhar humano como forma de biometria exclusiva.

### 3.6 Escolha do Algoritmo de Aprendizado de Máquina

Três algoritmos, *K-nearest neighbor* (k-NN), Rede Neural ou *Multilayer Perceptron* (MLP) e *Support Vector Machines* (SVM) foram avaliados levando em consideração o número de instâncias classificadas corretamente (ICC), em cada caso, como principal critério de avaliação de desempenho. O algoritmo que apresentou resultados satisfatórios na classificação do *Dataset* vii, enumerado na Tabela 2, foi escolhido como principal ferramenta de classificação dos indivíduos e assim aplicado no restante da metodologia apresentada neste trabalho. O processo de busca pelo melhor algoritmo foi automatizado utilizando variações de parâmetros específicos para cada implementação. Como resultado, uma lista de combinações de parâmetros e o resultado obtido na classificação do maior conjunto de indivíduos (140 indivíduos) e o maior conjunto de atributos (combinação vii) foi obtida e o melhor score de ICCs apontou o algoritmo escolhido.

O *workbench* WEKA (*Waikato Environment for Knowledge Analysis*) para aprendizado de máquina da universidade de Waikato (HALL et al., 2009) fornece implementações apresentadas na Tabela 3, dos algoritmos k-NN, MLP e SVM. A seguir são enumerados os principais parâmetros que cada implementação utiliza e que foram iterados para obter a melhor combinação, e quais dessas combinações apresentaram diferenças significantes durante a avaliação.

Tabela 3: Implementações dos algoritmos de Aprendizado de Máquina no WEKA

Algoritmo	Nome	Implementação WEKA
k-NN	IBk	<code>weka.classifiers.lazy.IBk</code>
MLP	MultilayerPerceptron	<code>weka.classifiers.functions.MultilayerPerceptron</code>
SVM	LibSVM	<code>weka.classifiers.functions.LibSVM</code>
SVM	SMO	<code>weka.classifiers.functions.SMO</code>

#### 1. Principais parâmetros para IBk

- (a) O número de vizinhos usados na classificação:  $-K$
- (b) Qual o método de ponderação de distância utilizado:
  - i.  $1 - distance = -F$
  - ii.  $1 / distance = -I$
- (c) Algoritmo *Nearest Neighbour* a ser usado:  $-A$  (“*BallTree*”, “*CoverTree*”, “*KDTree*” ou “*LinearNNSearch*”), com as funções de cálculo de distância (“*ChebyshevDistance*”, “*EuclideanDistance*”, “*ManhattanDistance*” e “*MinkowskiDistance*”).

## 2. Principais parâmetros para MultilayerPerceptron

- (a) O número de épocas para treino:  $-N$
- (b) Camadas ocultas da Rede Neural. Cada número natural fornece o número de nodos nesta camada oculta:  $-H$

## 3. Principais parâmetros para LibSVM

- (a) Tipo de SVM a ser usado: C-SVC =  $-C$ , nu-SVC =  $-N$
- (b) Tipo de Kernel usado:  $-K$  (“*linear*”, “*polinomial*”, “*radial*”, “*sigmoide*”).
- (c) Parâmetros usados:
  - i. O custo  $C$  do SVM C-SVC:  $-C$
  - ii. O parâmetro  $nu$  do SVM nu-SVC:  $-N$
  - iii. O valor do coeficiente do polinômio ou função do Kernel usado:  $-R$
  - iv. Grau do polinômio ou função do Kernel:  $-D$
  - v. Valor de gamma do polinômio ou função do Kernel:  $-G$
- (d) Uso dos parâmetros de acordo com o kernel usado:
  - i. *linear*:  $-C$  ou  $-N$
  - ii. *polinomial*:  $-C$  ou  $-N$ ,  $-G$ ,  $-D$ ,  $-R$
  - iii. *radial*:  $-C$  ou  $-N$ ,  $-G$
  - iv. *sigmoide*:  $-C$  ou  $-N$ ,  $-G$ ,  $-R$

## 4. Principais parâmetros para SMO

- (a) Parâmetro ou constante de complexidade:  $-C$
- (b) Tipo de Kernel usado:  $-K$ : (“*NormalizedPolyKernel*”, “*PolyKernel*”, “*Puk*”, “*RBFKernel*”).
- (c) Parâmetros usados:
  - i. Expoente:  $-E$

- ii. Omega:  $-O$
  - iii. Sigma:  $-S$
  - iv. Gama:  $-G$
- (d) Uso de parâmetros de acordo com o kernel usado:
- i. *NormalizedPolyKernel*:  $-E$
  - ii. *PolyKernel*:  $-E$
  - iii. *Puk*:  $-O, -S$
  - iv. *RBFKernel*:  $-G$

### 3.6.1 Execução e Teste

Considerando cada implementação dos algoritmos utilizados, o âmbito de valores de alguns principais parâmetros foi estabelecido como segue:

#### 1. IBk

- (a)  $K$ : de 1 a 20

considerando que as combinações entre os outros parâmetros suportadas pela implementação são: *BallTree*, *KDTree* e *CoverTree* com *EuclideanDistance* e ponderação de distâncias válidas  $1/distance$ ,  $1 - distance$  e nenhuma, e *LinearNNSearch* com *ChebyshevDistance*, *EuclideanDistance* e *MinkowskiDistance* com ponderação de distâncias suportadas  $1/distance$ ,  $1 - distance$  e nenhuma e *ManhattanDistance* com  $1/distance$  e nenhuma.

#### 2. MultilayerPerceptron

- (a) Épocas  $N$ : de 500 a 2000, intercalando em 100.
- (b) Camadas ocultas  $H$ : de 3 a 48, intercalando em 5.

#### 3. LibSVM - C-SVC

- (a) Custo  $C$ : de 1 a 11.
- (b) Gama  $G$ : de 0 a 1.1, intercalando em 0.1.
- (c) Grau  $D$ : de 0 a 4.
- (d) Coeficiente  $R$ : de 0 a 1.1, intercalando em 0.1.

#### 4. LibSVM - nu-SVC

- (a) nu  $N$ : de 0 a 1.1, intercalando em 0.1.
- (b) Gama  $G$ : de 0 a 1.1, intercalando em 0.1.
- (c) Grau  $D$ : de 0 a 4.

considerando para (3) e (4) o uso de acordo com os parâmetros suportados pelo kernel usado.

## 5. SMO

- (a) Complexidade  $C$ : de 1 a 100, intercalando em 1.
- (b) Kernel  $K$ : *Puk*, *NormalizedPolyKernel*, *RBFKernel*, *PolyKernel*.
- (c) Omega  $O$ : de 1.0 a 10.0, intercalando em 1.
- (d) Sigma  $S$ : de 1.0 a 10.0, intercalando em 1.
- (e) Expoente  $E$ : de 1.0 a 5.0, intercalando em 0.1
- (f) Gama  $G$ : de 0 a 0.5, intercalando em 0.01.

considerando o uso de Omega, Sigma, Expoente e Gama de acordo com os parâmetros suportados pelo kernel usado.

As combinações de parâmetros utilizados formam grupos de teste que disputam com outras combinações em sua “chave” de algoritmo. Para cada algoritmo, a combinação que apresentou o maior número de ICC resultante é escolhido para representar a sua chave na disputa seguinte, contra todos os outros algoritmos. Finalmente, o melhor resultado de ICC geral indica qual desses será usado para os experimentos seguintes.

Após as classificações, cada combinação de cada chave de algoritmo foi agrupada com seu resultado, possibilitando a verificação de seu ranking em relação às outras classificações. A metodologia apresentada avalia o pior caso entre o número de indivíduos, mas não leva em conta o pior caso entre o conjunto de atributos. Utilizando um conjunto diferente de atributos, diferentes combinações de parâmetros ou algoritmos diferentes podem apresentar melhores resultados de ICCs.

## 3.7 Comparação com Trabalhos Relacionados

A metodologia proposta no presente trabalho foi comparada com trabalhos relacionados, reproduzindo outras metodologias na base de dados obtida neste trabalho através das capturas do sensor Kinect e com comparação direta com resultados de trabalhos relacionados estado-da-arte e que utilizam o Kinect na obtenção da caminhada. Esses trabalhos propõem diferentes formas de classificação das pessoas através do caminhar humano, explorando alguns atributos não utilizados neste projeto e utilizam exemplos de caminhar capturados de maneira diferente ao proposto neste trabalho. As metodologias de captura e extração de parâmetros do caminhar foram detalhadas no Capítulo 3.

### 3.7.1 Reprodução de metodologias utilizando a base de dados obtida com o Kinect

#### 3.7.1.1 Caixa delimitadora

A técnica da *Caixa delimitadora* é classificada como livre de modelo e foi proposta primeiramente por Guillen, Padilla, Hernandez et al. em “*Gait Recognition System: Bundle Rectangle Approach*” (2009) e foi descrita no Capítulo 3, Seção 3.2.2, no presente trabalho, e foi revista por Nandini, Mukhopadhyay, Tanmay et al. em “*An Efficient Human Identification Using Gait Analysis*” (2011). A *Caixa delimitadora* é aplicada originalmente em silhuetas de indivíduos caminhando, em cada quadro, extraíndo alguns parâmetros da caixa que acompanha a silhueta. Normalmente, os parâmetros do caminhar extraídos de uma abordagem livre de modelo estão fortemente ligados a técnica de obtenção das silhuetas, tornando difícil sua aplicação quando a técnica utilizada para obter o caminhar é outra. A *Caixa delimitadora*, contudo, não depende diretamente da silhueta e pode ser substituída, por exemplo, por uma Figura de movimento ou *Stick figure*.

