

RODRIGO ALVES NUNES

**AVALIAÇÃO DE TÉCNICAS PARA O RECONHECIMENTO
DE PESSOAS PELA FORMA DE ANDAR (GAIT
RECOGNITION)**

CURITIBA

2010

RODRIGO ALVES NUNES

**AVALIAÇÃO DE TÉCNICAS PARA O RECONHECIMENTO
DE PESSOAS PELA FORMA DE ANDAR (GAIT
RECOGNITION)**

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre. Programa de Pós-Graduação em Informática, Setor de Ciências Exatas, Universidade Federal do Paraná.

Orientadora: Profa. Dra. Olga R. P. Bellon

Orientador: Prof. Dr. Luciano Silva

CURITIBA

2010

AGRADECIMENTOS

Agradeço primeiramente a Deus por ter me iluminado durante todo o período de desenvolvimento deste trabalho.

Agradeço aos meus orientadores Olga Regina Pereira Bellon e Luciano Silva pela oportunidade de trabalhar junto ao Grupo IMAGO de Pesquisa e por todo o apoio, paciência e compreensão durante o meu mestrado.

Aos meus pais, Edgard e Neusa, que sempre me incentivaram e apoiaram em todos os momentos da minha vida. Aos meus irmãos, Thiago e Edgard, pelas palavras de incentivo nos momentos em que eu mais precisava.

Aos grandes amigos que em Curitiba conquistei e que de alguma forma que seja colaboram com o desenvolvimento deste trabalho.

E por fim, agradeço ao CNPq pelo apoio financeiro durante meu mestrado.

SUMÁRIO

LISTA DE FIGURAS	viii
LISTA DE TABELAS	ix
RESUMO	x
ABSTRACT	xi
1 INTRODUÇÃO	1
2 BASES DE DADOS	5
2.1 UCSD Gait Data	5
2.2 Primeira versão da Soton Gait Data	7
2.3 Soton Database	8
2.4 CASIA Database	9
2.5 UMD Database	11
2.6 USF Database	11
3 ABORDAGENS DE RECONHECIMENTO	16
3.1 Baseline	19
3.2 Abordagens Implementadas	24
3.2.1 Averaged Silhouette	25
3.2.2 Width Vectors e Mass Vectors	27
3.3 Experimentos e Abordagens Desenvolvidas	31
3.3.1 Abordagem Averaged Width Vectors	32
3.3.2 Abordagem Averaged Mass Vectors	33
3.3.3 Adição Silhueta Média Mediana	35
3.3.4 Somente Silhueta Média Mediana	37
3.3.5 Média Medianas	38

3.3.6 Ciclos Intercalados	39
4 ANÁLISE DA QUALIDADE DAS SILHUETAS DA USF DATABASE	41
5 CONCLUSÃO	46
A RESULTADOS DAS FILTRAGENS PARA A ABORDAGEM <i>BASELINE</i>	53
B RESULTADOS DAS FILTRAGENS PARA A ABORDAGEM DE <i>SILHUETAS MÉDIAS</i>	57
C RESULTADOS DAS FILTRAGENS PARA A ABORDAGEM DE <i>VETORES DE LARGURA</i>	61

LISTA DE FIGURAS

1.1	Descrição do ciclo da forma de andar [33]	2
2.1	Quadro de exemplo de uma sequência de vídeo da <i>UCSD Gait Data</i> [21].	6
2.2	Imagem de exemplo da <i>Soton Gait Data</i> [9].	7
2.3	Imagens da base menor da <i>Soton Database</i>	9
2.4	Imagens do Conjunto A da <i>CASIA Database</i>	10
2.5	Imagens do Conjunto B da <i>CASIA Database</i>	10
2.6	Imagens do Conjunto C da <i>CASIA Database</i>	11
2.7	Imagens do primeiro (a, b, c, d) e segundo (e, f) conjunto de dados da <i>UMD Database</i>	12
2.8	Descrição do local das gravações da <i>USF Database</i> [38].	13
2.9	Imagens da <i>USF Database</i> com variação de superfície e câmera.	13
2.10	Possíveis combinações de variações para Maio e Novembro da <i>USF Database</i>	14
3.1	Fluxograma das principais etapas do <i>baseline algorithm</i>	19
3.2	Subtração de Fundo: (a) quadro de uma sequência da forma com o retângulo delimitador em vermelho, (b) imagem binária gerada no processo de subtração deste quadro e (c) silhueta extraída.	21
3.3	Exemplo de erros de segmentação de silhuetas que ocorrem no <i>baseline algorithm</i> . Na primeira linha (a-e) as imagens e na segunda linha (f-j) as silhuetas extraídas [38].	22
3.4	Periodicidade da forma de andar: (a) região de uma silhueta (quadrado vermelho) utilizada no cálculo da periodicidade da forma de andar e (b) gráfico da variação da quantidade de pixels classificados como de movimento dessa região.	23
3.5	Desempenho do <i>baseline algorithm</i> nos 12 experimentos [38].	24

3.6	Descrição do processo de geração das silhuetas médias a partir de uma sequência de silhuetas.	25
3.7	Amostras de silhuetas de um ciclo da forma de andar (a) e uma silhueta média (b) de um ciclo da forma de andar de um mesmo indivíduo [27]. . .	26
3.8	Taxas de reconhecimento alcançados pelo <i>Averaged Silhouette</i>	27
3.9	Exemplo de uma (a) silhueta e seus respectivos vetores de (b) largura e (c) massa.	27
3.10	Sinais de variação dos pixels classificados como de movimento de dois meio ciclos diferentes: (a) antes do alinhamento por <i>DTW</i> e (b) após o alinhamento.	29
3.11	Taxas de reconhecimento alcançados por pela abordagem de vetores de largura: (a) no período de maio e (b) nos períodos de maio e novembro do <i>HumanID Gait Challenge Problem</i>	30
3.12	Taxas de reconhecimento alcançados por pela abordagem de vetores de massa: (a) no período de maio e (b) nos períodos de maio e novembro do <i>HumanID Gait Challenge Problem</i>	31
3.13	Exemplo de uma (a) silhueta, (b) imagem binária de largura e (c) imagem média de imagens binárias de largura.	32
3.14	Taxas de reconhecimento alcançadas pela abordagem <i>Averaged Width Vectors</i>	33
3.15	Exemplo de uma (a) silhueta, (b) imagem binária de massas e (c) imagem média de imagens binárias de massa.	34
3.16	Taxas de reconhecimento alcançados pela abordagem <i>Averaged Mass Vectors</i>	35
3.17	Exemplo de um conjunto de silhuetas médias de uma sequência da forma de andar (a-f) e (g) uma silhueta média mediana gerada a partir desse conjunto.	36
3.18	Taxas de reconhecimento alcançados pela abordagem de silhuetas médias sem e com a adição de silhuetas médias medianas: (a) Experimentos A-F e (b) Experimentos G-H.	36

3.19	Taxas de reconhecimento alcançados pela abordagem de silhuetas médias sem e com o uso de apenas silhuetas médias medianas: (a) Experimentos A-F e (b) Experimentos G-H.	37
3.20	Taxas de reconhecimento alcançados pela abordagem de silhuetas médias sem e com a nova forma de cálculo de similaridade: (a) Experimentos A-F e (b) Experimentos G-H.	38
3.21	Taxas de reconhecimento alcançados pela abordagem de silhuetas médias sem e com a intercalação de ciclos: (a) Experimentos A-F e (b) Experimentos G-H.	40
3.22	Taxas de reconhecimento alcançados pela abordagem de silhuetas médias sem e com a intercalação e a nova forma de cálculo de similaridade: (a) Experimentos A-F e (b) Experimentos G-H.	40
4.1	Exemplos de silhuetas etiquetadas: (a) TAG_PERDA_PERNA, (b) TAG_PERDA_DUAS_PERNAS, (c) TAG_PERDA_CABECA, (d) TAG_SUB_VAZIA, (e) TAG_TODO_QUADRO_MOVIMENTO, (f) TAG_RUIDOS, (g) TAG_DEGRADACAO_1 e (h) TAG_DEGRADACAO_2.	42
4.2	Diagrama do processo de filtragem das listas de sequências.	43
A.1	Taxas de reconhecimento da abordagem <i>baseline</i> após filtragem da etiqueta TAG_PERDA_PERNA.	53
A.2	Taxas de reconhecimento da abordagem <i>baseline</i> após filtragem da etiqueta TAG_PERDA_DUAS_PERNAS.	53
A.3	Taxas de reconhecimento da abordagem <i>baseline</i> após filtragem da etiqueta TAG_PERDA_CABECA.	54
A.4	Taxas de reconhecimento da abordagem <i>baseline</i> após filtragem da etiqueta TAG_SUB_VAZIA.	54
A.5	Taxas de reconhecimento da abordagem <i>baseline</i> após filtragem da etiqueta TAG_ERRO_TODO_QUADRO_MOVIMENTO.	55

A.6	Taxas de reconhecimento da abordagem <i>baseline</i> após filtragem da etiqueta TAG_DEGRADACAO_1.	55
A.7	Taxas de reconhecimento <i>baseline algrithm</i> após filtragem da etiqueta TAG_DEGRADACAO_2.	56
B.1	Taxas de reconhecimento da abordagem de silhuetas médias após filtragem da etiqueta TAG_PERDA_PERNA.	57
B.2	Taxas de reconhecimento da abordagem de silhuetas médias após filtragem da etiqueta TAG_PERDA_DUAS_PERNAS.	57
B.3	Taxas de reconhecimento da abordagem de silhuetas médias após filtragem da etiqueta TAG_PERDA_CABECA.	58
B.4	Taxas de reconhecimento da abordagem de silhuetas médias após filtragem da etiqueta TAG_SUB_VAZIA.	58
B.5	Taxas de reconhecimento da abordagem de silhuetas médias após filtragem da etiqueta TAG_ERRO_TODO_QUADRO_MOVIMENTO.	59
B.6	Taxas de reconhecimento da abordagem de silhuetas médias após filtragem da etiqueta TAG_DEGRADACAO_1.	59
B.7	Taxas de reconhecimento da abordagem de silhuetas médias após filtragem da etiqueta TAG_DEGRADACAO_2.	60
C.1	Taxas de reconhecimento da abordagem de vetores de largura após filtragem da etiqueta TAG_PERDA_PERNA.	61
C.2	Taxas de reconhecimento da abordagem de vetores de largura após filtragem da etiqueta TAG_PERDA_DUAS_PERNAS.	61
C.3	Taxas de reconhecimento da abordagem de vetores de largura após filtragem da etiqueta TAG_PERDA_CABECA.	62
C.4	Taxas de reconhecimento da abordagem de vetores de largura após filtragem da etiqueta TAG_SUB_VAZIA.	62
C.5	Taxas de reconhecimento da abordagem de vetores de largura após filtragem da etiqueta TAG_ERRO_TODO_QUADRO_MOVIMENTO.	63

C.6	Taxas de reconhecimento da abordagem de vetores de largura após filtragem da etiqueta TAG_DEGRADACAO_1.	63
C.7	Taxas de reconhecimento da abordagem de vetores de largura após filtragem da etiqueta TAG_DEGRADACAO_2.	64

LISTA DE TABELAS

2.1	Visão geral das bases de dados da forma de andar	6
2.2	Conjunto de 12 experimentos da <i>USF Database</i>	15
4.1	Etiquetas utilizadas no processo de classificação da qualidade das silhuetas obtidas pelo <i>baseline algorithm</i>	41
4.2	Quantidade de sequências antes e após processo de filtragem em cada lista.	43
4.3	Total de silhuetas classificadas nas listas de teste e galeria para cada tipo de etiqueta.	44

RESUMO

A possibilidade de uso da forma de andar de seres humanos como característica biométrica para a identificação de indivíduos é o foco de estudo deste trabalho. Além de apresentar vantagens em relação a outras biometrias, como o reconhecimento por face, impressão digital ou íris, o reconhecimento de pessoas pela forma de andar possibilita a extração de características biométricas à distância de forma não invasiva, e não necessita de imagens de alta resolução.

O presente trabalho apresenta um estudo das principais e atuais abordagens de reconhecimento de pessoas pela forma de andar, livres de modelos (*model-free*). Foram exploradas as principais bases de dados utilizadas atualmente assim como os métodos estado da arte. Cada uma das bases apresenta diferentes variações nos ambientes das filmagens (interno ou externo), tipo de superfície, tipo de calçado, ângulos de câmera e a variação de datas de gravação das sequências. Dessa forma, pode ser medido qual a influência de cada uma destas variações no processo de reconhecimento de pessoas pela forma de andar.

São apresentadas em detalhes as etapas de funcionamento das abordagens de reconhecimento definidas como *baseline*, silhuetas médias, vetores de largura e vetores de massa. Modificações na abordagem de silhuetas médias foram propostas e os resultados obtidos são discutidos em detalhes.

Após a análise de qualidade das silhuetas extraídas da *USF Database*, foi efetuada a classificação de uma parcela das silhuetas da base de acordo com a presença de erros de segmentação. Por meio dessa classificação são apresentadas as taxas de reconhecimento obtidas após a remoção de cada tipo de erro da base, de forma a poder avaliar a influência destes erros nos resultados dos métodos implementados.

ABSTRACT

The potential of using the human gait as a biometric feature for identification is the goal of this work.

The human gait recognition allows the extraction of biometric features at distance in a non-invasive way without requiring high resolution images, and also presents advantages in comparison to other biometrics, such as face, fingerprint and iris recognition.

This work presents a study of the main current approaches of gait recognition, model-free. The main databases employed in state-of-the-art methods were exploited. Each database presents different environment variations, such as indoor or outdoor sets, flooring, shoes, viewpoints and time instants. So the influence of each variation in the gait recognition process can be evaluated.

The working steps of approaches defined as baseline, averaged silhouettes, width vectors and mass vectors is also presented. Modification in the averaged silhouettes approach were proposed, and the results are discussed in details.

After a quality analysis of the silhouettes extracted from USF Database, a portion of the resulting database silhouettes were classified according to the presence of segmentation errors. The recognition rates are presented after removing each error type of the database separately, in order to evaluate the influence of these errors in the results of the implemented methods.

CAPÍTULO 1

INTRODUÇÃO

O estudo da *forma de andar* ("gait" em inglês) de seres humanos para ser utilizado como biometria está adquirindo destaque entre as biometrias existentes, pois possibilita a identificação de indivíduos a distância e sem a necessidade de imagens de alta qualidade. Essas condições não são favoráveis à outras biometrias como o reconhecimento por face, impressão digital ou pela íris. Estudos de várias áreas como psicologia, medicina e até literatura demonstram indícios da possibilidades de que os seres humanos podem ser reconhecidos pela sua forma de andar.

De acordo com Nixon *et al.* [32], Shakespeare já demonstrava em vários de seus textos indícios da capacidade dos seres humanos em reconhecer outros seres pela sua forma de andar, como exemplo em *A Tempestade* "*High'st Queen of state, Great Juno comes; I know her by her gait*" e *Twelfth Night* "*wherein, by the color of his beard, the shape of his leg, the manner of his gait, the expresse of his eye, forehead, and complexion, he shall find himself most feelingly personated*".

Ainda segundo Nixon *et al.* [32], Aristóteles pode ser considerado um dos primeiros a estudar a forma de andar. Outros grandes nomes como Leonardo da Vinci e Galileu contribuíram indiretamente para o avanço das pesquisas, pois Leonardo da Vinci estudou vetores de força e Galileu iniciou a aplicação de seus conhecimentos da mecânica na biomecânica. Borelli (1608-1679) foi um dos primeiros a estudar a locomoção humana e estava interessado nos princípios mecânicos da locomoção. Posteriormente os irmãos Weber (1836) investigaram a forma de andar dos seres humanos, tanto o andar como o correr, utilizando instrumentos simples tentaram descrever o movimento dos membros inferiores relacionando-os ao movimento de um pêndulo. Porém era necessário a utilização de técnicas matemáticas mais avançadas e instrumentos mais precisos para que essas ideias pudessem ser comprovadas

Murray *et al.* [31] desenvolveram pesquisas na área da medicina com o objetivo de classificar as componentes da forma de andar dos seres humanos para o tratamento de pacientes com anormalidades patológicas. A intenção era classificar as componentes da forma de andar de pacientes com condições patológicas normais e descobrir suas faixas de variação para posteriormente comparar com as componentes e variações de pacientes com condições patológicas anormais [30]. Para efetuar a coleta de informações as pessoas envolvidas no estudo deveriam utilizar marcadores espalhados ao corpo, tática comum para estudos feitos em laboratórios, porém inadequada para propósitos de reconhecimento.

Alguns termos envolvidos ao ciclo da forma de andar são descritos na Figura 1.1. O ciclo da forma de andar (*gait cycle*) é definido como o espaço de tempo entre o momento em que o pé direito ou esquerdo de uma pessoa entra em contato com o solo durante o ato de andar até que este mesmo pé entre novamente em contato com o solo. Durante os ciclos as pernas podem ser classificadas em dois momentos: apoiada ao solo suportando o peso do corpo (*stance*) ou em movimento (*swing*).

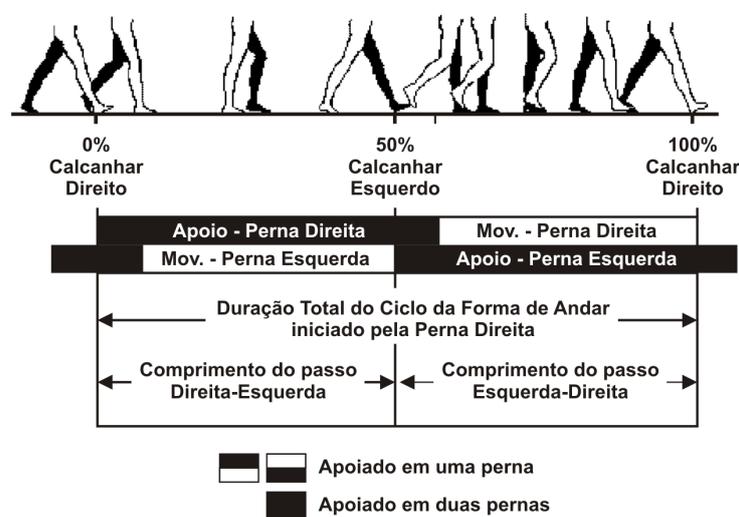


Figura 1.1: Descrição do ciclo da forma de andar [33]

De acordo com Nixon *et al.* [32], por meio do estudo feito por Murray *et al.* [31], pode ser concluído que se todos os movimentos da forma de andar forem levados em consideração, cada indivíduo possui uma forma de andar única. Porém algumas dessas componentes só podem ser mensuradas a partir de uma visão superior dos indivíduos, como exemplo os movimentos de rotação da pélvis e do tórax. Murray *et al.* [31]

demonstrou que os movimentos destas componentes, movimentos de rotação da pélvis e do tórax, são altamente variáveis de um indivíduo para outro entretanto não aparentam ser constantes em ciclos repetidos da forma de andar. Já os movimentos de rotação do tornozelo, inclinação da pélvis e o deslocamento espacial possuem uma consistência individual em ciclos repetidos da forma de andar, porém medir esses movimentos por meio de imagens reais (sem o uso de marcadores) é muito difícil [32].

Resultados obtidos por Johanson [22] em pesquisas na área da Psicologia demonstraram que os seres humanos possuem a capacidade de reconhecer por meio de pontos de luz em movimento o padrão de locomoção humano entre qualquer outro padrão de movimento. Cutting e Kozlowski [11] conseguiram demonstrar em seus estudos que além da capacidade de reconhecer padrões de movimento, eles conseguiam reconhecer amigos. Posteriormente outros experimentos mostraram que mais características poderiam ser extraídas tais como, gênero, direção do movimento e condições de transporte de objetos.

Estudos mais atuais [39] por meio do uso de vídeo ao invés de pontos de luz em movimento demonstraram que realmente os seres humanos possuem a capacidade de reconhecer pessoas pela sua forma de andar, e além disso memorizar esta forma para posteriores reconhecimentos. Também foi possível demonstrar que mesmo sobre condições adversas, como variação de iluminação ou de espaço de tempo exposto a uma sequência de vídeo, o reconhecimento ainda poderia ser alcançado.

Com o passar do tempo grandes avanços foram alcançados na área de visão computacional, possibilitando dessa forma análises mais aprofundadas do comportamento dos movimentos do corpo humano e por consequência o desenvolvimento de várias abordagens para a análise [1, 14, 18, 29, 41] e reconhecimento da forma de andar dos seres humanos [25].

O uso da forma de andar de seres humanos como característica biométrica para a identificação de indivíduos apresenta vantagens em relação a outras biometrias, como o reconhecimento por face, impressão digital ou íris, pois possibilita a extração de características biométricas à distância de forma não invasiva e não necessita de imagens de alta resolução. Estas vantagens são os principais motivadores para o desenvolvimento

deste trabalho.

O presente trabalho tem por finalidade apresentar um estudo sobre o uso da forma de andar dos seres humanos como uma característica biométrica na identificação de indivíduos, um levantamento das principais abordagens de reconhecimento desenvolvidas e as bases utilizadas para avaliação destas abordagens. Outro objetivo é propor modificações que melhorem os resultados de reconhecimento das abordagens estudadas.

Este trabalho está organizado conforme descrito a seguir. O Capítulo 2 apresenta quais foram as primeiras bases de dados utilizadas para o estudos da forma de andar e quais são as atuais bases utilizadas na avaliação das abordagens de reconhecimento. No Capítulo 3 são apresentadas os métodos de reconhecimento estudados, as etapas de processamento dos principais métodos e os resultados de experimentos desenvolvidos. No Capítulo 4 é feito uma análise da qualidade das silhuetas extraídas de uma parcela da *USF Database*. Após esta análise, são apresentadas as taxas de reconhecimento obtidas após a filtragem da base. Por fim, no Capítulo 5 são apresentadas as conclusões finais e os propostos trabalhos futuros.

CAPÍTULO 2

BASES DE DADOS

Inicialmente, para que estudos relacionados a análise da forma de andar dos seres humanos possam ser desenvolvidos, é necessário que exista previamente uma base de dados com várias sequências de imagens (vídeos) sobre o assunto.

As primeiras bases (*UCSD Gait Data* e a primeira versão da *Soton Gait Data*, ver Seções 2.1 e 2.2) de dados surgiram com a intenção de descobrir se o reconhecimento pela análise do andar poderia ser realmente alcançado. Já as atuais bases de dados (*Soton Database*, *CASIA Database*, *UMD Database* e *USF Database*, ver seções 2.3, 2.4, 2.5 e 2.6) foram desenvolvidas com o objetivo de serem usadas na avaliação dos métodos de reconhecimento. As atuais bases são maiores em relação a quantidade de indivíduos, apresentam variações de ambientes (interno e externo), variações e até segmentação manual de sequências de vídeo para uso em *ground truths*. A Tabela 2.1 apresenta uma visão geral sobre as bases de dados da forma de andar que serão abordadas neste capítulo.

2.1 UCSD Gait Data

A *UCSD Gait Data* [26] foi desenvolvida em 1998 pelo *Visual Computing Group* da *University of California San Diego*. Esta foi a primeira base de dados a ser disponibilizada e surgiu inicialmente composta por 5 indivíduos e 6 sequências de vídeo por indivíduo. Em seguida suas sequências foram estendidas de 5 para 6 por indivíduo e um novo indivíduo foi adicionado à base, este o único com 7 sequências de vídeo.

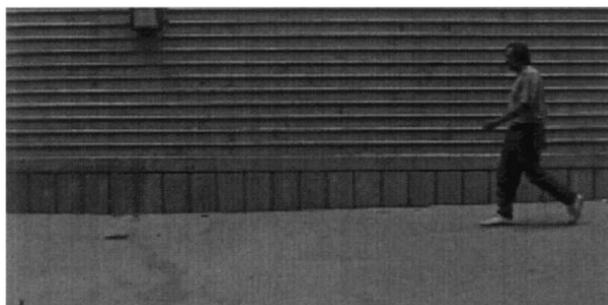
As gravações foram divididas em duas sessões de datas diferentes. Para a aquisição das imagens foi utilizada uma câmera Sony Hi8 que foi posicionada em direção a um muro de um pátio externo. As pessoas deviam andar em um caminho circular em torno da câmera e somente uma pessoa por vez deveria passar pelo campo de visão da câmera. A câmera foi posicionada de forma que as pessoas fossem filmadas andando com a sua

Tabela 2.1: Visão geral das bases de dados da forma de andar

Nome	Número de Indivíduos	Ambiente	Data	Variações
UCSD Gait Data	7	Externo	1998	–
Primeira versão da Soton Gait Data	10	Interno (<i>chromakey</i>)	1996	–
Soton Database	116	Externo, interno (<i>chromakey</i>) e em uma esteira ergométrica	2001	2 ângulos de visão (fronto paralela e inclinada)
	12	Interno	2001	Carregando mochila ou pasta, tipos de calçados, tipos de roupas, variação de velocidades e ângulos de visão (fronto paralela, fronto paralela elevada, inclinada e frontal)
CASIA Database	20	Externo	2001	3 ângulos de visão (fronte paralela, inclinada e frontal)
	124	Interno	2005	11 ângulos de visão, carregando mochilha e usando casaco
	153	Externo e noturno (infravermelho)	2005	velocidade (normal, rápido e lento) e carregando mochila
UMD Database	25	Externo	2001	2 ângulos de visão (frontal e fronte paralela)
	55	Externo	2001	2 ângulos de visão (frontal e fronte paralela)
USF Database	74	Externo	2001	2 ângulos de visão (direita e esquerda), 2 tipos de superfície e dois tipos de calçados
	122 (74+48)	Externo	2001	2 ângulos de visão (direita e esquerda), 2 tipos de superfície, dois tipos de calçados, carregando ou não pasta e dois intervalos de coleta das imagens

fronte paralela a posição da câmera, isto é, a direção em que as pessoas deveriam andar é perpendicular à direção em que a câmera estava posicionada. Durante as filmagens as pessoas percorreram o caminho circular durante aproximadamente quinze minutos e foram removidas as duas primeiras passagens em frente a câmera para aguardar a estabilização da câmera e também para que as pessoas estabeleçam um andar mais estável.

As sequências de vídeo da base possuíam aproximadamente 100 quadros a uma taxa de 30 quadros por segundo com 640x480 pixels de dimensão. A Figura 2.1 exemplifica um quadro de uma sequência de vídeo da base.

Figura 2.1: Quadro de exemplo de uma sequência de vídeo da *UCSD Gait Data* [21].

2.2 Primeira versão da Soton Gait Data

A primeira versão da *Soton Gait Data* surgiu em Southampton (*University of Southampton*) no Reino Unido em 1996 e foi desenvolvida para ser utilizada em estudos baseados na abordagem *model-based* de reconhecimento da forma de andar. Os vídeos foram obtidos por meio de uma câmera CCD ligada a um gravador de vídeo para posteriormente serem digitalizados. O posicionamento da câmera foi o mesmo adotado na *UCSD Gait Data*, as pessoas envolvidas nas filmagens passavam em frente a câmera com a sua frente paralela a posição da câmera, porém as filmagens desta base foram feitas em um ambiente fechado com um fundo de pano e a iluminação era controlada.

Para contornar problemas como a dificuldade de distinguir a perna direita da perna esquerda e a oclusão de uma perna pela outra, as pessoas a serem filmadas utilizavam uma calça especial com uma tarja preta em sua lateral (Fig. 2.2).



Figura 2.2: Imagem de exemplo da *Soton Gait Data* [9].

Foram gravadas dez sequências de cada pessoa passando em frente a câmera, porém as três primeiras e as três últimas sequências foram descartadas restando quatro sequências por pessoa. As primeiras e últimas sequências são descartadas na intenção de obter somente sequências onde as pessoas apresentem um andar mais estável e a uma velocidade constante. Mesmo sem definir um caminho circular ao qual as pessoas deveriam percorrer como foi feito na *UCSD Gait Data*, as sequências selecionadas conseguiram registrar o ato de andar a uma velocidade constante.

A *UCSD Gait Data* e a primeira versão da *Soton Gait Data* ainda estão disponíveis, mas atualmente estão superadas por bases mais atuais tanto na quantidade de indivíduos e também em termos de variações da forma de andar (fatores que exercem influência no desempenho do reconhecimento da forma de andar).

2.3 Soton Database

A *Soton Database*¹ possui sequências de vídeo de ambientes internos e externos e está dividida em duas bases: uma base com um grande número de pessoas (*large-subject database*) e outra com um número menor de pessoas (*small-subject database*). As sequências de vídeo foram armazenadas em formato de vídeo digital (DV) e algoritmos para capturar os quadros de cada sequências estão disponíveis em Python e C junto com a base.

A base maior possui 116 indivíduos e foi criada com a intenção de estudar o potencial da forma de andar como biometria e analisar as variações da forma de andar entre pessoas diferentes. As filmagens dessa base foram obtidas em três tipos de cenários: trajeto em ambiente externo, trajeto em ambiente interno e em uma esteira ergométrica.

O trajeto definido para o ambiente interno e externo são os mesmos, o indivíduo a ser filmado deveria percorrer um trajeto em linha reta e quando chegar ao seu final retornaria fazendo o caminho inverso. Porém para que o indivíduo não parasse quando fosse fazer o caminho inverso, ao final deste trajeto ele deveria retornar desenvolvendo um percurso circular, mantendo assim a sua forma de andar estável e contínua. Para o cenário da esteira ergométrica a pessoa a ser filmada deveria andar a uma velocidade de 4.1 km/h e a uma inclinação de 3° na intenção de se obter um andar natural.

Nos cenários de ambiente interno (trajeto em ambiente interno e em esteira ergométrica) foram utilizados refletores de luz para controlar a iluminação do ambiente e as paredes e o chão foram cobertos por um fundo de cor verde (*chromakey*). Todos esses cuidados foram tomados para reduzir a intensidade das sobras no ambiente das filmagens e para que os processos de subtração de fundo e extração de informação apresentem bons resultados. Já nas filmagens externas, não há controle de iluminação e outros fatores influenciam a qualidade dos resultados dos processos de extração de informação das sequências de vídeo. Esses fatores do mundo real (variações de iluminação, variações nas sombras, movimentos de galhos de árvores e objetos se movimentando ao fundo da cena) devem ser particularmente tratados para que melhores resultados sejam obtidos.

¹<http://www.gait.ecs.soton.ac.uk/data.php3> (Último acesso em 07/01/2009)

Em cada, cenário duas câmeras (Canon MV30i e Sony TRV900E) foram posicionadas em ângulos diferentes, uma perpendicular a direção em que as pessoas andavam (fronto-paralela a direção da câmera) e a outra a um ângulo mais inclinado.

A base menor possui 12 indivíduos e foi criada com a intenção de analisar a variação da forma de andar em situações diferentes, como carregando uma mochila, carregando uma pasta de documentos, usando calçados diferentes, usando roupas diferentes e andando a velocidades diferentes (Fig. 2.3). As filmagens dessa base foram obtidas em um ambiente interno, o mesmo utilizado na base maior com o mesmo trajeto, com iluminação controlada por refletores de luz e o uso de *chromakey*. Porém além das duas câmeras utilizadas nas filmagens da base anterior, mais duas câmeras foram adicionadas, uma posicionada na direção em que as pessoas andam (frontal) e a outra perpendicular a direção em que as pessoas andam a uma altura mais elevada (fronto-paralela elevada).

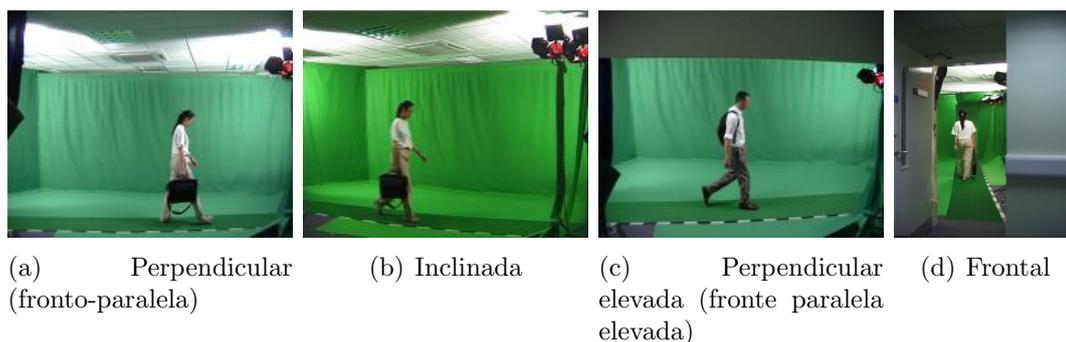


Figura 2.3: Imagens da base menor da *Soton Database*.

2.4 CASIA Database

A *CASIA Database*² foi desenvolvida pelo Instituto de Automação da Academia Chinesa de Ciências (CASIA). As pessoas a serem filmadas deveriam percorrer um caminho em linha reta, somente uma pessoa aparece em cada sequência de vídeo sem nenhuma ocorrência de oclusão. A base é composta por 3 conjuntos de dados: Conjunto A, Conjunto B e Conjunto C.

O *Conjunto A* foi gerado em 10 de dezembro de 2001 e possui 20 pessoas, no total de 240 sequências de vídeo, 12 sequências por pessoa (4 sequências para cada ângulo

²<http://www.sinobiometrics.com/english/Gait%20Databases.asp> (Último acesso em 07/01/2009)

de visão). As sequências foram capturadas em um ambiente externo a partir de três ângulos diferentes (Fig. 2.4): perpendicular ao caminho a ser percorrido (frente paralela), inclinado e frontal. Todas as sequências deste conjunto foram separadas em quadros e estão disponíveis no formato PNG.



Figura 2.4: Imagens do Conjunto A da *CASIA Database*.

O *Conjunto B* foi gerado em janeiro de 2005 e é composto por sequências de vídeo de 124 pessoas. As sequências de vídeo estão disponíveis no formato AVI e foram capturadas em um ambiente interno a partir de 11 ângulos diferentes. Além de possuir variações de ponto de visão este conjunto possui variações de tipo de roupa (casaco) e o carregamento ou não de uma mochila (carregando peso) (Fig. 2.5).

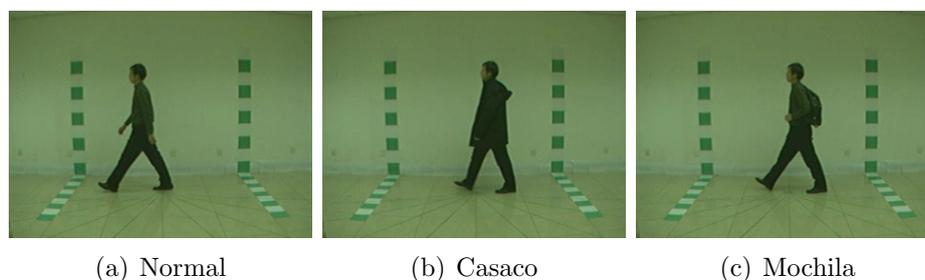


Figura 2.5: Imagens do Conjunto B da *CASIA Database*.