Considerando isso, foi aplicada a técnica da *Caixa delimitadora* no modelo de esqueleto obtido pelo sensor Kinect na captura das caminhadas. Em cada quadro capturado, foi construída uma caixa imaginária que possui como divisas os pontos mais externos da figura de movimento do indivíduo capturado, segundo as Equações 37, 38, 39 e 40:

$$b_1 = (\min([x_1, \dots, x_i]_{artic}), \min([y_1, \dots, y_i]_{artic}), z_{cab}) \quad (37)$$

$$b_2 = (\max([x_1, \dots, x_i]_{artic}), \min([y_1, \dots, y_i]_{artic}), z_{cab}) \quad (38)$$

$$b_3 = (\min([x_1, \dots, x_i]_{artic}), \max([y_1, \dots, y_i]_{artic}), z_{cab}) \quad (39)$$

$$b_4 = (\max([x_1, \dots, x_i]_{artic}), \max([y_1, \dots, y_i]_{artic}), z_{cab}) \quad (40)$$

onde  $[x_1, \dots, x_i]_{artic}$  e  $[y_1, \dots, y_i]_{artic}$  são todos os pontos  $x$  e  $y$  no espaço 3D de captura e  $z_{cab}$  é a posição  $z$  referente a cabeça do indivíduo capturado, no modelo de esqueleto.

Para cada quadro capturado, os pontos  $b_1$ ,  $b_2$ ,  $b_3$  e  $b_4$  compõem as divisas para a *Caixa delimitadora* naquele quadro. A partir dela, são extraídos os parâmetros descritos nas Equações 8, 9, 10, 11 e 12 na Seção 3.2.2.

O sistema de autenticação pelo caminhar humano proposto por Guillen, Padilla, Hernandez et al. (2009) foi testado em grupos de 1 a 50 indivíduos com 5 exemplos

de caminhar cada. Primeiramente é gerado uma base de dados com valores dos atributos de cada indivíduo, considerada pelos autores como fase de treino. Em seguida, cada indivíduo que requisita autenticação é testado em uma árvore de decisão que verifica se o valor dos atributos desse indivíduo encontra-se ou não nos intervalos armazenados previamente como seu identificador. A Equação 41 mostra se o valor do atributo  $A$  está entre os intervalos armazenados para cada indivíduo  $i$ .

$$[A_{valor} - A_{intervalo}]_i \leq A \leq [A_{valor} + A_{intervalo}] \quad (41)$$

A Equação 41 é executada para cada atributo utilizado: altura, largura, ângulo da diagonal principal, área e frequência da área da Caixa delimitadora.  $A_{intervalo}$  é uma média aritmética entre os valores máximo e mínimo do atributo para aquele indivíduo.

A árvore é percorrida testando os atributos altura, frequência da área, área, largura e ângulo da diagonal principal, na ordem citada, para cada indivíduo que é testado. Se o indivíduo é falso para um dos atributos ele sai da árvore e não é reconhecido pelo sistema. Um indivíduo pode ser considerado, no sistema proposto pelos autores: (i) um Verdadeiro Positivo ( $VP$ ) se pertence ao intervalo testado e encontra-se em todas as listas de intervalos de atributos geradas, (ii) Falso Positivo ( $FP$ ) se ele pertence ao intervalo mas não está em uma das listas de intervalos de atributos, ( $VN$ ) se ele não pertence ao intervalo e não está em uma das listas de intervalos e ( $FN$ ) se o indivíduo não pertence ao intervalo mas encontra-se em todas as listas de intervalos cadastradas. Finalmente, a confiabilidade de reconhecimento, como é referenciado pelos autores (GUILLEN et al., 2009) é testado para o resultado final segundo a Equação 42.

$$confiabilidade = 100\% - FP - FN \quad (42)$$

### 3.7.1.2 Figuras de movimento de Yoo e Nixon

Em “*Automated Markerless Analysis of Human Gait Motion for Recognition and Classification*”, Yoo e Nixon (2011) descrevem sua metodologia baseada em modelo para extrair características espaço-temporais e cinemáticas de figuras de movimento, utilizando o banco de dados do caminhar SOTON, ambos relatados no Capítulo 2 e 3. As figuras de movimento são retornadas diretamente pela NUI *Skeleton* API, o que possibilita a extração direta dos parâmetros utilizados por Yoo e Nixon (2011) do banco de dados do caminhar capturado no presente trabalho.

Além dos atributos espaço-temporais: duração de ciclo de caminhar, tamanho da passada e velocidade; foram também utilizados parâmetros cinemáticos na forma de média geral dos ângulos das articulações inferiores, o coeficiente de correlação entre os ângulos formados pelo joelho da perna direita e esquerda e as coordenadas

centrais do ciclograma do quadril-jelho.

A correlação cruzada é um método para detectar periodicidades entre dois sinais de interesse, como medida de similaridade entre duas formas de onda em função do tempo (JOSHI; ANAND, 2010). O coeficiente da correlação cruzada  $R$  é uma quantidade que informa a qualidade da melhor adaptação de uma curva a um conjunto de pontos, e pode ser definido segundo a Equação 43:

$$R = \frac{ss_{xy}^2}{ss_{xx} \times ss_{yy}} \quad (43)$$

onde  $ss_{xx}$ ,  $ss_{yy}$  e  $ss_{xy}$  são definidos pelas Equações 44, 45 e 46 respectivamente

$$ss_{xx} = \sum_{i=1}^N x^2 - ix'^2 \quad (44)$$

$$ss_{yy} = \sum_{i=1}^N y^2 - iy'^2 \quad (45)$$

$$ss_{xy} = \sum_{i=1}^N xy - ix'y' \quad (46)$$

onde  $N$  é o número de pontos  $(x_i, y_i)$  e  $x', y'$  a sua média. Essas Equações são formas de variância e covariância de  $x$  e  $y$  (WEISSTEIN, 2014).

Os ciclogramas são formados ignorando o eixo do tempo e apontando diretamente pares de ângulos e podem ser usados para refletir a cinemática do ciclo de caminhar e, a partir dele, extrair atributos para representar o caminhar humano. Exemplos de atributos extraídos a partir do ciclograma são o perímetro e as coordenadas centrais (JOSHI; ANAND, 2010). Logo, a coordenada central do ciclograma quadril-jelho é dada pela Equação 47:

$$Centro_{QJ} = (Centro_{Quadril}, Centro_{Jelho}) \quad (47)$$

onde  $Centro_{Quadril}$  e  $Centro_{Jelho}$  são os pontos centrais do conjunto de ângulos de  $0$  a  $N$  da articulação do quadril e da articulação do jelho, respectivamente, dadas pelas Equações 48 e 49

$$Centro_{quadril} = \frac{(\min([Q_0, \dots, Q_N]) + \max([Q_0, \dots, Q_N]))}{2} \quad (48)$$

$$Centro_{jelho} = \frac{(\min([J_0, \dots, J_N]) + \max([J_0, \dots, J_N]))}{2} \quad (49)$$

Esses dois métodos são usados para criar atributos do cinemáticos do caminhar humano, a partir das figuras de movimento extraídas dos vídeos de indivíduos cami-

nhando do banco de dados SOTON, como explicado no Capítulo 3, Seção 3.1.2, e foram reproduzidos diretamente na figura de “esqueleto” criada a partir dos pontos capturados pelo sensor Kinect presentes na base de dados deste trabalho.

### **3.7.2 Comparação da metodologia proposta com trabalhos relacionados que utilizam o Kinect na obtenção da caminhada**

Dos 140 indivíduos pertencentes à base de dados construída, foram selecionados grupos com o mesmo número de indivíduos utilizados nesses trabalhos relacionados e aplicada a metodologia proposta no presente trabalho, escolhendo o *dataset* e o algoritmo de aprendizado de máquina que obteve o melhor desempenho.

Esses trabalhos relacionados fornecem o número de instâncias classificadas corretamente (ICC) como resultado de suas metodologias. Assim, comparamos o o número de ICCs obtidas com a aplicação de nossa metodologia, com os valores fornecidos pelos autores, para cada número de indivíduos analisado.

## **4 RESULTADOS**

Os resultados alcançados a partir da aplicação da metodologia apresentada no Capítulo 3 são descritos e comentados nas seções seguintes.

### **4.1 Utilização da Base de Dados do Esqueleto com outras Metodologias**

Com a base de dados criada para os experimentos deste trabalho, foram testadas outras metodologias para extração de atributos do caminhar, diferentes das apresentadas no método proposto. Para cada abordagem existente, baseada em modelo ou livre de modelo, um método foi escolhido para ser reproduzido.

Como os métodos livres de modelo estão diretamente ligados a sua forma de extração (por silhueta, energia da imagem, etc), foi escolhida uma metodologia livre de modelo que podia ser aplicada na figura de movimento ou esqueleto extraída com o sensor Kinect: a Caixa delimitadora, descrita na Seção 2.3.6 do Capítulo 2. A técnica baseada em modelo escolhida utiliza figuras de movimento semelhantes as obtidas pelo Kinect, descritas no Capítulo 2, Seção 2.3.3.

A metodologia utilizada para a reprodução de tais técnicas foi apresentada no Capítulo 3, Seção 3.7. As seções seguintes apresentam os resultados obtidos reproduzindo as mesmas técnicas com a base de dados criada para o presente trabalho.

#### **4.1.1 Reprodução da técnica de caixa delimitadora**

Para reproduzir o experimento realizado em Guillen, Padilla, Hernandez et al. (2009) foi necessário criar dois grupos distintos, com a mesma quantidade de indivíduos cada, sendo que um grupo representa o grupo de “pessoas autenticadas”, cujos atributos são conhecidos e calculados previamente, e o seguinte o grupo de “pessoas intrusas” ou não cadastradas.

O grupo autenticado foi treinado, onde as medidas da caixa delimitadora de cada um dos indivíduos desse grupo foram agrupadas nas tabelas de intervalos, como explicado no Capítulo 3, Seção 3.7.1.1. O grupo seguinte, de intrusos, é testado seguindo

a árvore de decisão descrita na seção mencionada, para obter os valores de falsos positivos (FP), pessoas intrusas que conseguem autenticação, e verdadeiros negativos (VN), pessoas intrusas que não conseguem autenticação.

O próprio grupo autenticado também é testado, obtendo os valores de verdadeiros positivos (VP), pessoas autorizadas que conseguem autenticação, e falsos negativos (FN), pessoas autorizadas que não conseguem autenticação. Os resultados podem ser visualizados na Tabela 4. Os resultados obtidos pelos autores aplicando a Equação 42 podem ser vistos na Tabela 5.

Tabela 4: Comparação dos resultados obtidos utilizando a Caixa delimitadora

Tamanho dos Grupos	Total de indivíduos	VP	VN	FN	FP
3	6	2	3	1	0
5	10	1	5	4	0
10	20	5	8	5	2
20	40	18	3	2	17

O tamanho de cada grupo, autenticado ou intruso, varia conforme é mostrado na Tabela 4. O total de indivíduos é a soma do tamanho desses dois grupos. Com o aumento do número de indivíduos, as faixas de intervalos aplicadas na Equação 41 tendem a satisfazer um número maior de indivíduos, aumentando o número de falsos positivos (FP).

Tabela 5: Comparação dos resultados obtidos utilizando a Caixa delimitadora

Total de indivíduos	Guillen et al.	Reprodução
6	96.4%	83,4%
10	95.8%	60%
20	97.6%	65%
40	98.8%	52,5%

Os resultados obtidos utilizando a Equação 42 são apresentados na Tabela 5. A Equação 42 utiliza valores de falso negativos (FN) e falsos positivos (FP) apresentados na Tabela 4.

Os resultados obtidos por Guillen, Padilla, Hernandez et al. (2009) aumentam com o acréscimo de indivíduos, enquanto que os resultados obtidos na reprodução do experimento decaem com o aumento de indivíduos. Para a maioria dos algoritmos de classificação, o aumento do número de indivíduos faz com que a taxa de acertos caia. Os valores percentuais de acerto obtidos com a reprodução estão dentro do esperado. Em contrapartida os resultados dos autores aumentam proporcionalmente ao aumento de indivíduos, apresentando um comportamento atípico de classificadores conhecidos. Guillen, Padilla, Hernandez et al. (2009) não fornecem maiores deta-

lhes do banco de dados utilizado para segmentação das imagens, tão pouco das características de altura dos indivíduos, que podem contribuir para os altos resultados obtidos.

#### 4.1.2 Reprodução da técnica de figuras de movimento

A reprodução do experimento realizado por Yoo e Nixon (2011) pode ser feita diretamente utilizando o banco de dados capturado no presente trabalho, devido a aparente semelhança entre o modelo utilizado pelos autores - as figuras de movimento - com o esqueleto composto pelas articulações rastreadas pelo sensor Kinect.

Os atributos foram extraídos dos exemplos de caminhadas do banco de dados obtido para o presente trabalho da maneira apontada no Capítulo 3, Seção 3.7.1.2. e os resultados obtidos para cada grupo de indivíduos são apresentados nas Tabelas 6, 7 e 8.

Tabela 6: Comparação de 30 Indivíduos

K	Yoo e Nixon	Reprodução
1	96.7%	59.3%
3	93.3%	58.6%
5	96.7%	62%

Tabela 7: Comparação de 60 Indivíduos

K	Yoo e Nixon	Reprodução
1	91.7%	45.8%
3	86.7%	46.8%
5	85.7%	47.1%

Tabela 8: Comparação de 100 Indivíduos

K	Yoo e Nixon	Reprodução
1	84.0%	36.4%
3	80.0%	38.6%
5	82.0%	39.8%

Os valores de instâncias classificadas corretamente (ICCs) obtidas por Yoo e Nixon (2011) variam de acordo com o valor de K escolhido e tendem a diminuir conforme o número de indivíduos aumenta. Os valores de ICCs obtidos com a reprodução do experimento refletem o mesmo comportamento com acurácias de menor valor. Uma possível justificativa para esse comportamento é que as figuras de movimento obtidas

diretamente do sensor Kinect passaram pelo menor número de filtros possíveis, sem a preocupação de encaixar esse modelo em algum tipo de atributo, ao contrário das figuras de movimento de Yoo e Nixon (2011) que são extraídas visando a extração dos atributos escolhidos pelos autores.

Da mesma forma, a qualidade do caminhar humano no banco de dados SOTON utilizado pelos autores é diferente da qualidade do caminhar capturado com o Kinect: no SOTON o ambiente é controlado e as silhuetas são melhoradas até fornecerem uma figura de movimento bem definida, o que não ocorre com as capturas utilizando o Kinect.

## 4.2 Ajuste dos Algoritmos de Aprendizado de Máquina

Para escolher o algoritmo de aprendizado de máquina que melhor desempenhasse a classificação do maior número de indivíduos e de todos os atributos escolhidos como metodologia, foi executado o procedimento apresentado no Capítulo 3, Seção 3.6. A Tabela 9 apresenta os melhores resultados para cada algoritmo, apresentando destacado a configuração escolhida para cada grupo de algoritmo. A Figura 54 apresenta graficamente esses resultados.

Tabela 9: Configuração dos algoritmos e melhores resultados obtidos

Algoritmo	Configuração	ICCs	ICCs (%)
MLP	<b>-H 43, -N 500</b>	<b>595</b>	<b>84,75%</b>
	-H 48, -N 1200	595	84,75%
	-H 48, -N 1800	595	84,75%
	-H 48, -N 1900	595	84,75%
k-NN	-K 2, -A LinearNNSearch ManhattanDistance, -l	598	85,2%
	-K 3, -A LinearNNSearch ManhattanDistance, -l	602	85,75%
	<b>-K 4, -A LinearNNSearch ManhattanDistance, -l</b>	<b>616</b>	<b>87,74%</b>
	-K 5, -A LinearNNSearch ManhattanDistance, -l	613	87,32%
C-SVC	-K Polinomial, -C 1, -D 1, -R 0.1, -G 0.0	366	52,2%
	-K Polinomial, -C 1, -D 1, -R 0.2, -G 0.0	366	52,2%
	-K Polinomial, -C 1, -D 1, -R 0.3, -G 0.0	366	52,2%
	<b>-K Polinomial, -C 1, -D 1, -R 0.4, -G 0.0</b>	<b>366</b>	<b>52,2%</b>
nu-SVC	-K Polinomial, -N 0.7, -D 1, -R 0.9, -G 0.6	363	51,7%
	-K Polinomial, -N 0.7, -D 1, -R 1.0, -G 0.6	363	51,7%
	-K Polinomial, -N 0.7, -D 1, -R 0.0, -G 0.7	363	51,7%
	<b>-K Polinomial, -N 0.7, -D 1, -R 0.1, -G 0.7</b>	<b>363</b>	<b>51,7%</b>
SMO	<b>-K Puk, -C 20.0, -O 3.0, -S 6.0</b>	<b>613</b>	<b>87,3%</b>
	-K NormalizedPolyKernel, -C 50.0, -E 2.0	588	83,7%
	-K RBFKernel, -C 100.0, -G 0.5	608	86,6%
	-K PolyKernel, -C 100.0, -E 1.0	613	87,3%

O algoritmo e a configuração escolhida foi *k-NN* com  $k = 4$ , por possuir maior

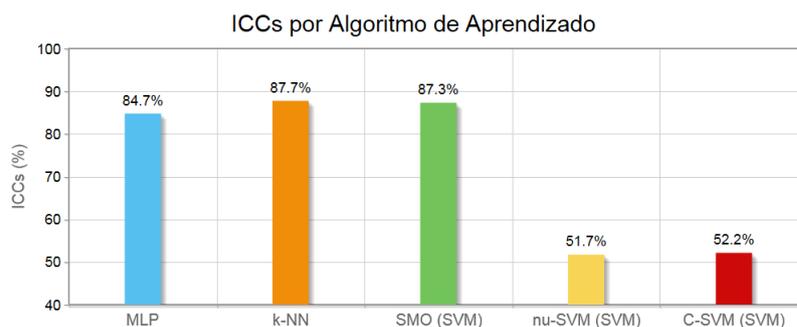


Figura 54: Resultado dos ICCs por algoritmos de aprendizado de máquina utilizando todos os atributos e o maior número de indivíduos do conjunto.

número de ICCs, 616 de 702 exemplos classificados corretamente, obtendo uma taxa de acerto de 87,74%, como pode ser visto na Tabela 9. Embora a taxa de ICCs para SMO seja muito próxima ao resultado obtido à configuração escolhida para  $k$ -NN, o tempo gasto para construir o modelo  $k$ -NN é inferior (menor que 1 segundo) ao tempo gasto para construir o modelo usando SMO (maior que 60 segundos). Além do número de ICCs ligeiramente maior, esse diferencial apontou a melhor configuração e o algoritmo  $k$ -NN foi escolhido como metodologia de classificação do presente trabalho.

### 4.3 Resultados Utilizando os Atributos do Método Proposto

Utilizando o algoritmo e a configuração escolhida,  $k$ -NN com  $k = 4$ , Algoritmo de busca *LinearNNSearch* com função de distância *ManhattanDistance* e o peso igual a  $1/distancia$ , foram classificados todos os *datasets*, e a principal escolha entre eles, contendo todos os atributos. O número de indivíduos classificados é 140, onde cada indivíduo contém em média, 5 exemplos de caminhar, totalizando 702 exemplos. Os resultados das taxas de ICCs são apresentados na Figura 55.