O último conjunto, *Conjunto C*, foi gerado durante os meses de julho e agosto de 2005 e é composto por 153 pessoas. As sequências de vídeo deste conjunto estão disponíveis em formato AVI e foram capturadas em um ambiente externo por meio de câmeras térmicas (infravermelho) durante a noite. Este conjunto possui quatro condições de variação da forma de andar (Fig. 2.6): andar normal, andar rápido, andar lento e carregando um mochila.

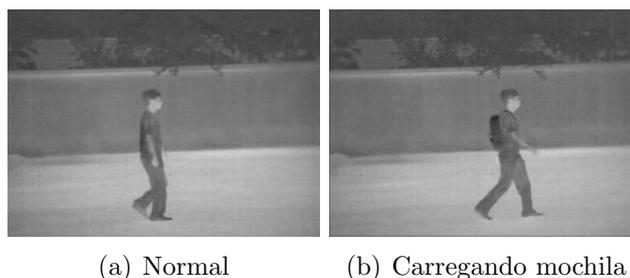


Figura 2.6: Imagens do Conjunto C da *CASIA Database*.

2.5 UMD Database

A *UMD Database*³ foi desenvolvida pela *University of Maryland* e é composta por dois conjuntos de dados. O primeiro conjunto possui 25 pessoas filmadas individualmente por uma câmera de vigilância (Phillips G3 EnviroDome) durante os meses de fevereiro e maio de 2001. Foram definidos quatro caminhos aos quais as pessoas deveriam percorrer durante as filmagens (Fig. 2.7a-d): frontal/ida, frontal/volta, lateral/direção direita-esquerda e lateral/direção esquerda-direita. O segundo conjunto é composto por 55 pessoas filmadas individualmente por duas câmeras de vigilância (Phillips G3 EnviroDome) nos meses de junho e julho de 2001. As câmeras foram ortogonalmente (90°) posicionadas uma em relação a outra a uma altura de 1,5 metros do solo. Em cada sequência de vídeo as pessoas foram filmadas por dois ângulos diferentes percorrendo um caminho em forma de “T” (Fig. 2.7e-f). Todas as sequências foram gravadas a uma resolução de 640×480 pixels.

2.6 USF Database

Conhecida como a base do *HumanID Gait Challenge Problem*, a *USF Database*⁴ foi gravada na *University of South Florida*. Inicialmente era composta de 452 sequências de vídeo de 74 pessoas, porém posteriormente foi expandida para 1870 sequências de vídeo de 122 pessoas. As gravações foram divididas em duas sessões de datas diferentes, a primeira em maio de 2001 e a segunda em novembro de 2001.

Um percurso elíptico com aproximadamente 15 metros de comprimento por 4 metros

³<http://www.umiacs.umd.edu/labs/pirl/hid/data.html> (Último acesso em 07/01/2009)

⁴<http://figment.csee.usf.edu/GaitBaseline/Data%20Set.htm> (Último acesso em 07/01/2009)

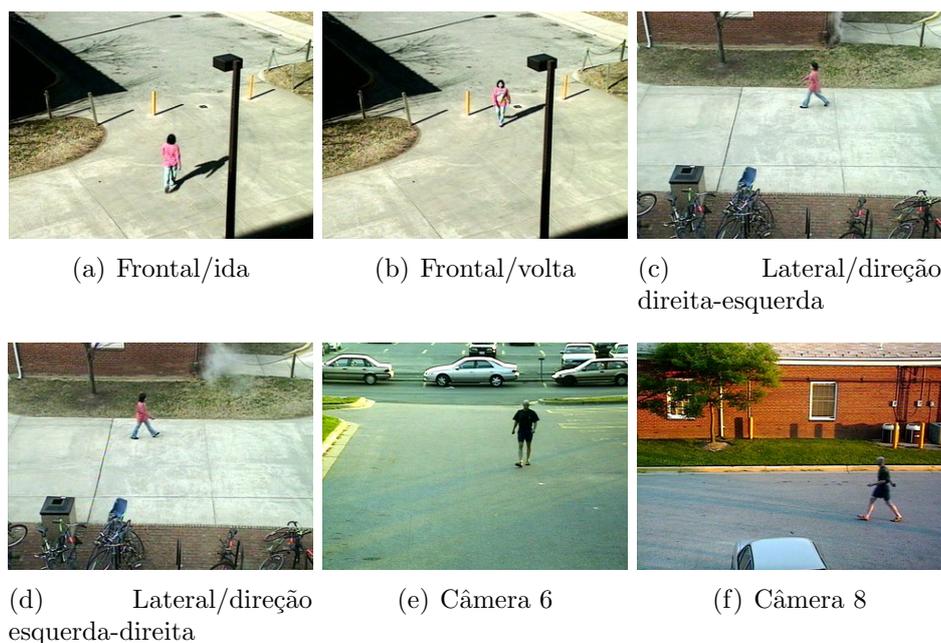


Figura 2.7: Imagens do primeiro (a, b, c, d) e segundo (e, f) conjunto de dados da *UMD Database*.

de largura foi definido como caminho ao qual as pessoas deveriam andar durante as gravações. Duas câmeras dispostas em ângulos diferentes foram utilizadas para filmar as pessoas enquanto andavam no percurso definido. Ao fundo da cena foram colocadas duas placas com o desenho de um tabuleiro de xadrez para serem usados em experimentos que necessitem da calibração das câmeras. Cada câmera foi posicionada a uma distância de aproximadamente 15 metros da extremidade mais próxima do percurso elíptico, de forma que todo o percurso estivesse em seu campo de visão (Fig. 2.8).

Todas as sequências foram gravadas em ambientes externos (Fig. 2.9) sobre a influência de movimentos no fundo das sequências e de sombras geradas por nuvens e pela luz do sol. Foi feita a escolha por ambientes externos por já existirem várias bases em ambientes internos e abordagens com desempenho de reconhecimento muito bons nessas bases. Outro motivo é o interesse em usar o reconhecimento pela forma de andar em ambientes externos e a distância, condições que não são favoráveis a outras biometrias.

A *USF Database* possui cinco variações, as quais foram escolhidas baseadas na hipótese de que exercem influência na forma de andar ou na extração de características da forma de andar a partir de imagens. As cinco variações foram selecionadas a partir de uma grande lista de variações, lista que surgiu após discussões entre vários pesquisadores do

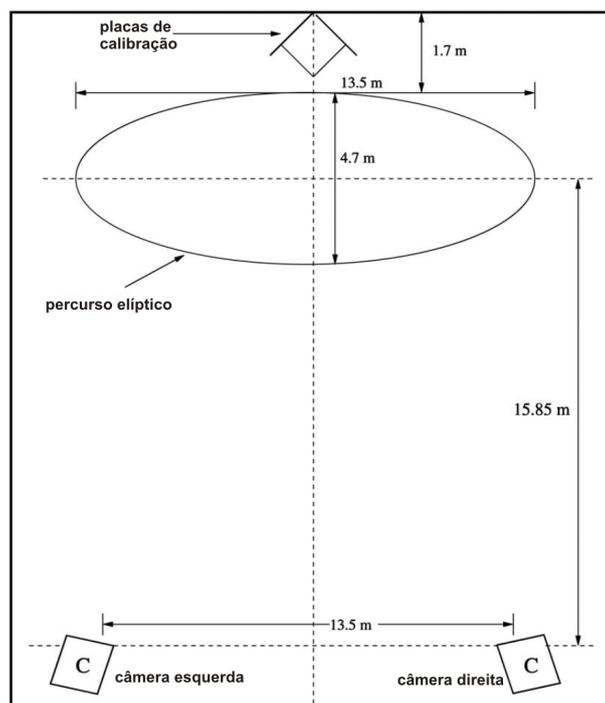


Figura 2.8: Descrição do local das gravações da *USF Database* [38].

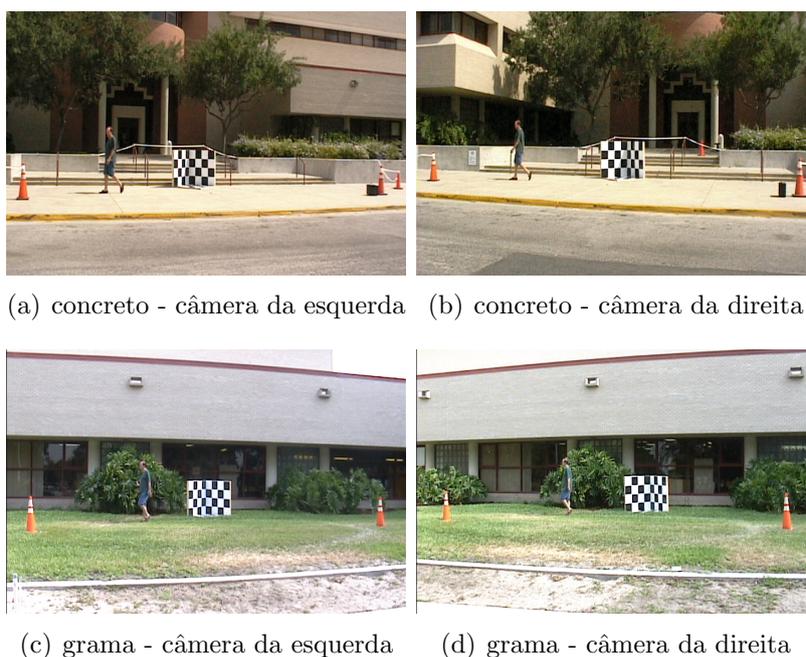


Figura 2.9: Imagens da *USF Database* com variação de superfície e câmera.

HumanID Program sobre a importância das variações na análise da forma de andar. Combinando as cinco variações existem 32 possíveis combinações diferentes, porém nem todos os indivíduos da base possuem sequências de vídeo de todas as combinações. Para cada variação existem duas possíveis variações (Fig. 2.10⁵): dois tipos de superfície

⁵<http://figment.csee.usf.edu/GaitBaseline/Data%20Set.htm> (Último acesso em 07/01/2009)

(concreto e grama), dois tipos de calçado, carregando ou não uma pasta de documentos (carregando peso), dois diferentes ângulos de câmera e as sequências foram coletadas em duas datas diferentes (intervalo de tempo de seis meses).

	Maio 2001				Novembro 2001				
	Não Carregando Pasta (NB)		Carregando Pasta (BF)		Não Carregando Pasta (NB)		Carregando Pasta (BF)		
Tênis A (A)	C,A,L, NB	G,A,L, NB	C,A,L, BF	G,A,L, BF	C,A,L, NB	G,A,L, NB	C,A,L, BF	G,A,L, BF	Câmera Esquerda (L) Câmera Direita (R)
Tênis B (B)	C,B,L, NB	G,B,L, NB	C,B,L, BF	G,B,L, BF	C,B,L, NB	G,B,L, NB	C,B,L, BF	G,B,L, BF	
Tênis A (A)	C,A,R, NB	G,A,R, NB	C,A,R, BF	G,A,R, BF	C,A,R, NB	G,A,R, NB	C,A,R, BF	G,A,R, BF	
Tênis B (B)	C,B,R, NB	G,B,R, NB	C,B,R, BF	G,B,R, BF	C,B,R, NB	G,B,R, NB	C,B,R, BF	G,B,R, BF	
	Concreto (C)	Grama (G)	Concreto (C)	Grama (G)	Concreto (C)	Grama (G)	Concreto (C)	Grama (G)	

Figura 2.10: Possíveis combinações de variações para Maio e Novembro da *USF Database*.

Além da grande quantidade de sequências de vídeo e variações da forma de andar, a *USF Database* possui para cada sequência da base um arquivo contendo as coordenadas da região (*bounding box*) onde se localiza a pessoa a ser analisada em cada quadro da sequência. A base também disponibiliza a rotulação manual de um ciclo da forma de andar de 72 indivíduos da base sob a influência de diferentes variações. As silhuetas foram rotuladas separando por meio de cores diferentes as seguintes partes do corpo: cabeça, tórax, braço direito, braço esquerdo, coxa esquerda, coxa direita, perna esquerda e perna direita. Por meio das rotulações manuais é possível comparar os resultados obtidos pelos segmentadores automáticos e também detectar possíveis erros gerados durante o processo de segmentação, além de possibilitar o estudo da relação entre o desempenho do reconhecimento e a qualidade da segmentação das silhuetas.

Com o objetivo de possibilitar a análise de desempenho das abordagens de reconhecimento da forma de andar o projeto *HumanID Gait Challenge Problem*⁶ disponibiliza um conjunto de 12 experimentos e um algoritmo de reconhecimento (*baseline*

⁶<http://www.gaitchallenge.org> (Último acesso em 12/07/2010)

algorithm) para ser usado na *USF Database*. O *baseline algorithm* é disponibilizado com a intenção de que seus resultados de desempenho nos 12 experimentos sejam utilizados como uma referência de comparação com outras abordagens de reconhecimento.

Os 12 experimentos estão classificados de acordo com o nível de dificuldade de reconhecimento, estes níveis são alcançados por meio da combinação de várias variações da forma de andar existentes na *USF Database* (Tab. 2.2⁷). Este conjunto de experimentos possibilita a avaliação das abordagens de reconhecimento na presença de cada uma das variações da base. Dessa forma, é possível analisar quais abordagens de reconhecimento mais adequadas a um determinado cenário controlado. Outra possibilidade é utilizar estes experimentos para estudar quais as componentes da forma de andar são alteradas após a presença de uma determinada variações.

Tabela 2.2: Conjunto de 12 experimentos da *USF Database*.

Experimento	Qtd. Pessoas		Variações
	Total	Somente Maio	
A	122	71	ângulo de visão
B	54	41	calçado
C	54	41	ângulo de visão e calçado
D	121	70	superfície
E	60	44	superfície e calçado
F	121	70	superfície e ângulo de visão
G	60	44	superfície, calçado e ângulo de visão
H	120	70	carregando pasta
I	60	47	calçado e carregando pasta
J	120	70	ângulo de visão e carregando pasta
K	33	33	variação de tempo (calçado e roupa)
L	33	33	superfície e variação de tempo

Dentre as bases apresentadas, a *USF Database* foi a escolhida para avaliar o desempenho das abordagens e os experimentos que serão apresentados neste trabalho. Esta base de dados foi escolhida por apresentar uma grande quantidade de indivíduos, variações e um conjunto definido de experimentos que possibilitam a avaliação das abordagens de reconhecimento com e sem a presença das variações citadas anteriormente.

⁷<http://marathon.csee.usf.edu/GaitBaseline/Experiments.htm> (Último acesso em 11/07/2010)

CAPÍTULO 3

ABORDAGENS DE RECONHECIMENTO

As abordagens de reconhecimento da forma de andar podem ser classificadas em dois tipos: baseados em modelos e livres de modelo. Os métodos baseados em modelos são caracterizados pela construção de um modelo prévio para descrever os movimentos do corpo humano. Este modelo é responsável em manter informações sobre a posição e trajetória dos membros humanos extraídos em cada quadro de uma dada sequência de vídeo. Um exemplo deste tipo de abordagem é apresentado por Cunado *et al.* [9] o qual descrevem a movimentação das pernas das pessoas como a movimentação de um pêndulo. As características utilizadas no reconhecimento são extraídas de componentes de frequência da variação de inclinação das pernas. Estas informações são obtidas com a ajuda da transformada de Fourier. Já Johnson and Bobick [4, 5], extraem parâmetros estáticos do corpo durante o ato de andar para serem utilizados no processo de reconhecimento. Nesta abordagem somente os parâmetros estáticos foram utilizados sem analisar diretamente a dinâmica dos padrões da forma de andar. Posteriormente, BenAbdelkader *et al.* [2, 3] utilizaram auto-similaridade e parâmetros estruturais (largura de passo e ritmo do andar) para descrever e reconhecer a forma de andar de indivíduos.

As abordagens livres de modelos se caracterizam por não necessitar da construção de um modelo prévio do corpo humano. Estes métodos analisam as sequências de imagens geradas durante o processo de caminhar, através da detecção de movimento e/ou forma que os indivíduos produzem, para extrair características necessárias ao reconhecimento. Uma das vantagens das abordagens baseadas em modelos em relação as livres de modelos é a imunidade a mudanças de roupa e de ponto de visão (ângulo de câmera), porém, por outro lado exige um grande custo computacional.

Lee e Grimson [25] apresentaram um método simples para a descrição da forma de andar. Este método está diretamente relacionado ao apresentado por Little e Boyd [26],

entretanto ao invés analisar as silhuetas por inteiro as silhuetas são divididas em regiões para serem processadas individualmente. Para cada uma das regiões da silhueta é feito o encaixe de uma elipse sobre a área classificada como de movimento e características como a variação do centro e largura das elipses são utilizadas no processo reconhecimento.

Phillips *et al.* [34,35,38] apresentaram o projeto *Human Gait Challenge Problem*. Este projeto é composto por uma grande base de dados sobre a forma de andar (*USF Database*), uma coleção de 12 experimentos classificados em ordem de dificuldade e um algoritmo de reconhecimento, o *baseline algorithm*. Este projeto tem como objetivo principal definir uma forma de mensurar o desempenho das abordagens de reconhecimento, possibilitando a comparação de desempenho das abordagens em cada um dos experimentos propostos. O *baseline algorithm* utiliza a correlação espaço-temporal de silhuetas para alcançar o reconhecimento da forma de andar.

Logo em seguida, Foster *et al.* [12, 13] apresentaram uma abordagem que combina a simplicidade do *baseline algorithm* com o cálculo de área (quantidade de pixels classificados como de movimento) de regiões específicas das silhuetas. Os sinais gerados pela variação da área em cada região são utilizados como uma assinatura da forma de andar.