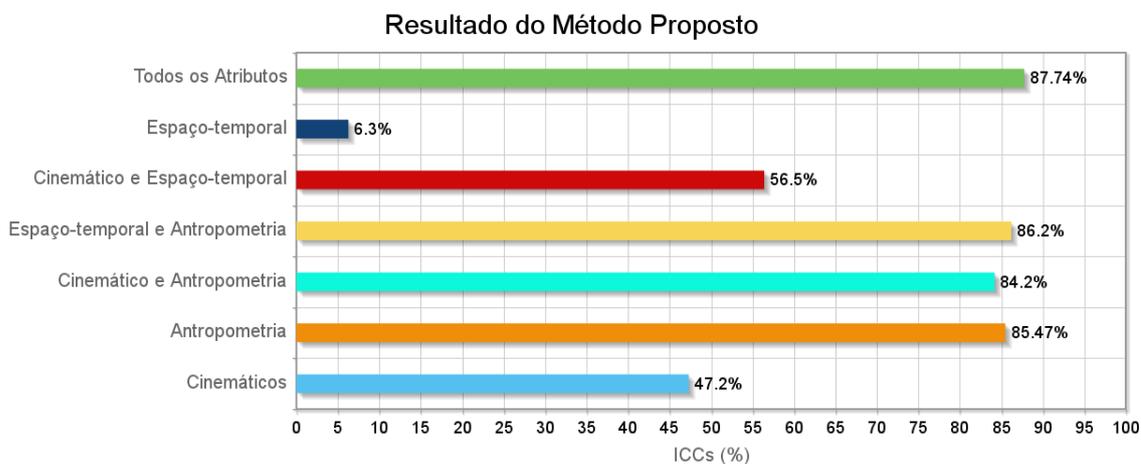


Figura 55: Resultado dos ICCs utilizando  $k$ -NN ( $k=4$ ) por *datasets* com diferentes atributos e o método proposto (todos os atributos).

O uso da antropometria combinada aos atributos do caminhar humano apresenta um aumento no número de ICCs em comparação a estes *datasets* usados individualmente. De fato, o uso de atributos cinemáticos ou espaço-temporais individualmente, para um grande número de indivíduos, mostrou-se impraticável, com taxas equivalentes à 47.2% e 6.3% respectivamente. Atributos espaço-temporais combinados com atributos antropométricos apresentam o maior ganho do ponto de vista da classificação por antropometria, chegando a um resultado de 86.2% de ICCs, próximo a classificação com todos os atributos, 87,74%.

Ao aplicar um recurso de seleção de subconjunto baseado em correlação (HALL, 1998), de um total de 80 atributos antropométricos e do caminhar humano, foram selecionados 14 atributos. O número de instâncias classificadas corretamente utilizando esses 14 atributos, para os 140 indivíduos, foi de 87.6%, apenas 0.1 pontos abaixo da taxa de ICCs utilizando todos os atributos. A Tabela 10 mostra os atributos selecionados e a mudança de acurácia quando o atributo é removido.

Tabela 10: Atributos selecionados através da seleção de subconjunto baseado em correlação e o quanto remover cada um destes atributos afeta a taxa de ICCs, em pontos percentuais.

Atributo	ICCs(%)	Mudança(%)
Tamanho da passada (ciclo)	83,1%	-4,64%
Pé direito	84,0%	-3,74%
Mão direita	84,7%	-3,04%
Pescoço	85,0%	-2,74%
Coluna superior	85,7%	-2,04%
Ombro direito	85,8%	-1,94%
Quadril esquerdo	85,8%	-1,94%
Altura	86,1%	-1,64%
Perna direita	86,3%	-1,44%
Picos altos do tornozelo esquerdo	86,9%	-0,84%
Antebraço esquerdo	87,0%	-0,74%
Fêmur direito	87,0%	-0,74%
Perna esquerda	87,1%	-0,64%
Fêmur esquerdo	87,74%	0,0%

Entre os 14 atributos selecionados, 12 são antropométricos e apenas 2 relacionados ao caminhar humano: tamanho da passada (ou tamanho do ciclo de caminhar) e a média dos picos altos do tornozelo esquerdo. Isso reforça a ideia de que atributos do caminhar humano estão correlacionados com atributos antropométricos mas menos confiáveis para identificação. Contudo, o atributo que causa maior queda de acurácia ao ser retirado é um atributo do caminhar humano: o tamanho da passada (ou tamanho do ciclo de caminhar).

A seleção de atributos mostra que nem todos os atributos são importantes, mas uns são mais necessários do que outros. Quando é preciso selecionar apenas alguns atributos, a avaliação executada mostra que ao menos o tamanho do ciclo e alguns atributos antropométricos são necessários para uma melhor classificação.

#### 4.3.1 Efeito da Divisão por Número de Indivíduos

A partir do número total de indivíduos selecionados para as classificações, foram gerados grupos com diferentes números de indivíduos, variando de 5 em 5 indivíduos. Para cada grupo, 10 grupos distintos foram classificados e a mediana das instâncias classificadas corretamente são apresentadas na Figura 56. A Figura 57 apresenta detalhes da classificação dos *datasets* contendo atributos antropométricos cujos resultados apresentaram-se muito próximos.

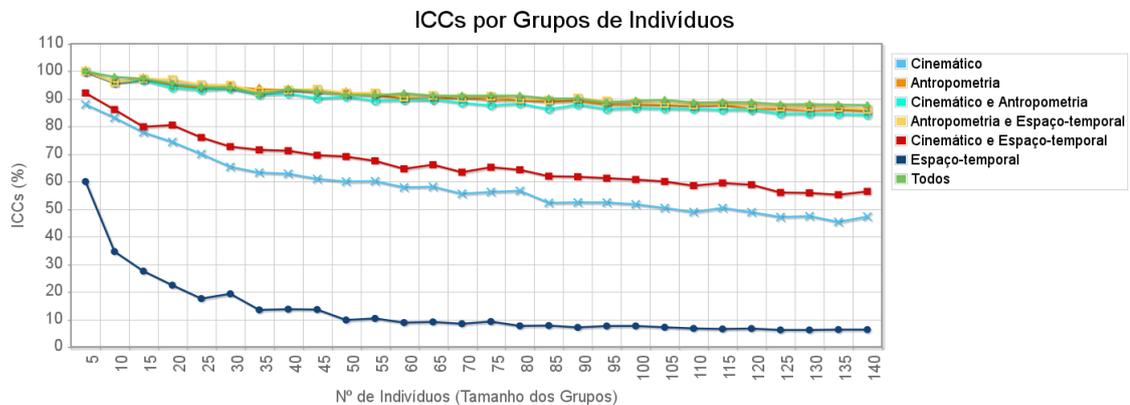


Figura 56: Mediana dos ICCs de cada grupo de indivíduos.

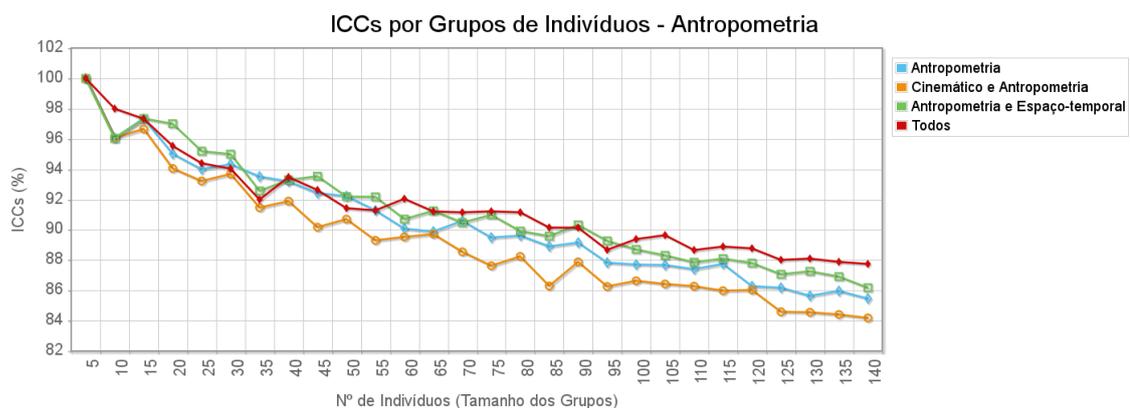


Figura 57: Mediana dos ICCs de cada grupo de indivíduos para *datasets* contendo atributos antropométricos.

Para classificação de indivíduos com *datasets* que utilizam só caminhar humano (atributos espaço-temporais e cinemáticos), o tamanho máximo do grupo ou número máximo de indivíduos a serem classificados é aproximadamente 50, como mostra a

Figura 56. A partir de 50 indivíduos, o número de ICCs começa a decair chegando a valores próximos a 50% com 140 indivíduos. A utilização de atributos antropométricos aumenta o número de ICCs possibilitando a classificação de um maior número de indivíduos, mantendo as taxas de acertos muito próximas com um decaimento suave.

#### 4.4 Exploração do Número de Quadros e Ciclos de Caminhar

O *dataset* de indivíduos capturados é constituído de 5 a 6 exemplos de caminhadas para cada pessoa. Cada exemplo de caminhada possui um número diferente de quadros capturados, variando de 300 a 1000 quadros por captura e de 2 a 23 ciclos de caminhar. As Figuras 59 e 58 mostram a distribuição dos quadros e ciclos por número de exemplos. A partir desta distribuição foi possível obter a faixa de quadros e ciclos com o maior número de exemplos para análise da quantidade de quadros e ciclos necessários para classificar corretamente.

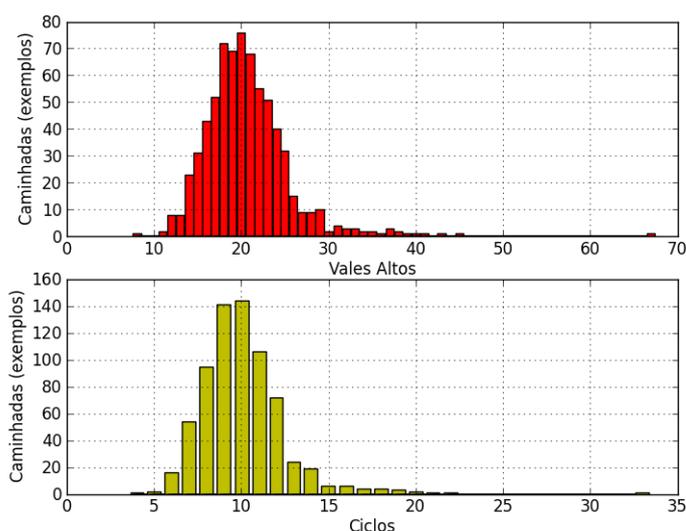


Figura 58: Histogramas com número de exemplos de caminhar distribuídos por número de vales altos e por ciclos.