Kale *et al.* [23] apresenta uma abordagem baseada no uso de vetores de largura (*width vectors*) de silhuetas binarizadas como uma característica básica para a representação da forma de andar. Devido a velocidade durante o ato de andar não ser constante de pessoa para pessoa e entre ciclos da mesma pessoa, foi utilizado o método de “dynamic time warping (DTW)”¹ para alinhar os vetores de largura antes de calcular a distância entre cada vetor.

Sundaresan *et al.* [40] desenvolveram uma abordagem que utiliza “Hidden Markov Models (HMM)” para gerar um modelo probabilístico para o reconhecimento da forma de andar. Nesta abordagem as silhuetas binarizadas são utilizadas para treinar e efetuar o reconhecimento. Já em Kale *et al.* [24], foram utilizados vetores de largura e as silhuetas

¹Dynamic Time Warping (DTW) é um método para calcular uma função não linear que alinha da melhor forma duas sequências temporais (sinais) de tamanhos diferentes. Essa função de ajuste também pode ser usada para calcular a similaridade entre duas sequências ou para encontrar regiões correspondentes à elas.

binarizadas para treinar e efetuar o reconhecimento.

Liu e Sarkar [27] apresentam uma abordagem de reconhecimento baseado no cálculo de silhuetas médias (*Averaged Silhouette*) a partir de uma sequência de silhuetas. Essa abordagem alcançou taxas de reconhecimento equivalentes as do *baseline algorithm* porém a uma velocidade muito maior. A mesma abordagem foi utilizada no trabalho de Veres *et al.* [8] que analisou o potencial das silhuetas médias para o reconhecimento e por Han *et al.* [15] o qual a definiu como *Gait Energy Image* (GEI).

Zhao *et al.* [43] apresentam um método de reconhecimento baseado no espectro de amplitude (*Fourier spectrum*) do domínio da frequência. Neste método, ao invés de utilizar todas as silhuetas de um ciclo da forma de andar, somente quatro silhuetas (silhuetas chaves) são utilizadas para descrever cada ciclo. Após extraídas as quatro silhuetas chaves é calculada a amplitude média do espectro de frequência de cada silhueta. Ao final as quatro amplitudes médias de cada ciclo são utilizados no processo de reconhecimento.

Han e Bhanu [16] apresentaram um método de reconhecimento que utiliza GEI [15] em conjunto com GEIs sintéticas. As GEIs sintéticas são geradas por meio da simulação da adição de ruídos e distorção as GEIs originais. O processo de reconhecimento é aprimorado após a utilização das GEIs originais em conjunto com as sintéticas. Ainda nesta abordagem, foram utilizadas duas técnicas de redução de dimensionalidade combinadas, “Principal Component Analysis (PCA)” com “Multiple Discriminant Analysis (MDA)” [21].

Hong *et al.* [19] desenvolveram um algoritmo de reconhecimento da forma de andar baseado no uso de vetores de massa (*mass vectors*). Essa abordagem foi inspirada na abordagem apresentada por Kale *et al.* [23] (vetores de largura) e também utiliza DTW para alinhar os vetores antes de calcular a distância entre cada vetor.

Inspirados por [16], Yang *et al.* [42] apresentaram uma abordagem que utiliza GEI [15], GEIs sintéticas e Gabor. Ao invés de utilizar as GEIs como característica principal para descrever a forma de andar, Yang *et al.* [42] utilizaram o período do espectro de Gabor. Posteriormente, utilizaram “Locality Preserving Projections (LPP)” [17] para projetar as

características em um subespaço menor.

Hong *et al.* [20] desenvolveram uma abordagem que utiliza ao invés de um conjunto de vetores de largura como apresentado por [23], apenas um vetor de largura médio (*width vector mean*) extraído de uma sequência de silhuetas. Após calculados os vetores de largura médios da base, os vetores são projetados em um subespaço menor por meio da técnica de redução de dimensionalidade PCA.

Na próxima seção será descrito em maiores detalhes a abordagem de reconhecimento *baseline algorithm*, bem como suas etapas de extração de fundo, extração das silhuetas, estimação do período da forma de andar e o cálculo de similaridade entre as sequências de silhuetas.

3.1 Baseline

O *baseline algorithm* é um algoritmo baseado no encaixe de modelos de silhueta (*silhouette template matching*) e o seu processo de funcionamento pode ser dividido em quatro etapas principais [38]: (1) definição dos retângulos delimitadores (*bounding box*) ao redor da pessoa em movimento em cada quadro de uma sequência; (2) extração das silhuetas de dentro de cada retângulo delimitador; (3) cálculo do período da forma de andar por meio das silhuetas extraídas e (4) cálculo de similaridade entre sequências da forma de andar por meio da correlação espaço temporal (*spatial-temporal*) entre as mesmas. A Figura 3.1 apresenta um fluxograma das etapas principais do *baseline algorithm*.



Figura 3.1: Fluxograma das principais etapas do *baseline algorithm*

A definição dos retângulos delimitadores (*bounding boxes*) é feita previamente de forma semi-manual e tem como função delimitar a região onde se localiza a pessoa em movimento em cada quadro de uma sequência da forma de andar. Inicialmente são definidos de forma manual três retângulos delimitadores em cada sequência: o inicial, o da posição central e o final. A partir destes três retângulos delimitadores são calculados todos os outros

restantes por meio da interpolação linear dos pontos superior esquerdo e inferior direito de cada retângulo delimitador [38].

Para extrair as silhuetas de uma sequência da forma de andar é necessário analisar a sequência duas vezes: uma para construir um modelo de fundo e outra para efetuar a subtração do fundo de cada quadro da sequência. O modelo de fundo composto pela média $\mu_B(x, y)$ e a covariância $\sum_B(x, y)$ das componentes RGB das posições (x, y) de cada quadro de uma sequência que não estejam dentro dos retângulos delimitadores demarcados anteriormente. Após construído o modelo de fundo inicia-se o processo de subtração de fundo de cada quadro da sequência, onde para cada pixel (x, y) de um quadro é calculado a distância de Mahalanobis [28] entre esse pixel e o modelo de fundo e o resultado é salvo em uma matriz de distâncias. Em uma versão mais antiga, a imagem binária (Fig. 3.2b) resultante do processo de subtração de fundo era gerada por meio de um limiar definido pelo usuário aplicado a matriz de distâncias. Já a versão atual não necessita da definição de um limiar pois utiliza um algoritmo de *Expectation Maximization* (EM) para classificar as distâncias da matriz de distâncias como fundo ou movimento [38].

Após a cada subtração de fundo, a imagem binária (Fig. 3.2b) é submetida aos seguintes processos: remoção de ruídos e extração da silhueta. A remoção de ruídos tem o objetivo de remover pequenas regiões de ruído que possam surgir após a subtração do fundo, dessa forma é mantido somente a maior região e todas as outras menores são removidas. Logo após é feito a extração e o redimensionamento da silhueta por meio de sua altura para que fique com o tamanho padrão de 128x88 pixels. Por último, a silhueta é a centralizada horizontalmente. O processo de redimensionamento torna a abordagem de reconhecimento mais resistente a variação de escala das silhuetas e a centralização contribuirá posteriormente facilitando o processo de similaridade. Ao final da etapa de subtração de fundo de uma sequência da forma de andar, as silhuetas extraídas são agrupadas em sequências de silhuetas. A Figura 3.2 apresenta a imagem binária (Fig. 3.2b) e a silhueta (Fig. 3.2c) extraídas de um quadro (Fig. 3.2a) de uma sequência da forma de andar da *USF Data base*.

A estratégia adotada no processo de extração das silhuetas, na maioria dos casos,



Figura 3.2: Subtração de Fundo: (a) quadro de uma sequência da forma com o retângulo delimitador em vermelho, (b) imagem binária gerada no processo de subtração deste quadro e (c) silhueta extraída.

apresentou bons resultados na qualidade das silhuetas. Porém, algumas silhuetas apresentaram erros de segmentação como os apresentados na Figura 3.3. Estes erros ocorreram devido aos seguintes fatores [38]:

1. a presença de sombras (especialmente em sequências filmadas em superfície de concreto);
2. incapacidade de segmentar regiões que possuem coloração muito próxima a cor do fundo, por isso são classificadas como fundo e são removidas;
3. presença de objetos em movimentos que pertencem ao fundo da sequência (folhas em movimento, pessoas andando ao fundo entre outros objetos);
4. ruídos gerados pelo formato de compressão de vídeo digital (*DV compression artifacts*²).

O ciclo da forma de andar, descrito na Seção 1 e ilustrado pela Figura 1.1, é definido como o espaço de tempo entre o momento em que o pé direito ou esquerdo de uma pessoa entra em contato com o solo durante o ato de andar até que este mesmo pé entre novamente em contato com o solo. Este ciclo sempre se repete a cada dois passos completos e pode se repetir várias vezes em uma sequência da forma de andar. Uma forma de calcular a periodicidade (N_{gait}) dos ciclos da forma de andar é analisar a variação da quantidade dos pixels classificados como de movimento $N_f(t)$ de cada silhueta de uma sequência de

²*Compression artifacts* são gerados após a compactação de dados com uso de algoritmos de compactação com perda de dados (*lossy data compression algorithms*).

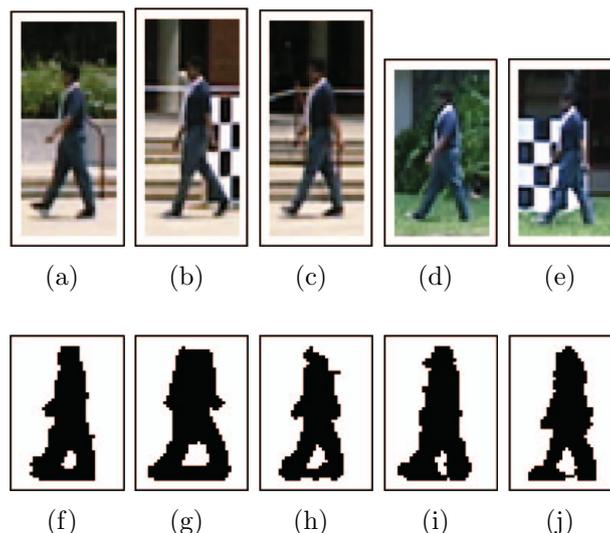


Figura 3.3: Exemplo de erros de segmentação de silhuetas que ocorrem no *baseline algorithm*. Na primeira linha (a-e) as imagens e na segunda linha (f-j) as silhuetas extraídas [38].

silhuetas. É possível identificar que essa variação possui um certo padrão cíclico (um pico seguido de um vale), apresenta valores mínimos (vales) nos momentos em que as pessoas estão com as pernas juntas (sobrepostas) e valores máximos (picos) nos momentos em que as pernas estão afastadas. Quando analisada somente a metade inferior das silhuetas (Fig. 3.4a), ou seja, calcular somente a variação dos pixels classificados como de movimento que estão contidos na região das pernas das silhuetas, a variação se apresenta menos ruidosa e mais precisa. A Figura 3.4 apresenta a região das silhuetas que é utilizada no cálculo da periodicidade da forma de andar e um gráfico da variação da quantidade de pixels classificados como de movimento dessa região de uma sequência de silhuetas da *USF Database*.

Com base no exposto acima, é possível calcular o tamanho de cada ciclo por meio da quantidade de silhuetas a cada dois vales consecutivos. Essa estratégia funciona para sequências filmadas fronto-paralela a direção da câmera ou com leve inclinação (caso da *USF Database*). Após calculados os tamanhos de cada ciclo de uma sequência de silhuetas, o valor mediano dos tamanhos é considerado como o período da forma de andar (N_{gait}).

Para calcular as taxas de similaridade as sequências de silhuetas extraídas nos processos anteriores são separadas em dois grupos: sequências da galeria (*gallery*) e sequências de teste (*probe*). O processo de reconhecimento consiste em calcular a

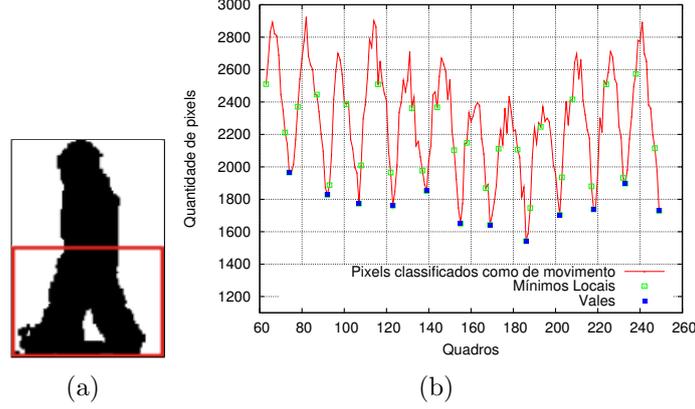


Figura 3.4: Periodicidade da forma de andar: (a) região de uma silhueta (quadrado vermelho) utilizada no cálculo da periodicidade da forma de andar e (b) gráfico da variação da quantidade de pixels classificados como de movimento dessa região.

similaridade entre as sequências de teste e todas as sequências da galeria. Uma sequência de teste de M quadros é descrita como $S_P = \{S_P(0), \dots, S_P(M-1)\}$ e uma sequência da galeria de N quadros é descrita como $S_G = \{S_G(0), \dots, S_G(N-1)\}$. As sequências de teste são particionadas em subsequências de N_{gait} quadros contíguos, onde N_{gait} é o período estimado do ciclo da forma de andar da sequência de teste. As k subsequências de teste são descritas como $S_{P_k} = \{S_P(kN_{gait}), \dots, S_P((k+1)N_{gait}-1)\}$.

O processo de cálculo de similaridade é dividido em três passos: cálculo de similaridade entre quadros de silhueta, cálculo de similaridade entre S_{P_k} e S_G e o cálculo de similaridade entre S_P e S_G . O cálculo de similaridade entre quadros de silhueta é obtido através da razão do número de pixels gerados pela intersecção de $S_P(i)$ e $S_G(j)$ pelo número de pixels gerados pela união de $S_P(i)$ e $S_G(j)$ (Eq. 3.1). O próximo passo é o cálculo da similaridade entre S_{P_k} e S_G por meio da similaridade entre quadros de silhuetas (Eq. 3.2) e por último o cálculo da similaridade entre a sequência de teste S_P e uma sequência da galeria S_G (Eq. 3.3).

$$FrameSim(S_P(i), S_G(j)) = \frac{Num(S_P(i) \cap S_G(j))}{Num(S_P(i) \cup S_G(j))} \quad (3.1)$$

$$Corr(S_{P_k}, S_G)(l) = \sum_{j=0}^{N_{gait}-1} FrameSim(S_P(k+j), S_G(l+j)) \quad (3.2)$$

$$Sim(S_P, S_G) = Median_k(\max_l Corr(S_{Pk}, S_G)(l)) \quad (3.3)$$

As taxas de identificação (*Cumulative Match Scores* [6]) do *baseline algorithm* na *USF Database* variaram de 3% (no experimento de maior dificuldade de reconhecimento) à 78% (no experimento de menor dificuldade de reconhecimento). A Figura 3.5 apresenta um gráfico com as taxas de reconhecimento do *baseline algorithm* nos 12 experimentos propostos pelo projeto *HumanID Gait Challenge Problem*.

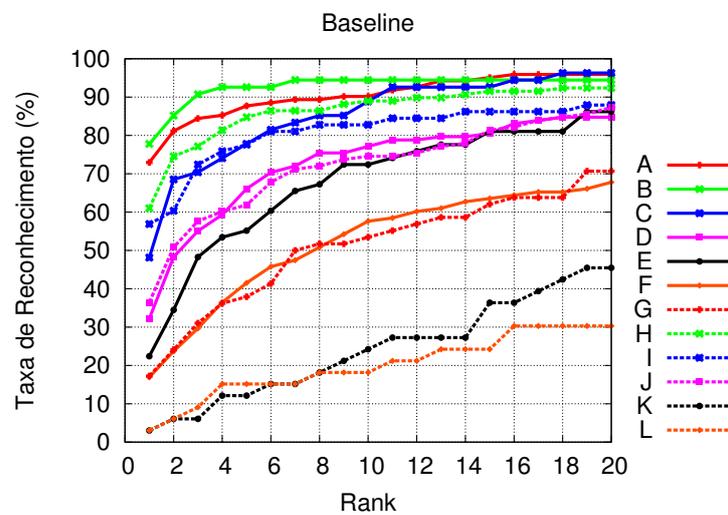


Figura 3.5: Desempenho do *baseline algorithm* nos 12 experimentos [38].

3.2 Abordagens Implementadas

A seguir serão apresentadas as abordagens implementadas neste trabalho e os resultados alcançados por cada uma. No total foram implementadas três abordagens de reconhecimento: silhuetas médias (ou GEI), vetores de largura e vetores de massa. O interesse em implementar essas abordagens é justificado devido ao uso dos seus conceitos em várias abordagens: abordagem de silhuetas médias em [15, 16, 27, 42] e vetores em [19, 20, 23, 24].

3.2.1 Averaged Silhouette

A abordagem de reconhecimento *Averaged Silhouette* (silhueta média) apresentada por Liu e Sarkar [27] é baseada no cálculo de silhuetas médias a partir de uma sequência de silhuetas. É uma abordagem simples e alcançou desempenhos quase equivalentes as do *baseline algorithm* porém a uma velocidade de reconhecimento muito maior.

O processo de subtração de fundo, redimensionamento das silhuetas e o cálculo do período do ciclo da forma de andar são os mesmos utilizados no *baseline algorithm* [38]. Após essas três etapas são calculadas as silhuetas médias a partir de uma sequência de silhuetas. Por meio do período do ciclo da forma de andar N_{gait} calculado anteriormente, uma sequência de silhuetas S com M quadros, descrita como $S = \{S(0), \dots, S(M - 1)\}$, é particionada gerando $S_{P_k} = \{S(k), \dots, S(kN_{gait} - 1)\}$ subseqüências. Para cada subseqüência S_{P_k} de S é gerado um silhueta média $AS(k)$, dessa forma a partir de S é gerado um conjunto de $\frac{M}{N_{gait}}$ silhuetas médias (Eq. 3.4). A Figura 3.6 apresenta um diagrama do processo de geração das silhuetas médias.

$$AS(i) = \frac{1}{N_{gait}} \sum_{k=iN_{gait}}^{(i+1)N_{gait}-1} S(k) \quad (3.4)$$

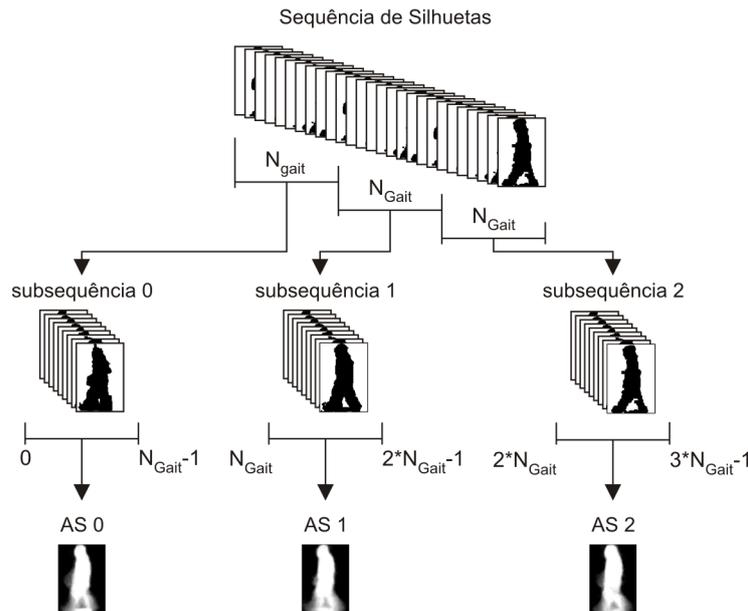


Figura 3.6: Descrição do processo de geração das silhuetas médias a partir de uma sequência de silhuetas.

O resultado da geração de uma silhueta média a partir de um ciclo da forma de andar é apresentado na Figura 3.7, em resumo a silhueta média é gerada pela média de todas as silhuetas de uma subsequência S_{Pk} .

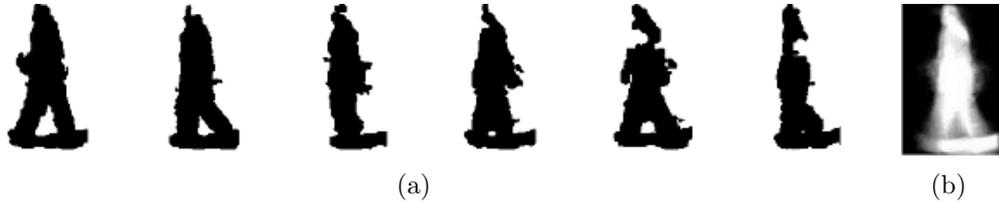


Figura 3.7: Amostras de silhuetas de um ciclo da forma de andar (a) e uma silhueta média (b) de um ciclo da forma de andar de um mesmo indivíduo [27].

Para calcular a similaridade entre uma sequência de silhuetas de teste e uma sequência de silhuetas da galeria, as suas respectivas silhuetas médias são utilizadas nesse processo. Dado um conjunto de silhuetas médias de teste AS_P com N_P silhuetas médias e um conjunto de silhuetas médias da galeria AS_G com N_G silhuetas médias, a similaridade entre esses dois conjuntos é obtida pela mediana negativa da distância Euclidiana entre as silhuetas médias de teste e da galeria (Eq. 3.5).

$$Sim(AS_P, AS_G) = -Median_{i=1}^{N_P} \left(\min_{j=1}^{N_G} \|AS_P(i) - AS_G(j)\| \right) \quad (3.5)$$

A abordagem de silhuetas médias alcançou taxas de reconhecimento equivalentes ao do *baseline algorithm* nos 12 experimentos propostos pelo *HumanID Gait Challenge Problem* (Fig. 3.8), porém a uma velocidade 30 vezes maior [27]. De acordo com Liu e Sarkar [27], em testes executados em um SunFire Server de 800 MHz a comparação de duas sequências por meio do *baseline algorithm* demora em média 4.63s, já pela abordagem de silhuetas médias 0.14s. Em alguns experimentos como A (ângulo de visão), B (calçado) e H (carregando ou não uma pasta de documentos) a abordagem de silhuetas médias apresentou taxas de reconhecimento maiores do que o *baseline algorithm* e apresentou taxas menores somente nos experimentos D (superfície) e K (tempo).

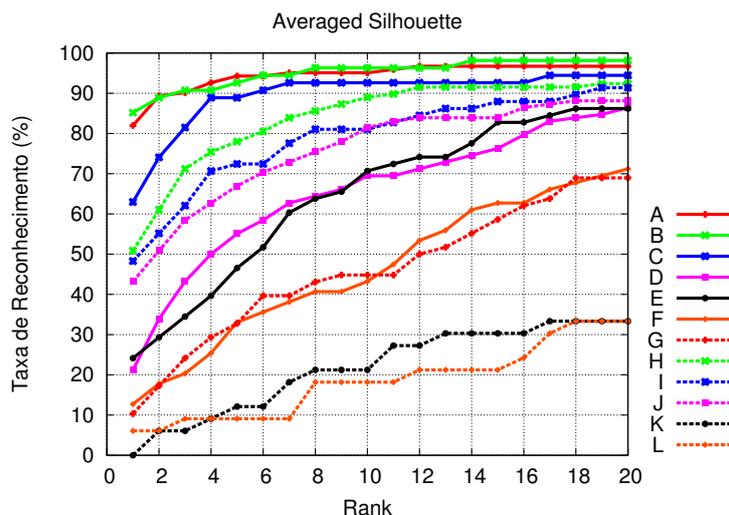


Figura 3.8: Taxas de reconhecimento alcançados pelo *Averaged Silhouette*.

3.2.2 Width Vectors e Mass Vectors

As abordagens *Width Vectors* [23] (vetores de largura) e *Mass Vectors* [19] (vetores de massa) são abordagens muito parecidas em relação ao processo de descrição dos ciclos da forma de andar, as duas utilizam vetores para descrever cada silhueta de uma sequência de silhuetas da forma de andar. A Figura 3.9 apresenta uma silhueta e seus respectivos vetores de largura e de massa.

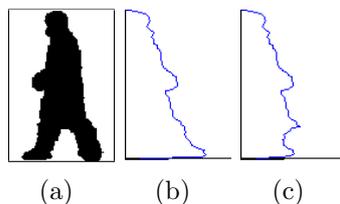


Figura 3.9: Exemplo de uma (a) silhueta e seus respectivos vetores de (b) largura e (c) massa.

Um vetor de largura WV possui M posições, cada uma delas referente a cada linha de uma silhueta S com M linhas e N colunas. Cada posição i de um vetor de largura $WV(i)$ registra a largura da linha i de uma silhueta. A largura de uma linha de uma silhueta é a distância entre o primeiro pixel classificado como de movimento e o último pixel classificado como de movimento dessa linha. Para descrever a forma de andar, são criados descritores WVS contendo os vetores de largura das silhuetas de cada meio ciclo da forma de andar.

Já um vetor de massa MV possui M posições, cada uma delas referente a cada linha de um silhueta S com M linhas e N colunas. Cada posição i de um vetor de massa $MV(i)$ registra a quantidade de pixels classificados como de movimento da linha i de uma silhueta. Para descrever a forma de andar, são criados descritores MVS contendo os vetores de massa das silhuetas de cada meio ciclo da forma.

Para reduzir a influência de ruídos durante o processo de geração dos vetores de largura e massa, cada vetor é suavizado por meio do uso do filtro da média.

Após calculados os descritores inicia-se o processo do cálculo de similaridade entre os descritores de cada abordagem. Devido as pessoas apresentarem velocidades diferentes durante o ato de andar os descritores gerados apresentaram quantidades diferentes de vetores. Desta forma, não é possível calcular diretamente a distância entre os vetores de um descritor em relação a outro descritor. Para resolver este problema, as duas abordagens utilizam o algoritmo *Dynamic Time Warping* (DTW) [37] para alinhar os vetores de cada descritor antes de calcular a distância entre eles. Após calculado o alinhamento (*warping path*), é definido como valor de similaridade entre os dois descritores o somatório das distâncias dos vetores alinhados.

Dado dois descritores D_A e D_B (descritores de vetores de largura ou massa) o algoritmo *DTW* alinha seus vetores por meio da construção da matriz de custo acumulado (*warping matrix*) e do cálculo do *warping path*. A matriz de custo acumulado M_{custo} possui A_n linhas e B_n colunas, onde A_n é a quantidade de vetores do descritor D_A e B_n é a quantidade de vetores do descritor D_B . Cada posição (x, y) de M_{custo} é definido como:

$$M_{custo}(x, y) = dist(D_A(x), D_B(y)) + mCusto(x, y, M_{custo}) \quad (3.6)$$

$$mCusto(x, y, M) = \begin{cases} 0 & \text{se } x = 1, y = 1 \\ M(x - 1, 1) & \text{se } x > 1, y = 1 \\ M(1, y - 1) & \text{se } x = 1, y > 1 \\ \min(M(x - 1, y), M(x, y - 1), M(x - 1, y - 1)) & \text{se } x > 1, y > 1 \end{cases}$$

onde $dist(D_A(x), D_B(y))$ é a distância Euclidiana entre os vetores $D_A(x)$ e $D_B(y)$ e $mCusto(x, y, M)$ é o menor custo adjacente de $M_{custo}(x, y)$. O *warping path*, $WP(D_A, D_B) = \{(A_n, B_n), \dots, (1, 1)\}$, é caminho de menor custo entre as posições extremas (A_n, B_n) e $(1, 1)$ da matriz de custo acumulado com o início em (A_n, B_n) . Então, para calcular a similaridade entre D_A e D_B são calculadas as distâncias Euclidianas entre os vetores listados pelo *warping path* calculado. O valor acumulado dessas distâncias é definido como o valor de similaridades entre os dois descritores. A Figura 3.10 apresenta os sinais de variação dos pixels classificados como de movimento das silhuetas de dois meios ciclos de pessoas diferentes, antes e depois do alinhamento utilizando o algoritmo de *DTW*.

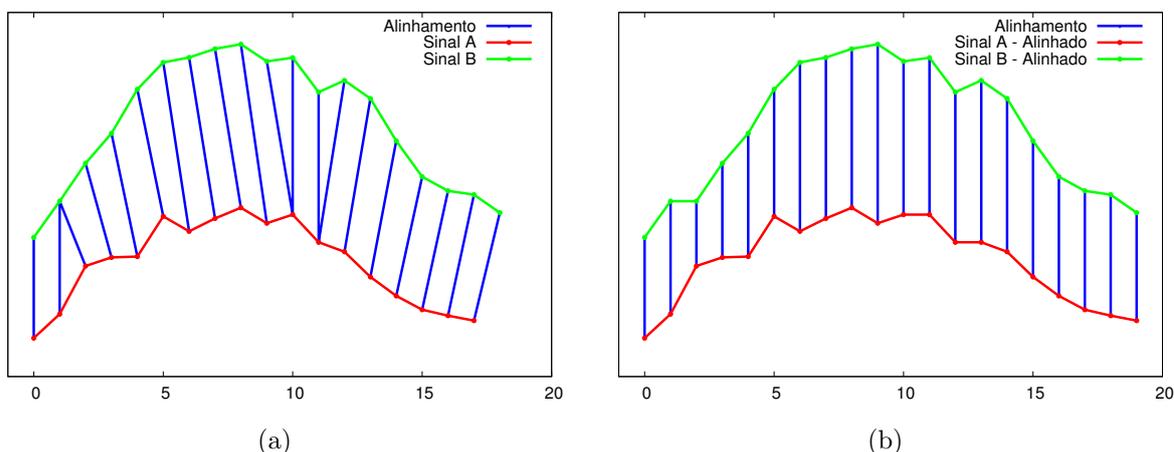


Figura 3.10: Sinais de variação dos pixels classificados como de movimento de dois meio ciclos diferentes: (a) antes do alinhamento por *DTW* e (b) após o alinhamento.

Para reduzir o grande consumo de processamento gasto no processo de cálculo da matriz de custo acumulado, os vetores de cada descritor são decompostos por meio do cálculo de seus auto valores e auto vetores. A partir deste momento, ao invés de calcular a matriz de custo acumulado por meio dos vetores de $D_A(x)$ e $D_B(y)$, a matriz será calculada por meio das projeções dos vetores de $D_A(x)$ e $D_B(y)$ no subespaço calculado para $D_A(x)$. Desta forma o processamento é reduzido devido a redução da dimensão dos vetores ao serem calculadas as distâncias da matriz de custo acumulado.

Dado um conjunto de descritores de teste $CD_P = \{D_P(1), \dots, D_P(N_P)\}$ com N_P descritores e um conjunto de descritores da galeria $CD_G = \{D_G(1), \dots, D_G(N_G)\}$ com

N_G descritores, a similaridade entre esses dois conjuntos é dada por:

$$Sim(CD_P, CD_G) = -Median_{i=1}^{N_P}(\min_{j=1}^{N_G} distWarp(D_P(i), D_G(j), WP(D_P(i), D_G(j)))), \quad (3.7)$$

onde $distWarp$ é o valor acumulado das distâncias Euclidianas entre os vetores de $D_P(i)$ e $D_G(j)$ listados pelo *warping path* $WP(D_P(i), D_G(j))$.

As Figuras 3.11 e 3.12 apresentam as taxas de reconhecimento alcançadas pelas abordagens implementadas de vetores de largura [23] e vetores de massa [19] nos experimentos do período de maio (7 experimentos com 71 pessoas na galeria) e nos experimentos dos períodos de maio e novembro (12 experimentos com 122 pessoas na galeria) propostos pelo *HumanID Gait Challenge Problem*. Para alcançar estes resultados os vetores foram suavizados com um mascara de tamanho 3 e projetados com o uso de sete auto vetores.

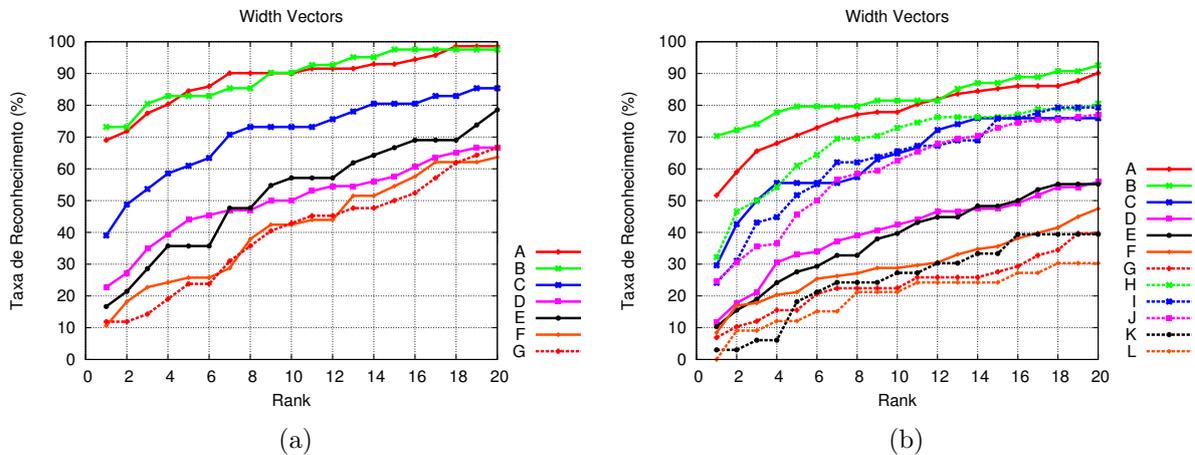


Figura 3.11: Taxas de reconhecimento alcançadas por pela abordagem de vetores de largura: (a) no período de maio e (b) nos períodos de maio e novembro do *HumanID Gait Challenge Problem*.