Para isolar quadros de caminhar, foi analisada a distribuição dos quadros capturados por caminhada, como mostra o histograma da Figura 59. O maior número de exemplos de caminhada possuía em comum aproximadamente 400 quadros. Dos 706 exemplos, 697 caminhadas podiam variar de 1 a 400 quadros sem perda de indivíduos. Foram então extraídos os parâmetros antropométricos - medidas dos segmentos do corpo e altura - para cada subgrupo de número de quadros, de 1 a 400, aplicando o algoritmo k-NN com a configuração utilizada no presente trabalho. A Figura 60 mostra o número de acertos em percentagem para cada subgrupo.

Para isolar ciclos de caminhar, foi utilizado o conjunto de ângulos do joelho direito

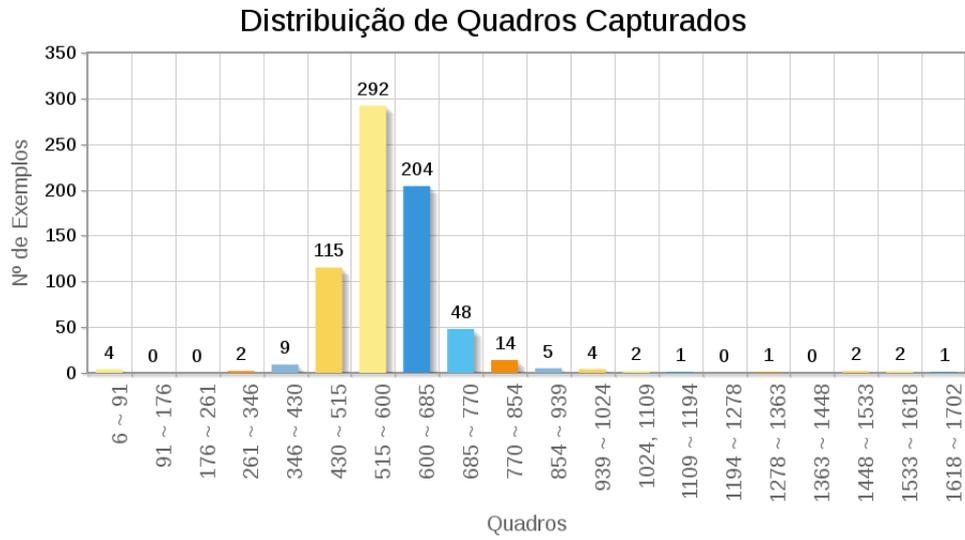


Figura 59: Histograma com número de exemplos de caminhar distribuídos por número de quadros capturados.



Figura 60: Resultados das instâncias classificadas corretamente (ICCs) utilizando atributos de antropometria por grupos de números de quadros capturados.

dos indivíduos e obtido o conjunto de vales altos a partir da técnica de gradientes, descrita no Capítulo 7, Seção 7.3. Cada quadro que contém um vale alto foi usado para marcar o início ou o fim de um ciclo de caminhar, como ilustra a Figura 58. A densidade de exemplos por número de vales altos é ilustrado na Figura 58 (a).

A análise de ciclos de caminhar consistiu em separar indivíduos com um número mínimo de exemplos que possuíssem de 1 até 10 ciclos de caminhar. Para cada indivíduo, foram separados 4 exemplos onde cada um possui até 10 ciclos por exemplo, totalizando 196 exemplos de 49 indivíduos distintos.

Para obter o número de ciclos de cada exemplo de caminhada foram separados os vales altos dos ângulos descritos pelo joelho direito. Os quadros que contém os vales altos foram armazenados e usados como marcadores de ciclos. Para cada número de ciclos, foram calculados os atributos de caminhar humano cinemáticos e espaço-temporais e posteriormente submetidos ao algoritmo de aprendizado de máquina escolhido para os experimentos. Juntamente com os atributos de caminhar humano, foi avaliado também o uso de antropometria com a unidade de ciclo. Os resultados são apresentados na Figura 61.

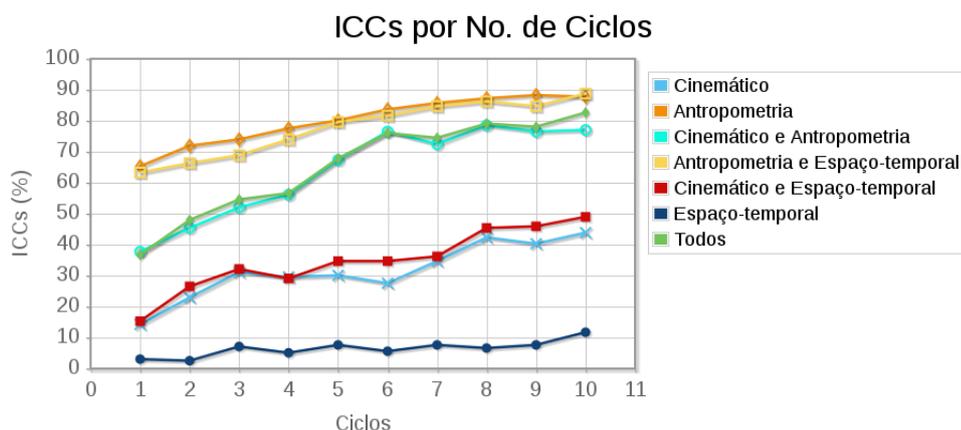


Figura 61: Número de instâncias classificadas corretamente (ICCs) para diferentes números de ciclos.

Na análise por número de ciclos, os *datasets* que importam são aqueles que contém atributos do caminhar humano. Os outros *datasets* foram mantidos para uma comparação com número de quadros. O *dataset* que combina todos os atributos do caminhar humano apresentou um aumento no número de ICCs a partir de 10 ciclos. O uso de atributos antropométricos em conjunto possibilita melhores resultados a partir de 5 ciclos de caminhar capturados.

Já na exploração por número de quadros capturados, é possível observar pela Figura 60 que 20 é o número mínimo de quadros necessários a serem capturados de um indivíduo em frente ao sensor Kinect, para uma classificação com altas acurácias, utilizando apenas atributos antropométricos. Isso significa que um indivíduo necessita

ao menos de 1 segundo em frente ao sensor Kinect para que seja classificado corretamente, já que o mesmo sensor captura até 30 quadros por segundo. Os resultados exibidos na Figura 60 mostram também que de 100 a 400 quadros, o número de ICCs parece estabilizar sem grandes modificações, indicando que muitos quadros não são extremamente necessários para bons resultados utilizando antropometria.

## 4.5 Comparação com Resultados de Métodos Estado-da-Arte

Os principais resultados obtidos com o método apresentado neste trabalho foram comparados com as metodologias apresentadas em Yoo e Nixon (2011), estado da arte em extração de parâmetros do caminhar humano baseado em modelo e as metodologias apresentadas em trabalhos recentes que também utilizam o Kinect para extrair o modelo de caminhar. O primeiro trabalho que utiliza o Kinect é uma técnica livre de modelo, proposta por Hofmann, Bachmann e Rigoll (2012) e apresentada na Seção 3.2.5; os DGHEIs. Em seguida é comparado com a metodologia proposta por Munsell et al. (2012) e finalmente com o trabalho de Preis, Kessel e Werner (2012).

### 4.5.1 Figuras de movimento de Yoo e Nixon

Em Yoo e Nixon (2011), utilizando o método de extração de parâmetros apresentado na Seção 3.1.2, no banco de dados SOTON citado na Seção 2.2.5, os autores obtiveram como parâmetros do caminhar humano: altura do corpo, duração do ciclo de caminhar, tamanho da passada, velocidade, média dos ângulos das articulações, variação dos ângulos dos quadris, o coeficiente de correlação entre os ângulos esquerdo e direito da perna, e as coordenadas centrais do ciclograma quadril Joelho. A Figura 62 apresenta uma comparação entre os resultados obtidos por Yoo e Nixon (2011), utilizando o algoritmo de aprendizado k-NN com  $k=1$  e diferentes números de indivíduos, comparado com o método proposto no presente trabalho, utilizando o mesmo número de indivíduos e configuração do algoritmo k-NN proposta.

A Tabela 11 apresenta os resultados obtidos para as instâncias classificadas corretamente (ICCs) por Yoo e Nixon (2011) e os resultados obtidos de ICCs através do método proposto neste trabalho, para o mesmo número de pessoas.

Tabela 11: Comparação entre as metodologias

No. de pessoas	Yoo e Nixon	Metodologia proposta
30	96.7%	94.0%
60	91.7%	91.4%
100	84.0%	89.4%

É possível observar pelo gráfico apresentado na Figura 62 que para um número

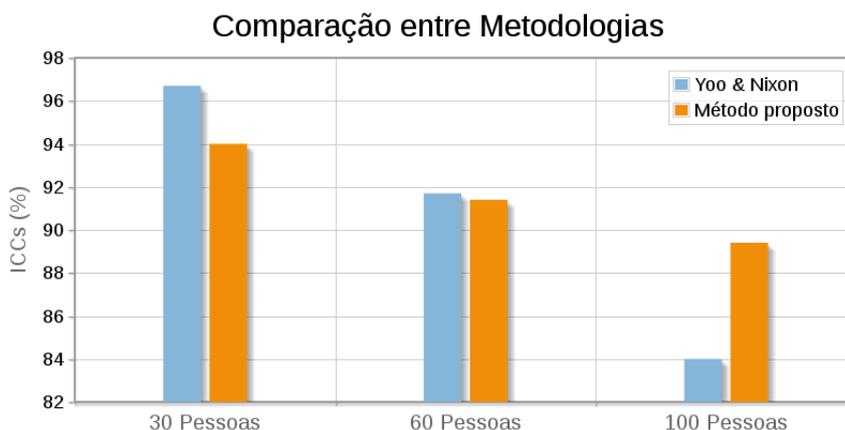


Figura 62: Resultados obtidos por Yoo e Nixon (2011) e resultados obtidos pelo método proposto: ICCs apresenta o número de instâncias classificadas corretamente para o mesmo número de indivíduos.

pequeno de pessoas, a metodologia de Yoo e Nixon é superior ao método proposto neste trabalho, atingindo 96.7% de ICCs contra 94% de ICCs do presente trabalho. Em contrapartida, conforme o número de pessoas aumenta, o ICC cai de forma brusca, atingindo 84%, diferentemente do resultado apresentado no presente projeto, que decai para 89.4%. Para um maior número de pessoas, o método proposto apresenta-se superior ao de Yoo e Nixon (2011) e mantém-se dessa forma com o aumento de indivíduos, como pode ser visto na Figura 56. Deve-se considerar também que as diferenças entre as metodologias para 30 e 60 pessoas não ultrapassam 5%.

#### 4.5.2 DGHEIs

O DGHEIs é um método livre de modelo apresentado na Seção 3.2.5 e introduzido por Hofmann, Bachmann e Rigoll (2012). A partir desse método, os autores extraíram histogramas de gradientes para caracterizar cada indivíduo, em um universo de 176 pessoas, classificando com o algoritmo *nearest-neighbor* e distância euclidiana. Os ICCs obtidos com os DGHEIs são apresentados na Tabela 12 e na Figura 63.

Embora o número de indivíduos utilizados no *dataset* de Hofmann, Bachmann e Rigoll (2012) seja ligeiramente superior ao utilizado no presente trabalho, os resultados apresentados na Figura 57 na Seção 4.3.1 mostram que existe pouca variação com o aumento de indivíduos, possibilitando equiparar os resultados obtidos quando um mesmo número de indivíduos fosse classificado. Para isso, foi aplicada uma curva de tendência utilizando uma função logarítmica que mais se adapta à curva com ICCs de todos os atributos para diferentes tamanhos de grupos de indivíduos, apresentado na Figura 56.

Segundo o gráfico da Figura 64, obteve-se a tendência da curva até 176 indivíduos com uma taxa de ICCs em pontos percentuais alcançados de aproximadamente 85%, possibilitando observar um resultado semelhante ou superior do método proposto

apresentado neste trabalho aos resultados obtidos por Hoffmann, Bachmann e Rigoll (2012) com o método dos DGHEIs.

Tabela 12: Comparação entre as metodologias

Método	No. de Indivíduos	ICCs
DGHEIs	176	81.0%
Proposto	140	87.7%

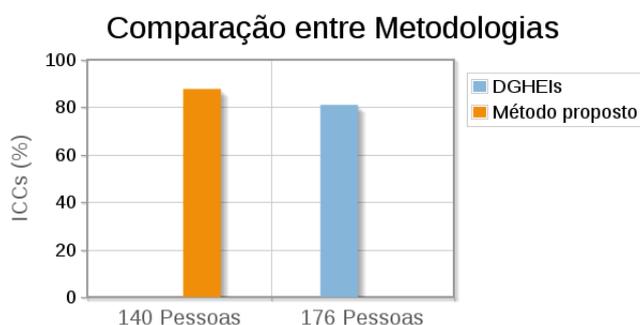


Figura 63: Número de ICCs em pontos percentuais para a classificação utilizando o método proposto, com 140 indivíduos, e o método DGHEIs, com 176 indivíduos.

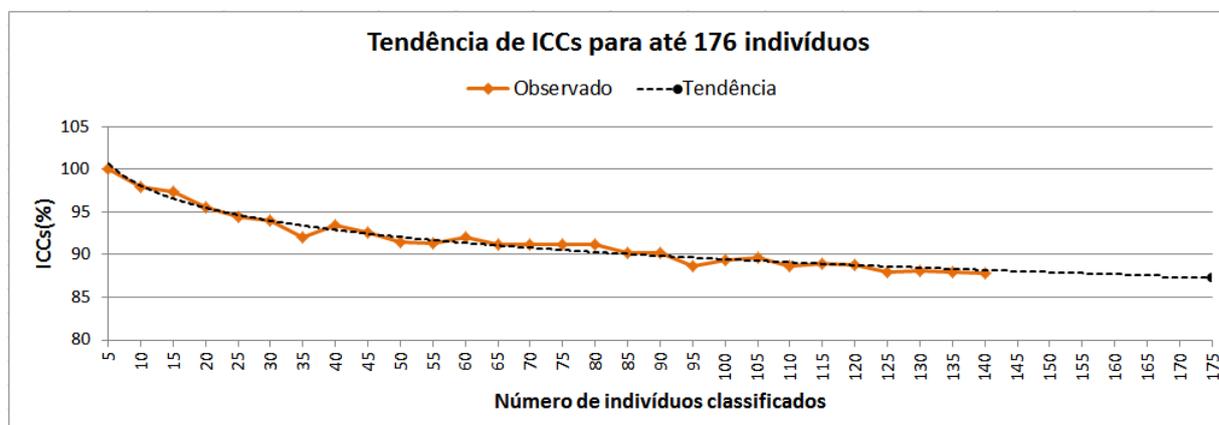


Figura 64: Aplicação de uma função logarítmica próxima a curva de ICCs obtida para todos os atributos do método proposto possibilita acompanhar um possível resultado com até 176 indivíduos.

#### 4.5.3 Outras metodologias utilizando o Sensor Kinect

Nos trabalhos de Munsell et al. (2012) e Preis et al. (2012), são utilizados pequenos bancos de dados com até 10 indivíduos compostos com o modelo de esqueleto retornado da NUI *Skeleton* API. Munsell et al. (2012) utilizaram atributos antropométricos juntamente com as rotações de todas as articulações, os quais os autores chamam de “atributos de movimento”. Em Preis et al. (2012), utilizam os atributos

antropométricos como altura, tamanho das pernas, braços e coluna, juntamente com atributos espaço-temporais de velocidade e tamanho do passo. Classificam esses atributos utilizando alguns algoritmos, cujos principais são C4.5 e *Naive Bayes*. Os melhores resultados obtidos pelos autores foram comparados com os resultados obtidos no método proposto neste trabalho, e podem ser visualizados na Tabela 13 e Figura 65.

Tabela 13: Comparação entre as metodologias

Método	No. de Indivíduos	ICCs
Munsell et al.	10	90.0%
Preis et al.	9	91.0%
Proposto	10	98.0%

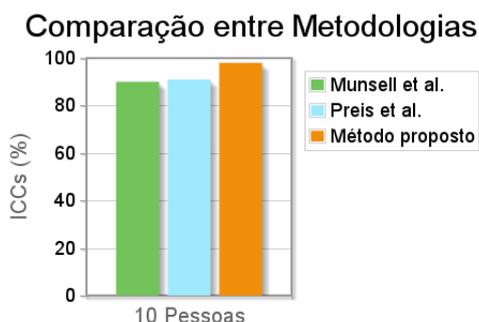


Figura 65: Percentual de ICCs para a classificação de 10 indivíduos, comparando o método proposto com o resultado obtido em Munsell et al.(2012) e Preis et al. (2012).

Os resultados obtidos com outras metodologias utilizando o sensor Kinect são inferiores ao do método proposto e não fornecem resultados com um número maior de indivíduos para uma melhor comparação.

## 4.6 Resultados Acessórios

As seções seguintes apresentam a classificação de indivíduos por antropometria e caminhar humano explorando a classificação por gênero e índice de massa corporal (IMC).

### 4.6.1 Classificação por gênero

Foi explorado no presente trabalho a classificação de indivíduos de acordo com seu gênero, utilizando os atributos por grupos, apresentados na Seção 3.5. Dos *datasets* de 140 indivíduos, foram separadas 44 mulheres com 109 exemplos e 44 homens com 112 exemplos, totalizando 221 exemplos de 88 pessoas, gerando dois grupos distintos

“g1” e “g2”, com total de 44 pessoas cada e 22 pessoas para cada gênero. Os dois grupos, “g1” e “g2”, foram classificados como treino e teste e logo após, invertidos, gerando dois resultados distintos. No final do processo a média dos ICCs obtidos foi calculada.

Utilizando k-NN com a configuração da metodologia proposta, o número de instâncias corretamente classificadas (ICCs) para um total de 221 exemplos é apresentado na Tabela 14. A Tabela 15 mostra o número de exemplos corretamente classificados como feminino e como masculino (diagonal principal), para a combinação dos grupos “g1” como treino e “g2” como teste. Igualmente, a Tabela 16 mostra os exemplos classificados para a combinação contrária, “g2” como treino e “g1” como teste.

Tabela 14: Classificação por Gênero

<i>Dataset</i>	g1-g2 (%)	g2-g1 (%)	Média (%)
Cinemático	73.3	72.3	72.8
Antropometria	81.4	78.7	80.0
Cinemático e Antropometria	76.9	84.6	80.7
Cinemático e Espaço-temporal	79.6	74.2	76.9
Antropometria e Espaço-temporal	82.3	80.5	81.4
Espaço-temporal	71.0	75.1	73.0
Todos os atributos	80.9	85.0	83.0

Tabela 15: Matriz de confusão g1-g2 para todos os atributos

Feminino	Masculino	← Classificado como
96	13	Feminino (109 exemplos)
29	83	Masculino (112 exemplos)

Tabela 16: Matriz de confusão g2-g1 para todos os atributos

Feminino	Masculino	← Classificado como
90	20	Feminino (109 exemplos)
13	98	Masculino (112 exemplos)

É possível observar que os diferentes resultados para os diferentes tipos de *data-sets* são melhores que 50%, podendo ser utilizados em conjunto ou separados para identificar com acurácia razoável o gênero do indivíduo.