Os resultados obtidos pela abordagem de vetores de largura nos experimentos envolvendo somente o período de maio (Fig. 3.11a) são equivalentes aos apresentados em [23]. Já os resultados referentes aos períodos de maio e novembro (Fig. 3.11a) não foram comparados por não estarem presentes em [23]. Em [19] são apresentados somente os resultados da abordagem de vetores de largura no *Conjunto A* de imagens da *CASIA*

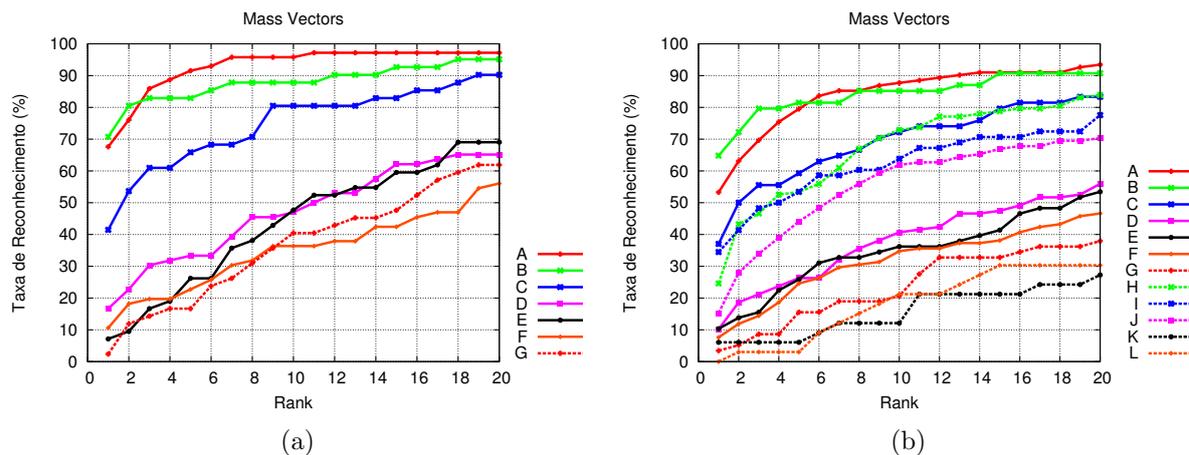


Figura 3.12: Taxas de reconhecimento alcançadas por pela abordagem de vetores de massa: (a) no período de maio e (b) nos períodos de maio e novembro do *HumanID Gait Challenge Problem*

Database (20 pessoas). Devido a essa razão não foi possível comparar os resultados originais com os obtidos pela abordagem implementada na *USF Database* (71 a 122 pessoas). O potencial das abordagens baseadas em vetores não deve ser desprezado. Em bases menores e menos ruidosas, as abordagens de vetores alcançaram taxas que variam de 80% a 95% em [23] e 88% a 96% em [19]. Já em [24] por meio do uso de HMM e vetores de largura, foram alcançadas taxas superiores as alcançadas pela abordagem de vetores de largura [23] e pelo *baseline algorithm* [38] nos experimentos do período de maio da *USF Database*.

3.3 Experimentos e Abordagens Desenvolvidas

Com a intenção de combinar características das abordagens estudadas para o desenvolvimento de novas abordagens de reconhecimento, vários experimentos foram desenvolvidos com o uso da *USF Database*. Duas novas abordagens de reconhecimento foram implementadas combinando características das abordagens de silhuetas médias, vetores de largura e vetores de massa. Posteriormente experimentos envolvendo modificações na estrutura de funcionamento da abordagem de silhuetas médias foram desenvolvidos e os resultados são apresentados a seguir. Os métodos de subtração de fundo e extração de silhuetas utilizados nas abordagens desenvolvidas e nos experimentos

são os mesmos apresentados no *baseline algorithm*.

3.3.1 Abordagem Averaged Width Vectors

A *Averaged Width Vectors* combina características dos descritores das abordagens de silhuetas médias (*Averaged Silhouettes*) e vetores de largura (*Width Vectors*). Os descritores desta abordagem são gerados por meio da construção de imagens médias de imagens binárias de largura de uma sequência de silhuetas. A Figura 3.13 apresenta um exemplo de imagem binária de largura (b) gerada a partir de uma silhueta (a) e a uma imagem média de imagens binárias de largura (c).

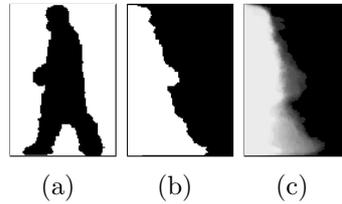


Figura 3.13: Exemplo de uma (a) silhueta, (b) imagem binária de largura e (c) imagem média de imagens binárias de largura.

Dada uma sequência de silhuetas S com M silhuetas, descrita como $S = \{S(0), \dots, S(M-1)\}$, é gerado para cada silhueta uma imagem binária de largura IBL . As imagens binárias de largura são construídas por meio do vetor de largura de cada silhueta. Cada linha i de uma imagem binária de largura IBL é preenchida da esquerda para a direita com n pixels de movimento, onde n é o valor de largura da posição i do vetor de largura de uma silhueta. Após geradas, as imagens binárias de largura IBL são agrupadas em sequência, $IBL = \{IBL(0), \dots, IBL(M-1)\}$, na mesma ordem da sequência de silhuetas S . Por meio do período forma de andar N_{Gait} as sequências de imagens binárias de largura são particionadas gerando $IBL_{PK} = \{IBL(k), \dots, IBL(kN_{Gait} - 1)\}$ subsequências. Finalmente, para cada subsequência IBL_{PK} de IBL é gerado uma imagem média de imagens binárias de largura $MIBL(k)$. Dessa forma, a partir de IBL é gerado um conjunto de $\frac{M}{N_{Gait}}$ imagens médias (Eq. 3.8), i.e.,

$$MIBL(i) = \frac{1}{N_{gait}} \sum_{k=iN_{gait}}^{(i+1)N_{gait}-1} IBL(k). \quad (3.8)$$

O processo de cálculo de similaridade é o mesmo utilizado na abordagem de silhuetas médias. A Equação 3.9 apresenta a equação de similaridade adaptada para esta abordagem. As taxas de reconhecimento obtidas pela abordagem nos 12 experimentos propostos pelo *HumanID Gait Challenge Problem* são apresentadas na Figura 3.14.

$$Sim(MIBL_P, MIBL_G) = -Median_{i=1}^{N_P} \left(\min_{j=1}^{N_G} ||MIBL_P(i) - MIBL_G(j)|| \right) \quad (3.9)$$

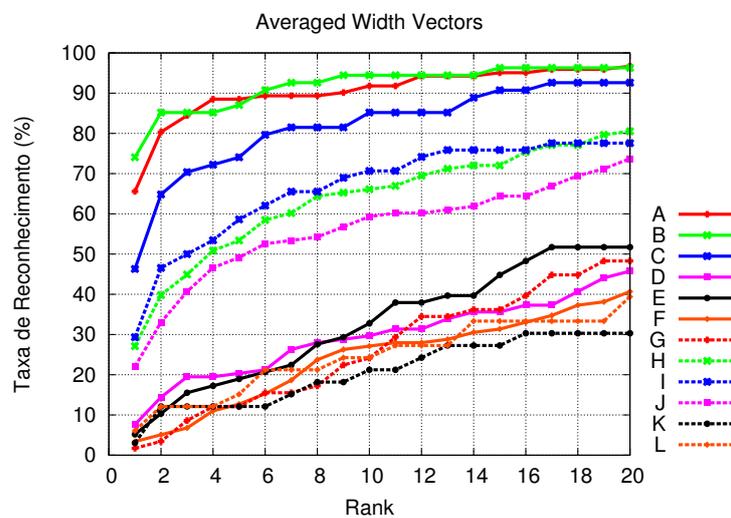


Figura 3.14: Taxas de reconhecimento alcançadas pela abordagem *Averaged Width Vectors*

Comparada a abordagem de vetores de largura, a abordagem implementada *Averaged Width Vectors* apresentou um aumento das taxas de reconhecimento de quatro experimentos: A (ângulo de visão), B (calçado), C (ângulo e calçado) e I (calçado e carregando pasta). Nota-se que o aumento ocorreu somente nos experimentos que envolvem a variação de ângulo de visão e calçado.

3.3.2 Abordagem Averaged Mass Vectors

A *Averaged Width Vectors* combina características dos descritores das abordagens de silhuetas médias (*Averaged Silhouettes*) e vetores de massa (*Mass Vectors*). Os descritores desta abordagem são gerados por meio da construção de imagens médias de imagens binárias de massa de uma sequência de silhuetas. A Figura 3.15 apresenta um exemplo de

imagem binária de massa (b) gerada a partir de uma silhueta (a) e a uma imagem média de imagens binárias de massa (c).

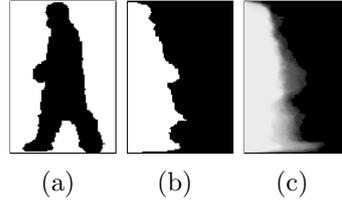


Figura 3.15: Exemplo de uma (a) silhueta, (b) imagem binária de masas e (c) imagem média de imagens binárias de massa.

Dada uma sequência de silhuetas S com M silhuetas, descrita como $S = \{S(0), \dots, S(M-1)\}$, é gerado para cada silhueta uma imagem binária de massa IBM . As imagens binárias de massa são construídas por meio do vetor de massa de cada silhueta. Cada linha i de uma imagem binária de massa IBM é preenchida da esquerda para a direita com n pixels de movimento, onde n é o valor de massa da posição i do vetor de massa de uma silhueta. Após geradas, as imagens binárias de massa IBM são agrupadas em sequência, $IBM = \{IBM(0), \dots, IBM(M-1)\}$, na mesma ordem da sequência de silhuetas S . Por meio do período forma de andar N_{Gait} as sequências de imagens binárias de massa são particionadas gerando $IBM_{PK} = \{IBM(k), \dots, IBM(kN_{Gait} - 1)\}$ subsequências. Ao final, para cada subsequência IBM_{PK} de IBM é gerado uma imagem média de imagens binárias de massa $MIBM(k)$. Dessa forma, a partir de IBM é gerado um conjunto de $\frac{M}{N_{Gait}}$ imagens médias (Eq. 3.10), i.e.,

$$MIBM(i) = \frac{1}{N_{gait}} \sum_{k=iN_{gait}}^{(i+1)N_{gait}-1} IBM(k). \quad (3.10)$$

O processo de cálculo de similaridade é o mesmo utilizado na abordagem de silhuetas médias. A Equação 3.11 apresenta a equação de similaridade adaptada para esta abordagem. As taxas de reconhecimento obtidas pela abordagem nos 12 experimentos propostos pelo *HumanID Gait Challenge Problem* são apresentadas na Figura 3.16.

$$Sim(MIBM_P, MIBM_G) = -Median_{i=1}^{N_P} \left(\min_{j=1}^{N_G} \|MIBM_P(i) - MIBM_G(j)\| \right) \quad (3.11)$$

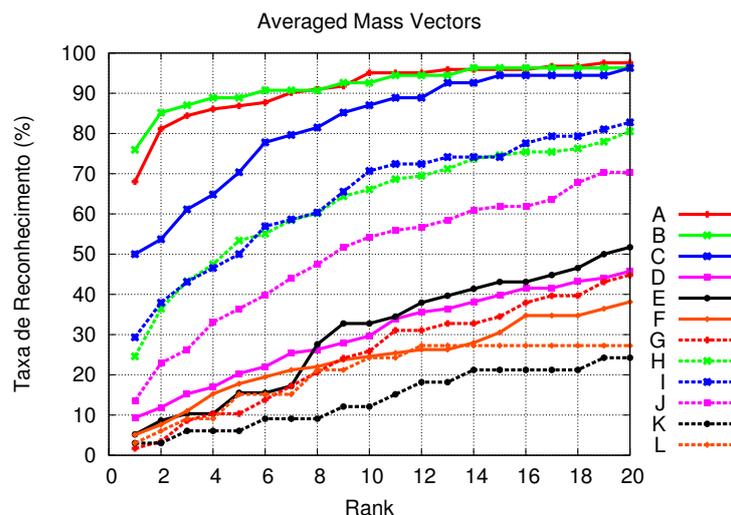


Figura 3.16: Taxas de reconhecimento alcançadas pela abordagem *Averaged Mass Vectors*

Comparada a abordagem de vetores de massa, a abordagem implementada *Averaged Mass Vectors* apresentou um aumento das taxas de reconhecimento de três experimentos: A (ângulo de visão), B (calçado), C (ângulo e calçado). Nas duas abordagens implementadas (*Averaged Width Vectors* e *Averaged Mass Vectors*) o aumento das taxas de reconhecimento ocorreram somente para os experimentos que envolvem a variação de ângulo de visão e calçado, já para os experimentos restantes há uma queda das taxas de reconhecimento das duas abordagens.

3.3.3 Adição Silhueta Média Mediana

O experimento “Adição Silhueta Média Mediana” é uma modificação aplicada a abordagem de silhuetas médias. Este experimento consiste em gerar uma silhueta média mediana (*Median Averaged Silhouette, MAS*) com base em um conjunto de silhuetas médias de uma mesma pessoa, e adicionar a silhueta ao conjunto de silhuetas médias. Esta ação tem o objetivo de gerar uma silhueta média com menos ruídos do que as que foram extraídas de uma sequência da forma de andar. Um exemplo desse processo é apresentado pela Figura 3.17.

Dado um conjunto de silhuetas médias AS com N silhuetas médias de uma mesma pessoa, onde $AS = \{AS(1), \dots, AS(N)\}$, o processo de geração da silhueta média mediana MAS para este conjunto consiste em calcular para cada posição i de MAS o valor mediano

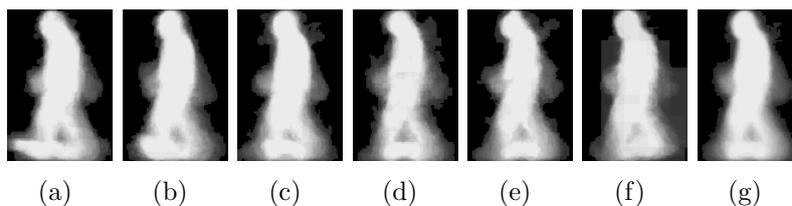


Figura 3.17: Exemplo de um conjunto de silhuetas médias de uma sequência da forma de andar (a-f) e (g) uma silhueta média mediana gerada a partir desse conjunto.

entre as posições i de cada silhuetas deste conjunto. Cada posição i de uma silhueta média mediana é denotado por:

$$MAS_i = \text{median}_{n=1}^N (f(i, AS(n))), \quad (3.12)$$

onde $f(i, AS)$ retorna o conteúdo da posição i de uma silhueta média AS . As taxas de reconhecimento alcançadas pela abordagem de silhuetas médias após esta modificação são apresentadas pela Figura 3.18. Nota-se que após a adição da silhueta média mediana a variação nas taxas de reconhecimento foram mínimas para a maioria dos 12 experimentos, as variações mais significativas ocorrem nos experimentos G (superfície, calçado e ângulo de visão), I (calçado e carregando pasta) e L (superfície e variação de tempo).

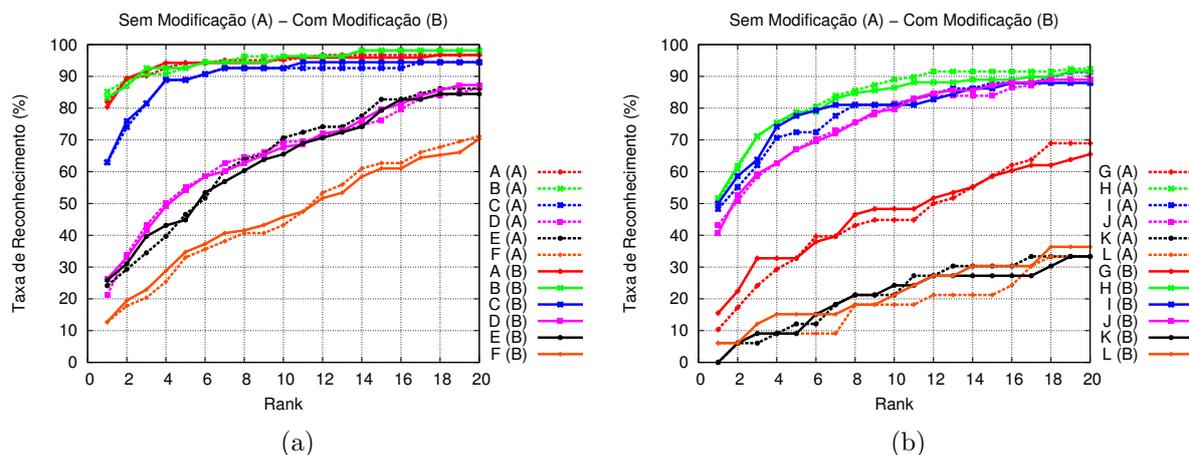


Figura 3.18: Taxas de reconhecimento alcançados pela abordagem de silhuetas médias sem e com a adição de silhuetas médias medianas: (a) Experimentos A-F e (b) Experimentos G-H.

3.3.4 Somente Silhueta Média Mediana

Este experimento consiste em substituir um conjunto de silhuetas médias de uma mesma pessoa pela silhueta média mediana *MAS* gerada para este conjunto. Esta ação tem como objetivo reduzir o ruído presente nas silhuetas médias e reduzir o número de comparações necessárias para calcular a similaridade entre duas pessoas, por consequência aumentar a velocidade do cálculo de similaridade. A velocidade do cálculo de similaridade aumenta devido a redução do número de comparações entre as silhuetas de $N_P \times N_G$ para apenas uma comparação, onde N_P é a quantidade de silhuetas médias de uma sequência de teste e N_G é a quantidade de silhuetas de médias de uma sequência da galeria. A Figura 3.19 apresenta as taxas de reconhecimento obtidas com o uso de somente uma silhueta média mediana para representar cada sequência da forma de andar. Nota-se que após o uso de somente silhuetas médias medianas a variação nas taxas de reconhecimento novamente foram mínimas para a maioria dos 12 experimentos, apresentando taxas inferiores somente nos experimentos D (superfície), E (superfície e calçado) e F (superfície e ângulo de visão). Devido a substituição de um conjunto de silhuetas médias por apenas uma silhueta média mediana, a velocidade do cálculo de similaridade entre duas sequências da forma de andar ficou em média 9 vezes mais rápida.