#### 4.6.2 Classificação por índice de massa corporal (IMC)

A possibilidade de se classificar um exemplo de caminhar humano através do índice de massa corporal (IMC) também foi explorado neste trabalho. O IMC é um

índice simples de massa corporal (peso) por altura que é comumente utilizado para classificar as pessoas adultas abaixo, acima e muito acima (obesidade) do peso ideal. As características de IMC do *dataset* capturado foram descritas no Capítulo 3, Seção 3.2.2, totalizando 112 indivíduos com informações de massa corporal. A distribuição do IMC dos indivíduos pode ser vista na Figura 43.

Das quatro categorias; saudáveis, sobrepeso, obesos e magros, foram criadas três classes correspondentes a exemplos de pessoas “Saudáveis”, “Sobrepeso-ou-Obesos” e “Magros”, devido a grande maioria dos exemplos pertencerem a pessoas saudáveis, e o menor número de exemplos corresponde aos grupos de sobrepeso e obesidade. Dessa forma, grupos distintos com exemplos para as 3 classes foram agrupados em “imc1” e “imc2”, com 20 exemplos para cada classe em cada grupo.

Utilizando o algoritmo k-NN e os parâmetros escolhidos na metodologia proposta, foi utilizado o grupo “imc1” como treino e “imc2” como teste e, logo após, a combinação inversa. Com a aplicação destes dois métodos foram obtidos os valores de ICCs para cada combinação, apresentados na Tabela 17.

Tabela 17: Classificação por Índice de Massa Corporal

<i>Dataset</i>	imc1-imc2 (%)	imc2-imc1 (%)	Média (%)
Cinemático	41.6	45	43.3
Antropometria	23.8	33.8	28.8
Cinemático e Antropometria	31.6	43.3	37.5
Cinemático e Espaço-temporal	40	45	42.5
Antropometria e Espaço-temporal	25	48.3	36.6
Espaço-temporal	51.6	53.3	52.5
Todos os atributos	26.6	46.6	36.6

As matrizes de confusão para as combinações de treino e teste imc1-imc2 e imc2-imc1 pode ser visualizadas nas Tabelas 18 e 19. A diagonal principal mostra os verdadeiros positivos para as classes de IMC utilizadas.

Tabela 18: Matriz de confusão imc1-imc2 para atributos espaço-temporais

Sobrepeso/Obeso	Saudável	Magro	← Classificado como
11	9	0	Sobrepeso/Obeso (20 exemplos)
5	9	6	Saudável (20 exemplos)
4	5	11	Magro (20 exemplos)

A classificação por IMC apresentou resultados acima de 30% em *datasets* que utilizam atributos do caminhar humano, enquanto que atributos antropométricos não são relevantes para a classificação por IMC. O *dataset* Espaço-temporal teve o melhor desempenho.

Tabela 19: Matriz de confusão imc2-imc1 para atributos espaço-temporais

Sobrepeso/Obeso	Saudável	Magro	← Classificado como
8	4	8	Sobrepeso/Obeso (20 exemplos)
5	10	5	Saudável (20 exemplos)
0	6	14	Magro (20 exemplos)

## 4.7 Publicações

A metodologia apresentada resultou na publicação dos seguintes trabalhos em conferências internacionais:

1. ***Towards skeleton biometric identification using the microsoft kinect sensor***, 28th Annual ACM Symposium On Applied Computing - ACM SAC 2013 (*Qualis A1*), onde foram publicados os primeiros experimentos com atributos antropométricos utilizando poucos indivíduos.
2. ***Anthropometric and human gait identification using skeleton data from Kinect sensor***, 29th Annual ACM Symposium On Applied Computing - ACM SAC 2014 (*Qualis A1*), onde foram publicados resultados utilizando atributos antropométricos e cinemáticos, com o banco de dados criado no presente trabalho.
3. ***Full Body Person Identification Using the Kinect Sensor***, 26th IEEE International Conference on Tools with Artificial Intelligence - ICTAI 2014 (*Qualis A2*), onde toda a metodologia proposta no presente trabalho foi publicada.

## 5 CONCLUSÃO

Foi apresentado neste trabalho uma introdução aos sistemas biométricos que utilizam antropometria e o caminhar humano como características biológicas para a identificação e distinção de indivíduos. Metodologias para extração de parâmetros do caminhar, baseada em modelo e livre de modelo, também foram enunciadas, assim como os principais bancos de dados do caminhar humano.

Com base nessas metodologias, foi proposto o uso do sensor Microsoft Kinect para simplificar as etapas de segmentação de imagem, visão computacional e obtenção de um modelo de “esqueleto” ou figura de caminhar para a extração de características antropométricas, espaço-temporais e cinemáticas de sequências de quadros capturados de indivíduos caminhando.

A partir da captura de 164 pessoas caminhando em frente ao sensor Kinect, que povoaram o banco de dados do caminhar apresentado neste trabalho, foi proposta uma metodologia composta de filtragem e seleção de indivíduos, extração de parâmetros antropométricos e parâmetros do caminhar humano, espaço-temporais e cinemáticos, que combinados ou individualmente, compuseram 6 *datasets* distintos para classificação de pessoas.

O *dataset* com todos os atributos e com o maior número de indivíduos e exemplos foi então usado para treinar algoritmos de aprendizado de máquina e suas eficácias em classificar cada indivíduo foram avaliadas. O algoritmo que teve o melhor desempenho foi então utilizado em todas as outras classificações feitas neste trabalho. Os algoritmos considerados foram: *k-Nearest Neighbor*, *Multilayer Perceptron* e *Support Vector Machines (SVM)*. O algoritmo *k-Nearest Neighbor* obteve o maior desempenho, seguido do algoritmo *Support Vector Machines*. O primeiro, além de apresentar o melhor desempenho, supera o segundo também na rapidez de execução e quantidade de memória consumida durante o processo.

Os experimentos realizados neste trabalho também levaram em consideração o número de indivíduos do grupo classificado: Atributos cinemáticos, relacionados ao caminhar humano tendem a diminuir o número de instâncias corretamente classificadas com o aumento de indivíduos. Atributos espaço-temporais não devem ser

utilizados separadamente dos cinemáticos e os resultados mostram que ambos aumentam sua taxa de ICCs quando utilizados em conjunto. Atributos antropométricos tendem a se manter mesmo com o aumento de indivíduos no grupo e quando utilizados juntamente com outros atributos também apresentam uma melhora, sendo a melhor combinação o *dataset* “Todos os atributos” seguido do *dataset* “Antropometria e espaço-temporal”.

As unidades principais para cada sistema biométrico também foram analisadas: para antropometria, foram agrupados os resultados das classificações de acordo com o número de quadros capturados de um indivíduo. Os resultados mostram que o número de ICCs é proporcional ao número de quadros utilizados, sendo necessário, no mínimo, 10 quadros para iniciar uma classificação com mais de 60% de acertos. Para o caminhar humano, o número de ciclos de caminhar capturados em uma caminhada também influencia no número de ICCs. Quanto mais ciclos de caminhar, melhor os atributos propostos parecem se comportar. Atributos antropométricos também foram considerados quando testados com número de ciclos, para saber quando se é possível utilizar antropometria usando o ciclo como unidade básica. O número mínimo de quadros contidos em um ciclo já proporciona uma taxa de ICCs de 65%.

Além disso foi possível classificar os exemplos por gênero e por índice de massa corporal (IMC). A classificação por gênero apresentou resultados satisfatórios utilizando a metodologia proposta com todos os atributos. Atributos de caminhar humano quando utilizados individualmente apresentaram taxas de ICCs relativamente aceitáveis, considerando o número de indivíduos, em comparação com a identificação por mesmo número de indivíduos. A classificação por IMC apresentou um resultado abaixo do esperado para atributos cinemáticos e dentro do esperado para atributos antropométricos, já que não tem influência direta com a massa corporal da pessoa.

Os resultados foram comparados diretamente com trabalhos estado-da-arte em caminhar humano, que possuem metodologias diferentes das propostas neste trabalho. A primeira comparação foi utilizando o banco de dados criado com métodos aplicados em bases do caminhar diferentes das obtidas com o sensor Kinect. A outra comparação deu-se comparando diretamente os resultados obtidos em trabalhos estado-da-arte e que também utilizam o sensor Kinect na captura do caminhar, com os obtidos no presente trabalho.

A partir dessa comparação foi possível entender que o desempenho da classificação dos atributos está diretamente relacionada ao banco de dados utilizado e à metodologia de extração desse atributos do caminhar. Os resultados obtidos utilizando a base de dados criada no presente trabalho com essas outras metodologias não foram satisfatórios, indicando diferenças influenciáveis na maneira que o caminhar humano é armazenado. Contudo, considerando que o uso do sensor Kinect é uma simplificação que diminui o custo de segmentação e extração de modelos do

caminhar de vídeos e os resultados obtidos com a metodologia proposta comparada aos métodos estado-da-arte são superiores ou semelhantes, a utilização do que foi proposto é relativamente viável, quando se busca rapidez, baixo custo e simplicidade.

De acordo com os resultados obtidos a partir do objetivo do presente trabalho, de analisar a viabilidade do uso do sensor Microsoft Kinect 360 (MICROSOFT, 2011) em sistemas biométricos, foi possível concluir que:

1. Sistemas que utilizam apenas os parâmetros espaço-temporais e cinemáticos do caminhar humano apresentados na presente metodologia: O grupo classificado não deve ultrapassar 50 indivíduos e é necessário capturar um número grande o suficiente de ciclos de caminhar (a partir de 10). Para esse tipo de captura, o sensor deve acompanhar algum dispositivo que permita o indivíduo permanecer mais tempo caminhando dentro de sua faixa de alcance.
2. Para aplicações que utilizam apenas antropometria como biometria: O número de indivíduos do grupo classificado pode aumentar consideravelmente, respeitando o número mínimo de 20 quadros capturados por indivíduo.
3. A combinação dos dois tipos de biometria: Deve levar em conta as observações apontadas para cada tipo de biometria, podendo diminuir o número de ciclos capturados para a obtenção dos parâmetros do caminhar humano.
4. Aspectos do funcionamento do sensor Kinect 360: A distância do indivíduo capturado de no máximo 3 metros deve ser respeitada, o ambiente de captura deve ser interno com iluminação natural controlada ou artificial.

Os resultados obtidos com a metodologia proposta podem ser considerados promissores quando a intenção é a elaboração de um sistema rápido de detecção de indivíduos, iniciando com a detecção da pessoa do vídeo, segmentação de um modelo e extração de atributos antropométricos e do caminhar. Sistemas biométricos que utilizam face, íris e voz não são aplicáveis quando não é possível proximidade ou contato com a pessoa. A biometria por face apresenta problemas em relação aos recursos utilizados para escondê-la, como máscaras simples ou complexas, balaclavas, bonés, ou simplesmente não mostrar o rosto para nenhuma câmera.

Até a conclusão deste trabalho não há conhecimento de nenhum trabalho que avalie a capacidade de um indivíduo mascarar seu caminhar para gerar um falso negativo ou imitar um caminhar conhecido para gerar um falso positivo, mas supondo que o mesmo consiga esconder qualquer padrão na maneira que caminha, a antropometria é um forte candidato no auxílio do reconhecimento biométrico.

Como trabalhos futuros, é proposto a investigação de outros parâmetros cinemáticos relevantes, bem como outros descritores que traduzam esses parâmetros em melhores atributos para a classificação dos indivíduos. Além disso, testes com o

uso de novas versões do sensor Microsoft Kinect, como o Kinect para Windows Versão 2 (MICROSOFT, 2014) que aprimora alguns aspectos da versão utilizada no presente projeto: precisão, responsividade, câmera colorida com melhor resolução (1080p), 25 articulações para até 6 indivíduos rastreados com maior precisão, ao mesmo tempo (MICROSOFT, 2014). Finalmente, trabalhos futuros deverão investigar a tolerância do sistema biométrico que utilize caminhar humano quanto à fraudes: mascarar ou mudar um caminhar e reproduzir um caminhar existente.

Ficam registrados no presente trabalho os agradecimentos ao Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPQ - processo 477937/2012-8.

## REFERÊNCIAS

ALBRIZIO, A. Biometry and anthropometry: from Galton to constitutional medicine. **Journal of Anthropological Sciences**, Roma, v.85, p.101–123, 2007.

ANDERSEN, M.; JENSEN, T.; LISOUSKI, P.; MORTENSEN, A.; HANSEN, M.; GREGERSEN, T.; AHRENDT, P. Kinect Depth Sensor Evaluation for Computer Vision Applications. **Electrical and Computer Engineering Technical report ECE-TR-6**, Universidade Aarhus, Dinamarca, 2012.

AZIMI, M. Skeleton joint smoothing white paper. **MSDN digital library**, Microsoft Inc., 2012.

BRADSKI, D. G. R.; KAEHLER, A. **Learning Opencv, 1st Edition**. 1.ed. Ohio: O'Reilly Media, Inc., 2008.

BURGES, C. A Tutorial on Support Vector Machines for Pattern Recognition. **Data Mining and Knowledge Discovery**, Boston, v.2, n.2, p.121–167, 1998.

CHANG, Y.; CHEN, S.; HUANG, J. A Kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities. **Research in developmental disabilities**, Louisiana, v.32, n.6, 2011.

CUNADO, D.; NIXON, M. S.; CARTER, J. N. Using gait as a biometric, via phase-weighted magnitude spectra. **Audio and Video-based Biometric Person Authentication**, New York, p.93–102, 1997.

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

FRAZER, P. Identification of Human Beings By the System of Alphonse Bertillon. , Pennsylvania, v.99, n.3, p.172–176, Mar. 1875.

GALLAGHER, S. Why facial recognition tech failed in the Boston bombing manhunt. **Ars Technica Technology Lab**, Chicago, Illinois, 2013.

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

HALL, M. A. Correlation-based Feature Subset Selection for Machine Learning. **Ph.D. Dissertation**, Hamilton, 1998.

HALL, M.; FRANK, E.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P.; WITTEN, I. H. The WEKA Data Mining Software: An Update. **SIGKDD Explorations**, Hamilton, v.11, 2009.

HAN, J.; BHANU, B. Individual recognition using gait energy image. **IEEE transactions on pattern analysis and machine intelligence**, Alamitos CA, v.28, n.2, p.316–22, Feb. 2006.

HAN, J.; KAMBER, M. Data mining: concepts and techniques. **San Francisco, CA, itd: Morgan Kaufmann**, Illinois, v.5, 2006.

HOFMANN, M.; BACHMANN, S.; RIGOLL, G. 2.5 d gait biometrics using the depth gradient histogram energy image. **Biometrics: Theory, Applications and Systems (BTAS), 2012 IEEE Fifth International Conference on**, Washington DC, p.399–403, 2012.

JAIN, A. K.; ROSS, A.; PRABHAKAR, S. An introduction to biometric recognition. **IEEE Transactions on Circuits and Systems for Video Technology**, Chicago, IL, v.14, n.1, p.4–20, 2004.

JOSHI, D.; ANAND, S. Cyclogram and cross correlation: A comparative study to quantify gait coordination in mental state. **Journal of Biomedical Science and Engineering**, Beijing, v.03, n.03, p.322–326, 2010.

KATZ, L. Anti-surveillance mask lets you pass as someone else. **CNET Tech Culture**, Somerville MA, 2014.

MICROSOFT. **Xbox 360 Project Natal 101**. Último acesso: junho 2012, <http://www.microsoft.com/games/en-US/e3/pr/projecnatal%20fact%20sheet%20may%2009.rtf>.

MICROSOFT. Getting Started with the Kinect for Windows SDK Beta from Microsoft Research. **Microsoft Research**, Microsoft Inc., jul 2011.

MICROSOFT. **Kinect for Windows**. Último acesso: junho 2012, <http://www.microsoft.com/en-us/kinectforwindows>.

MICROSOFT. **Kinect for Windows v2**. Microsoft Inc. Último acesso: agosto 2014, <http://www.microsoft.com/en-us/kinectforwindows/meetkinect/features.aspx>.

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

NANDINI, C.; MUKHOPADHYAY, P.; TANMAY, T.; RANJAN, S. K.; ROY, S. An Efficient Human Identification Using Gait Analysis. **International Journal of Research and Reviews in Computing Engineering (IJRRCE)**, India, v.1, n.2, 2011.

NG, H.; TONG, H.; TAN, W.; ABDULLAH, J. Improved Gait Classification with Different Smoothing Techniques. **International Conference on Advanced Science, Engineering and Information Technology**, Tirunelveli, India, 2011.

NYDCJS. Origins of the New York State Bureau of Identification - Internet Edition. **The New York State Division of Criminal Justice Services**, New York, 1997.

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

PRIMESENSE. **Natural Interaction**. Último acesso: maio 2012, [http://www.primesense.com/category/reference\\_design](http://www.primesense.com/category/reference_design).

SCHWARTZ, A. Chicago's Video Surveillance Cameras: A Pervasive and Poorly Regulated Threat to Our Privacy. **Northwestern Journal of Technology and Intellectual Property**, Chicago IL, v.11, n.2, 2013.

SHUTLER, J. D.; GRANT, M. G.; NIXON, M. S.; CARTER, J. N. On a large sequence-based human gait database. In: **Applications and Science in Soft Computing**. Nottingham, UK: Springer, 2004. p.339–346.

SINGH, Y. N.; SINGH, S. K. A taxonomy of biometric system vulnerabilities and defenses. **International Journal of Biometrics**, Ottawa, v.5, n.2, p.137, 2013.

SIVAPALAN, S.; CHEN, D.; DENMAN, S.; SRIDHARAN, S.; FOOKES, C. Gait energy volumes and frontal gait recognition using depth images. **2011 International Joint Conference on Biometrics IJCB**, Washington DC, p.1–6, 2011.

SUDHA, L.; BHAVANI, R. Performance comparison of SVM and kNN in automatic classification of human gait patterns. **International Journal of Computers**, Salem Oregon, v.6, n.1, 2012.

SUTROP, M.; LAAS-MIKKO, K. From identity verification to behavior prediction: Ethical implications of second generation biometrics. **Review of Policy Research**, Danvers MA, v.29, n.1, p.21–36, 2012.

VAUGHAN, C. L.; DAVIS, B. L.; O'CONNOR, J. C. **Dynamics of human gait**. Pennsylvania State University: Human Kinetics Publishers Champaign, Illinois, 1992.

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

WANG, L. Some issues of biometrics: technology intelligence, progress and challenges. **International Journal of Information Technology and Management**, Itta Bena MS, v.11, n.1, p.72–82, 2012.

WEISSTEIN, E. W. **Correlation Coefficient**. MathWorld—A Wolfram Web Resource. Último acesso: maio 2014, <http://mathworld.wolfram.com/CorrelationCoefficient.html>.

WHO. **BMI classification**. World Health Organization. Último acesso: agosto 2014, <http://apps.who.int/bmi/index.jsp?introPage=intro3.html>.

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

YEOH, T.; TAN, W.; NG, H.; TONG, H.; OOI, C. Improved gait recognition with automatic body joint identification. In: **Visual Informatics: Sustaining Research and Innovations**. Selangor, Malaysia: Springer, 2011. p.245–256.

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

ZHENG, S.; ZHANG, J.; HUANG, K.; HE, R.; TAN, T. Robust View Transformation Model for Gait Recognition. **Proceedings of the IEEE International Conference on Image Processing**, Beijing, 2011.