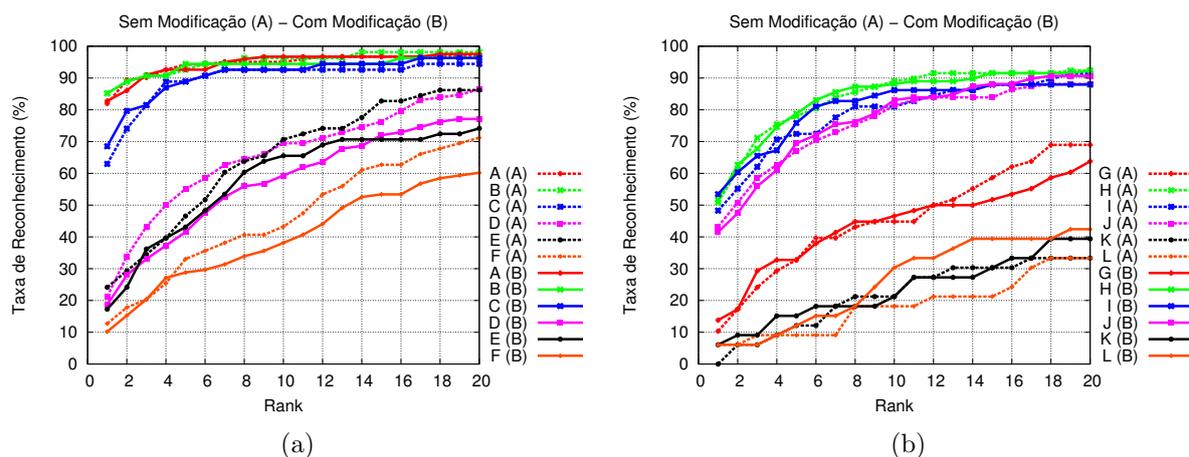


Figura 3.19: Taxas de reconhecimento alcançadas pela abordagem de silhuetas médias sem e com o uso de apenas silhuetas médias medianas: (a) Experimentos A-F e (b) Experimentos G-H.

3.3.5 Média Medianas

Este experimento consiste em alterar a forma que é calculado a similaridade entre as sequências da forma de andar. A Equação 3.5 descreve como é calculada a similaridade originalmente pela abordagem *Averaged Silhouettes* e a Equação 3.14 descreve qual a modificação realizada neste experimento. A ideia de modificar a forma que é calculada a similaridade surgiu devido que originalmente ao calcular $Sim(A, B)$ e $Sim(B, A)$ os resultados obtidos serem diferentes. Essa ideia foi implementada com base na abordagem apresentada em [7]. A Figura 3.20 apresenta as taxas de reconhecimento obtidas após alterar a forma de cálculo de similaridade. Após a modificação proposta, nota-se que as taxas de reconhecimento se elevam para a maioria dos experimentos.

$$SimA(AS_P, AS_G) = Median_{i=1}^{N_P} (\min_{j=1}^{N_G} ||AS_P(i) - AS_G(j)||) \quad (3.13)$$

$$SimB(AS_P, AS_G) = Median_{i=1}^{N_G} (\min_{j=1}^{N_P} ||AS_G(i) - AS_P(j)||)$$

$$SimFinal(AS_P, AS_G) = -\frac{1}{2}(SimA(AS_P, AS_G) + SimB(AS_P, AS_G))$$

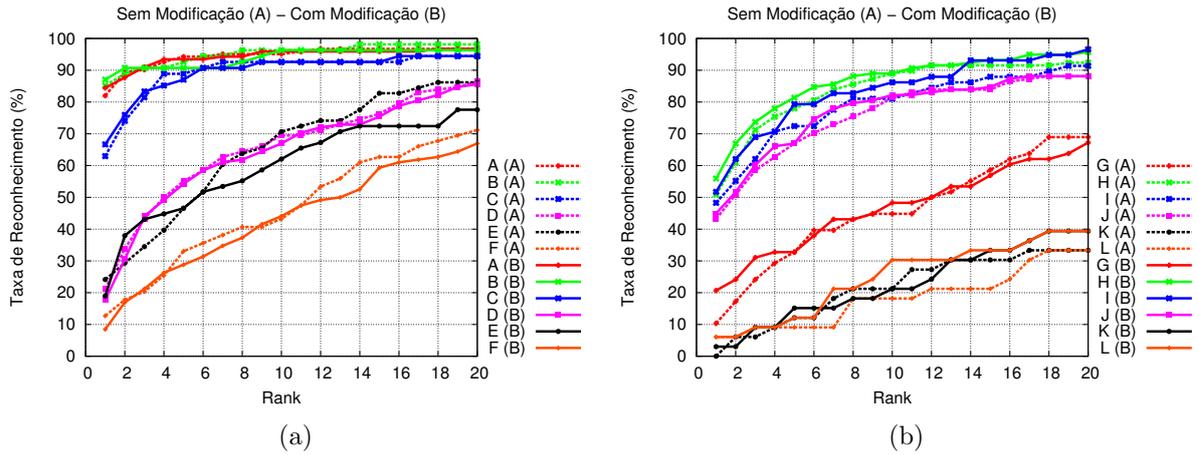


Figura 3.20: Taxas de reconhecimento alcançados pela abordagem de silhuetas médias sem e com a nova forma de cálculo de similaridade: (a) Experimentos A-F e (b) Experimentos G-H.

3.3.6 Ciclos Intercalados

Em [10, 13, 36, 38, 43] são relatados estudos sobre a dominância pedal³ (*foot dominance*) e uma possível relação de influência sobre as taxas reconhecimento e assimetria da forma de andar. Portanto existe a possibilidade de que pessoas possam apresentar uma certa variação na forma de andar de acordo com qual perna se inicia um ciclo da forma de andar.

De acordo com o exposto acima, este experimento consiste em calcular as silhuetas médias para os ciclos iniciados pela perna direita e esquerda em cada sequência de silhuetas. O processo de geração das silhuetas médias foi alterado e dividido em duas etapas: geração de silhuetas médias iniciadas pela perna X e geração de silhuetas médias iniciadas pela perna Y . O processo de geração de silhuetas médias para a perna X é o mesmo descrito na Seção 3.2.1 com o início da sequência de silhueta em $S(0)$. Já no processo de geração das silhuetas médias para ciclos iniciados pela perna Y , o início da sequência de silhuetas é alterado por meio da remoção das primeiras $\frac{N_{Gait}}{2} - 1$ silhuetas, dessa forma, o início da sequência a ser processada é alterado para $S(\frac{N_{Gait}}{2})$. Ao final serão geradas o total de $\frac{2N_{Gait}}{2} - 1$ silhuetas médias. As taxas de reconhecimento alcançadas pela abordagem após a modificação proposta é apresentado pela Figura 3.21. Nota-se que com exceção do experimento A, onde as taxas permaneceram idênticas, em todos os outros experimentos as taxas de reconhecimento aumentaram.

Após identificado o aumento das taxas de reconhecimento por meio do uso de ciclos intercalados, foi testado a fusão entre esta modificação e a proposta na Subseção 3.3.5. Os resultados obtidos são apresentados na Figura 3.22. Nota-se que as taxas de reconhecimento novamente são maiores em relação a abordagem original e também mais elevadas do que o teste anterior nos experimentos C (ângulo de visão e calçado), G (superfície, calçado e angulo de visão), H (carregando pasta), I (carregando pasta e calçado), J (carregando pasta e calçado) e L (superfície e tempo).

³Dominância pedal pode ser definida como a preferência em utilizar um determinado pé (direito ou esquerdo) no início de movimentos ou na manipulação de um objeto enquanto o outro pé é utilizado como apoio.

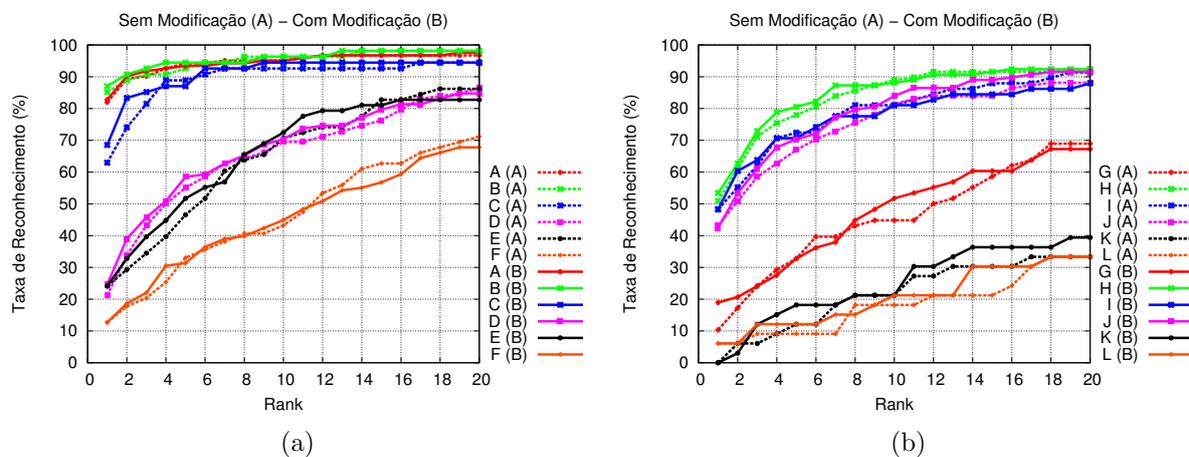


Figura 3.21: Taxas de reconhecimento alcançados pela abordagem de silhuetas médias sem e com a intercalação de ciclos: (a) Experimentos A-F e (b) Experimentos G-H.

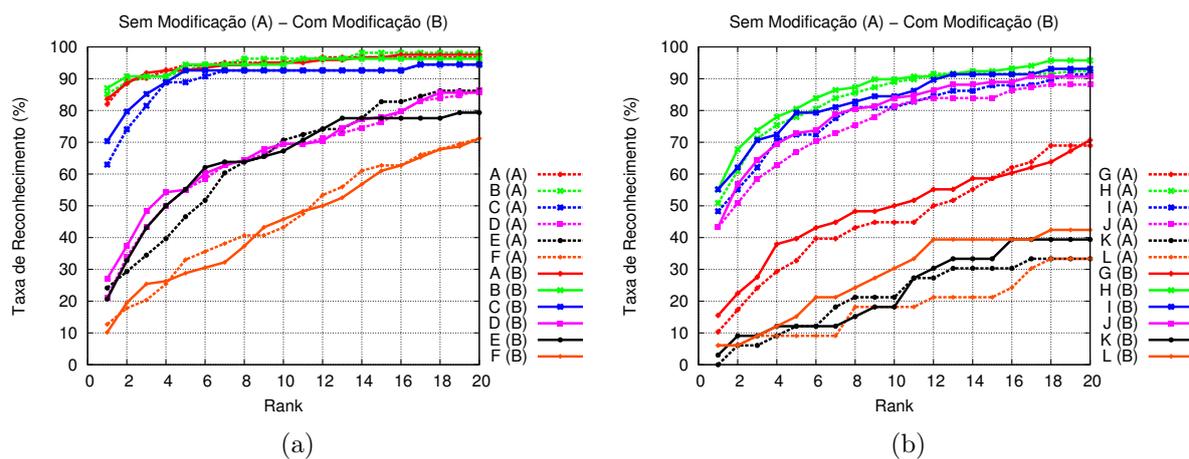


Figura 3.22: Taxas de reconhecimento alcançados pela abordagem de silhuetas médias sem e com a intercalação e a nova forma de cálculo de similaridade: (a) Experimentos A-F e (b) Experimentos G-H.

CAPÍTULO 4

ANÁLISE DA QUALIDADE DAS SILHUETAS DA USF DATABASE

Devido a um conjunto de fatores descritos anteriormente na Seção 3.1, uma parcela das silhuetas extraídas pelo processo de extração de silhuetas utilizado pelo *baseline algorithm* apresentam erros de segmentação. Para investigar a influência da presença desses erros sobre o processo de reconhecimento foi realizado uma análise de qualidade de um conjunto de sequências de silhuetas extraídas da *USF Database*.

Todas as sequências de silhuetas contidas na lista de sequências da galeria e na lista de teste do Experimento A (ProbeA, mudança de ângulo de visão) da *USF Database*, num total de 49900 silhuetas, foram analisadas com o objetivo de identificar quais os tipos de erros gerados pela influência dos fatores mencionados na Seção 3.1. Após identificados, para cada tipo de erro foi atribuído uma etiqueta de identificação. A Tabela 4.1 apresenta cada uma das etiquetas criadas e uma breve descrição do tipo de erro que ela representa. Já a Figura 4.1 apresenta exemplos de silhuetas etiquetadas para cada um dos erros identificados.

Tabela 4.1: Etiquetas utilizadas no processo de classificação da qualidade das silhuetas obtidas pelo *baseline algorithm*

Etiqueta	Descrição
TAG_PERDA_PERNA	Perda de uma das pernas
TAG_PERDA_DUAS_PERNAS	Perda das duas pernas
TAG_PERDA_CABECA	Perda da cabeça, como consequência redimensionamento incorreto da silhueta
TAG_SUB_VAZIA	Quadro com subtração vazia (nenhum movimento)
TAG_ERRO_TODO_QUADRO_MOVIMENTO	Todo ou quase todo o quadro é considerado como de movimento
TAG_RUIDOS	Ruídos gerados pela movimentação no fundo (adição de regiões de movimento no quadro que não pertencem a silhueta da pessoa)
TAG_DEGRADACAO.1	A forma da silhueta é degradada devido a regiões do fundo que possuem uma cor muito próxima as que a pessoa possui (perda de pedaços ou regiões pertencentes a silhueta da pessoa)
TAG_DEGRADACAO.2	A forma da silhueta é degradada devido a regiões do fundo que possuem uma cor muito próxima as que a pessoa possui (perda de pedaços ou regiões pertencente a silhueta da pessoa), a silhueta perde totalmente a sua forma devido ao redimensionamento da maior região resultante do processo de subtração

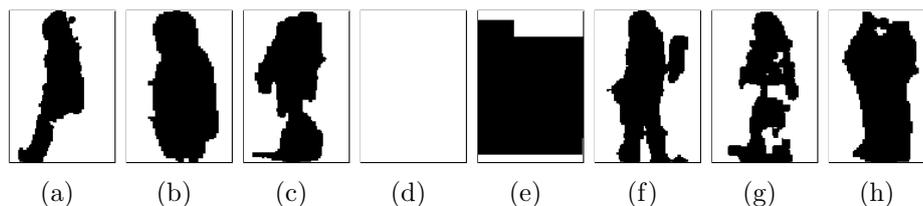


Figura 4.1: Exemplos de silhuetas etiquetadas: (a) TAG_PERDA_PERNA, (b) TAG_PERDA_DUAS_PERNAS, (c) TAG_PERDA_CABECA, (d) TAG_SUB_VAZIA, (e) TAG_TODO_QUADRO_MOVIMENTO, (f) TAG_RUIDOS, (g) TAG_DEGRADACAO_1 e (h) TAG_DEGRADACAO_2.

Um programa específico para auxiliar o processo de classificação foi desenvolvido, porém, todas as silhuetas foram classificadas de modo manual. Este programa foi desenvolvido com o objetivo de facilitar a visualização e padronização dos registros de classificação das silhuetas.

Após a classificação das silhuetas, são iniciados os processos de filtragem das listas de sequências da galeria e teste. Para cada etiqueta analisada são calculadas as taxas de reconhecimento de dois testes de filtragem:

1. T1 - Remoção de todas as sequências etiquetadas na lista da galeria e as sequências relacionadas na lista de teste;
2. T2 - Remoção de todas as sequências etiquetadas na lista da galeria e as sequências relacionadas na lista de teste; Remoção de todas as sequências etiquetadas na lista de teste da lista de teste

O “T1” tem o objetivo de calcular a taxa de reconhecimento alcançada após remover as sequências que foram classificadas com uma determinada etiqueta na galeria das listas da galeria e teste, dessa forma, garante que as sequências da galeria estão livres da presença de um determinado tipo de erro e a taxa alcançada neste teste será utilizada como referência após a próxima filtragem (T2). Já o “T2” tem o objetivo de remover da lista de teste as sequências que foram classificadas com a mesma etiqueta utilizada no “T1”, porém, sem remover suas sequências relacionadas da lista da galeria. Dessa forma será possível identificar o quanto a presença de um determinado erro influencia as taxas de reconhecimento por meio da comparação da taxa de reconhecimento do “T1” com a

do “T2”. A Figura 4.2 apresenta um diagrama do processo de filtragem das listas e o conteúdo de cada uma após cada teste de filtragem.

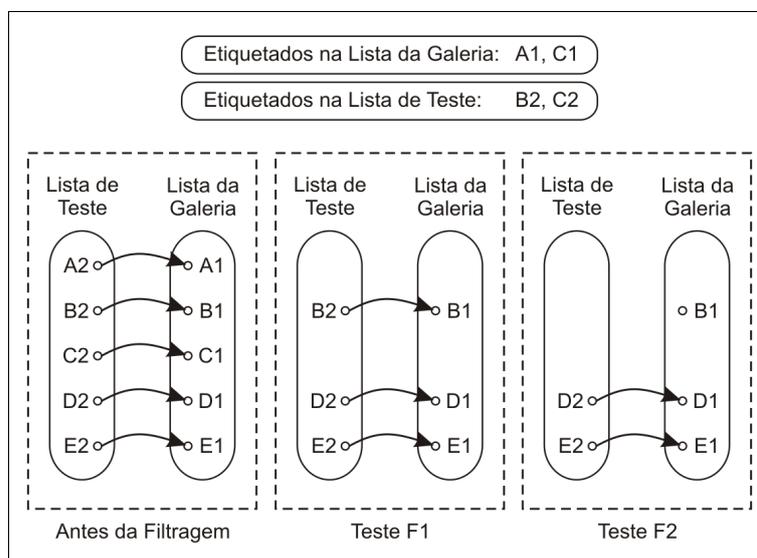


Figura 4.2: Diagrama do processo de filtragem das listas de seqüências.

A Tabela 4.2 apresenta a quantidade de seqüências nas listas de teste e galeria antes e após o processo de filtragem de cada etiqueta. Os tipos de erros mais encontrados nas duas listas foram TAG_DEGRADACAO_1, TAG_DEGRADACAO_2, TAG_PERDA_PERNA E TAG_RUIDOS, porém, devido ao alto número de silhuetas classificadas com a etiqueta “TAG_RUIDOS” tornou-se impossível analisar os resultados das taxas de reconhecimento após as filtrações para este tipo de erro. Na Tabela 4.3 é apresentado o total de silhuetas classificadas para cada tipo de etiqueta nas duas listas analisadas.

Tabela 4.2: Quantidade de seqüências antes e após processo de filtragem em cada lista.

Etiquetas	Qtd. de seqüências das listas						
	Lista da Galeria (122)			Lista de Teste (122)			
	Com	Sem	T1 Sem	Com	Sem	T1 Sem	T2 Sem
TAG_PERDA_PERNA	12	110	110	19	103	110	97
TAG_PERDA_DUAS_PERNAS	6	116	116	5	117	116	113
TAG_PERDA_CABECA	5	117	117	2	120	117	116
TAG_SUB_VAZIA	1	121	121	7	115	121	114
TAG_ERRO_TODO_QUADRO_MOVIMENTO	2	120	120	0	122	120	120
TAG_DEGRADACAO_1	27	95	95	33	89	95	78
TAG_DEGRADACAO_2	14	108	108	26	96	108	91
TAG_RUIDOS	103	19	19	120	2	19	0

Para entender os resultados obtidos, após cada filtragem são gerados dois tipos de gráficos: (a) gráfico de evolução e (b) gráfico de quantificação. O gráfico de evolução tem a função de apresentar a taxa de reconhecimento original obtida sem as filtrações e a

Tabela 4.3: Total de silhuetas classificadas nas listas de teste e galeria para cada tipo de etiqueta.

Etiquetas	Qtd. de silhuetas das listas	
	Lista de Treino (24950 silhuetas)	Lista do Exp. A (24950 silhuetas)
	Com	Com
TAG_PERDA_PERNA	36	76
TAG_PERDA_DUAS_PERNAS	12	17
TAG_PERDA_CABECA	13	9
TAG_SUB_VAZIA	1	18
TAG_ERRO_TODO_QUADRO_MOVIMENTO	3	—
TAG_DEGRADACAO_1	229	284
TAG_DEGRADACAO_2	42	78
TAG_RUIDOS	1287	3356

evolução das taxas de reconhecimento obtidas após os testes de filtragem “T1” e “T2”. Já o gráfico de quantificação tem a função de apresentar a quantidade de sequências removidas em cada posição do rank classificatório original pelos testes de filtragem “T1” e “T2”. Foram calculadas as taxas de reconhecimento dos testes de filtragem “T1” e “T2” para a abordagem de reconhecimento *baseline* (Figuras A.1-A.7), silhuetas médias (Figuras B.1-B.7) e vetores de largura (Figuras C.1-C.7).

Para exemplificar o processo de análise dos gráficos gerados, dois casos diferentes de resultados serão descritos. O primeiro caso a ser analisado será a filtragem da TAG_DEGRADACAO_1 da abordagem *baseline* na qual foram removidas 27 sequências das listas da galeria e de teste no teste de filtragem “T1”(Fig. A.6). As taxas de reconhecimento de “T1” aumentaram, porém, o que deve ser analisado é a evolução do teste “T2” em relação ao “T1”. No teste “T2”, além das 27 sequências removidas em “T1” foram removidas 17 novas sequências. Neste caso as taxas de “T2” foram inferiores a “T1” indicando que este tipo de erro não está influenciando as taxas de reconhecimento do Experimento A. A queda de “T2” em relação ao “T1” é causada devido a grande quantidade de sequências removidas no teste “T1” e novamente no teste “T2” da posição 1 do rank, isto é, foram removidas sequências de indivíduos que foram reconhecidas em um rank muito baixo (quanto mais baixo o rank melhor é a classificação de um indivíduo no processo de similaridade). O segundo caso a ser analisado será a filtragem da TAG_PERDA_PERNA da abordagem de vetores de largura na qual foram removidas 12 sequências das listas da galeria e de teste no teste de filtragem “T1”(Fig. C.1). As taxas de reconhecimento do teste “T2” aumentaram em relação a “T1”, porém, nota-se

que o rank 1 do teste “T2” (Fig. C.1a) não aumentou, isto é, a remoção das sequências de “T2” não colaboraram para que o rank de outros indivíduos fossem reduzidos aumentando a quantidade de indivíduos reconhecidos em ranks mais baixos (não houve alteração da ordem de classificação dos indivíduos, apenas a redução da quantidade de indivíduos na lista de teste e no rank 1).

Os resultados obtidos após calculadas as taxas de reconhecimento para os testes de filtragem “T1” e “T2” não apresentaram aumentos significativos nas taxas de reconhecimento.

Em todos os casos apresentados, o processo de filtragem resultou na remoção de indivíduos classificados em ranks muito baixos. Este fato demonstra que mesmo com a presença de erros muitos indivíduos são classificados com ranks baixos, isto é, a presença de determinado erro não influenciou no processo de classificação do indivíduo. Notou-se que a remoção das sequências com a presença de um determinado erro não colaborou para que as taxas de reconhecimento aumentassem, pois indivíduos diferentes continuaram a ser classificados como muito próximos nos cálculos de similaridade. Tudo indica que os erros analisados exercem influências mínimas sobre as taxas de reconhecimento das abordagens analisadas para o Experimento A da *USF Database*. Uma possível justificativa é que as abordagens testadas tenham uma certa resistência aos tipos de erros analisados. Caso seja necessário investigar novos tipos de erros ou fatores que possam influenciar o reconhecimento da forma de andar, a metodologia de análise (testes de filtragem T1 e T2) proposta neste trabalho pode ser facilmente estendida para a análise destes fatores

CAPÍTULO 5

CONCLUSÃO

Este trabalho apresenta um estudo sobre o uso da forma de andar dos seres humano como uma possível forma de identificação biométrica.

São apresentadas as principais bases de dados utilizadas para experimentos relacionados ao reconhecimento da forma de andar. Cada uma delas apresenta diferentes variações como o tipo de ambientes das filmagens (interno ou externo), tipo de superfície, tipo de calçado, ângulos de câmera e a variação de datas de gravação das sequências.

Para avaliar os resultados de reconhecimento das abordagens e dos experimentos desenvolvidos, foi escolhida a *USF Database* por esta apresentar uma grande quantidade de indivíduos, variações e um conjunto definido de experimentos que possibilitam a avaliação das abordagens de reconhecimento com e sem a presença das variações presentes nesta base.

Com base nas abordagens de reconhecimento estudadas foram implementadas abordagens que utilizam dois tipos de conceitos ou descritores: vetores e silhuetas médias. Nos testes feitos por meio da *USF Database*, entre as três abordagens implementadas a que apresentou melhores taxas de reconhecimento foi a abordagem de silhuetas médias. Por meio da combinação de características das abordagens vetores e silhuetas médias foram desenvolvidas duas novas abordagens de reconhecimento, *Averaged Width Vectors* e a *Averaged Mass Vectors*. As duas abordagens desenvolvidas alcançaram taxas de reconhecimento superiores as abordagens de vetores nos Experimentos A (ângulo de visão), B (calçado), C (ângulo e calçado) do *HumanID Gait Challenge Problem*.

Vários experimentos foram desenvolvidos por meio de modificações na estrutura de funcionamento da abordagem de silhuetas médias. Dentre os quatro experimentos desenvolvidos três apresentaram aumento das taxas de reconhecimento: o uso de somente uma silhueta média mediana (Subseção 3.3.4); a modificação do cálculo de similaridade

(Subseção 3.3.5); e a intercalação de ciclos (Subseção 3.3.6). As melhores taxas de reconhecimento foram obtidas após a combinação da intercalação de ciclos com a modificação do cálculo de similaridade.

Após identificado a presença de erros de segmentação nas silhuetas extraídas da *USF Database*, foi desenvolvida uma análise da influência desses erros sobre as taxas de reconhecimento das abordagens *baseline*, silhuetas médias e vetores de largura. Para cada tipo de erro foram calculadas as taxas de reconhecimento após a filtragem das sequências das listas da galeria e teste do Experimento A da *USF Database*. Por meio dos resultados obtidos identificou-se que a presença dos erros de segmentação analisados exercem influências mínimas sobre as taxas de reconhecimento das abordagens analisadas para o Experimento A da *USF Database*.

Ao final deste trabalho, conclui-se que seus objetivos foram alcançados após a apresentação de um estudo do uso da forma de andar dos seres humanos como uma característica biométrica para identificação de indivíduos. As principais bases de dados utilizadas para a avaliação das abordagens de reconhecimento foram descritas em detalhes e as principais abordagens de reconhecimento foram apresentadas. Após o desenvolvimento um conjunto de experimentos, foram propostas modificações na estrutura de funcionamento da abordagem de silhuetas médias que apresentaram melhorias em suas taxas de reconhecimento. Dentre as modificações, os melhores resultados foram alcançados por meio da combinação da intercalação de ciclos com a modificação do cálculo de similaridade

Um dos trabalhos futuros a ser desenvolvido é o estudo e implementação das abordagens que utilizam as GEIs em conjunto com as transformadas de Gabor e transformadas de redução de dimensionalidades devido as altas taxas de reconhecimento alcançadas na *USF Database*. Posteriormente estudar uma forma de combinar estas abordagens com características de abordagens baseadas em modelos.

Outro trabalho futuro é utilizar as imagens da base multi-biométrica desenvolvida pelo Grupo IMAGO de Pesquisa para construir uma abordagem de extração de características da forma de andar por meio de multi-câmera.

BIBLIOGRAFIA

- [1] J. Aggarwal and Q. Cai. Human motion analysis: A review. *Computer Vision and Image Understanding*, 73(3):428–440, 1999.
- [2] C. BenAbdelkader, R. Cutler, and L. Davis. Gait recognition using image self-similarity. *EURASIP Journal on Applied Signal Processing*, 2004(1):572–585, 2004.
- [3] C. BenAbdelkader, R. Cutler, H. Nanda, and L. Davis. Eigengait: Motion-based recognition using image self-similarity. In *Proceedings of the Third International Conference on Audio Visual Biometric Person Authentication*, volume 2091, pages 284–294, 2001.
- [4] A. Bobick and A. Johnson. Gait recognition using static, activity-specific parameters. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 423–430, 2001.
- [5] A. Bobick and A. Johnson. A multi-view method for gait recognition using static body parameters. In *Proceedings 3rd International Conference Audio and Video-Based Biometric Person Authentication*, pages 301–311, 2001.
- [6] R. M. Bolle, J. H. Connell, S. Pankanti, N. K. Ratha, and A. W. Senior. *Guide to Biometrics*. Springer, 2003.
- [7] N. V. Boulgouris, K. N. Plataniotis, and D. Hatzinakos. Gait recognition using dynamic time warping. In *IEEE 6th Workshop on Multimedia Signal Processing*, pages 263–266, 2004.
- [8] J.N. Carter, G. Veres, and M.S. Nixon. What image information is important in silhouette-based gait recognition. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 776–782, 2004.

- [9] D. Cunado, M.S. Nixon, and J.N. Carter. Using gait as a biometric, via phase-weighted magnitude spectra. In *Proceedings of the First International Conference on Audio Visual Biometric Person Authentication*, volume 1206, pages 95–102, 1997.
- [10] N. Cuntoor, A. Kale, and R. Chellappa. Combining multiple evidences for gait recognition. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 3, pages 33–36, 2003.
- [11] J.E. Cutting and L.T. Kozlowski. Recognition of friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, 9(5):353–356, 1977.
- [12] J.P. Foster, M.S. Nixon, and A. Prugel-Bennet. Recognising movement and gait by masking functions using area-based metrics. In *Lecture Notes in Computer Science (Proceedings of the Third International Conference on Audio Visual Biometric Person Authentication)*, volume 2091, pages 278–283, 2001.
- [13] J.P. Foster, M.S. Nixon, and A. Prugel-Bennet. Automatic gait recognition using area-based metrics. *Pattern Recognition Letters*, 24(14):2489–2497, 2003.
- [14] D. Gavrilu. The visual analysis of human movement: A survey. *Computer Vision and Image Understanding*, 73:82–98, 1999.
- [15] J. Han and B. Bhanu. Statistical feature fusion for gait-based human recognition. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 842–847, 2004.
- [16] J. Han and B. Bhanu. Individual recognition using gait energy image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(02):316–322, 2006.
- [17] X. He and P. Niyogi. Locality preserving projections. *Proceedings of the Conference on Advances in Neural Information Processing Systems*, 16:153–160, 2003.

- [18] A. Hilton and P. Fua. Modeling people toward vision-based understanding of a person's shape, appearance, and movement. *Computer Vision and Image Understanding*, 81:227–230, 2001.
- [19] S. Hong, H. Lee, I.F. Nizami, and E. Kim. A new gait representation for human identification: Mass vector. *IEEE Conference on Industrial Electronics and Applications*, pages 669–673, 2007.
- [20] Sungjun Hong, Heesung Lee, and Euntai Kim. Automatic gait recognition using width vector mean. In *IEEE Conference on Industrial Electronics and Applications*, pages 647–650, 25-27 2009.
- [21] P.S. Huang, C.J. Harris, and M.S. Nixon. Recognizing humans by gait via parametric canonical space. *Artificial Intelligence in Engineering*, 13:359–366, 1999.
- [22] G. Johansson. Visual motion perception. *Scientific Am.*, 232:75–88, 1976.
- [23] A. Kale, N. Cuntoor, B. Yegnanarayana, A. N. Rajagopalan, and R. Chellappa. Gait analysis for human identification. In *AVBPA'03: Proceedings of the 4th international conference on Audio- and video-based biometric person authentication*, pages 706–714, Berlin, Heidelberg, 2003. Springer-Verlag.
- [24] A. Kale, A.N. Rajagopalan, A. Sundaresan, N. Cuntoor, A. RoyChowdhury, V. Kruger, and R. Chellappa. Identification of humans using gait. *IEEE Transactions on Image Processing*, 13:1163–1173, 2004.
- [25] L. Lee and W. Grimson. Gait analysis for recognition and classification. In *Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition*, volume 1, pages 148–155, 2002.
- [26] J. Little and J. Boyd. Recognizing people by their gait: The shape of motion. *Videre: Journal of Computer Vision Research*, 1(2):1–32, 1998.

- [27] Z. Liu and S. Sarkar. Simplest representation yet for gait recognition: Average silhouette. In *Proceedings of the 17th International Conference on Pattern Recognition*, pages 211–214, 2004.
- [28] P. C. Mahalanobis. On the generalised distance in statistics. In *National Institute of Science*, volume 2, pages 49–55, 1936.
- [29] T. Moeslund and E. Granum. A survey of computer vision-based human motion capture. *Computer Vision and Image Understanding*, 81:231–268, 2001.
- [30] M.P. Murray. Gait as a total pattern of movement. *American Journal of Physical Medicine*, 46(1):290–332, 1967.
- [31] M.P. Murray, A.B. Drought, and R.C. Kory. Walking patterns of normal men. *Journal of Bone and Joint Surgery*, 46-A:335–360, 1964.
- [32] Mark S. Nixon, Tieniu Tan, and Rama Chellappa. *Human Identification Based on Gait*. Springer Science+Business Media, 2006.
- [33] M.S. Nixon, J.N. Carter, D. Cunado, P.S. Huang, and S.V. Stevenage. Automatic gait recognition. In A.K. Jain, R. Bolle, and S. Pankanti, editors, *Biometrics: Personal Identification in Networked Society*, chapter 11, pages 231–250. Kluwer Academic Publishing (Dordrecht Netherlands), 1999.
- [34] P.J. Phillips, S. Sarkar, I. Robledo, P. Grother, and K. Bowyer. Baseline results for then challenge problem of Human ID using gait analysis. In *Proceedings of IEEE International Conference Face and Gesture Recognition*, pages 137–143, 2002.
- [35] P.J. Phillips, S. Sarkar, I. Robledo, P. Grother, and K. Bowyer. The gait identification challenge problem: Data sets and baseline algorithm. In *Proceedings of 16th International Conference on Pattern Recognition*, pages 385–389, 2002.
- [36] H. Sadeghi, P. Allard, F. Prince, and H. Labelle. Symmetry and limb dominance in able-bodied gait: a review. *Gait Posture*, 12(1):34–45, 2000.

- [37] Hiroaki Sakoe and Seibi Chiba. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(1):43–49, 1978.
- [38] S. Sarkar, P.J. Phillips, Z. Liu, I.R. Vega, P. Grother, and K. Bowyer. The HumanId gait challenge problem: Data sets, performance and analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(2):162–177, 2005.
- [39] S.V. Stevenage, M.S. Nixon, and K. Vince. Visual analysis of gait as a cue to identify. *Applied Cognitive Psychology*, 13(6):513–526, 1999.
- [40] A. Sundaresan, A. RoyChowdhury, and R. Chellappa. A hidden markov model based framework for recognition of humans from gait sequences. In *Proceedings IEEE International Conference on Image Processing*, pages 93–96, 2003.
- [41] J.A. Webb and J.K. Aggarwal. Structure from motion of rigid and jointed objects. *Artificial Intelligence*, 19:107–130, 1982.
- [42] X. Yang, T. Zhang, Y. Zhou, and J. Yang. Gabor phase embedding of gait energy image for identity recognition. In *Computer and Information Technology, 2008. CIT 2008. 8th IEEE International Conference on*, pages 361–366, 8-11 2008.
- [43] G. Zhao, R. Chen, G. Liu, and L. Hua. Amplitude spectrum-based gait recognition. In *Proceedings of the IEEE International Conference Face and Gesture Recognition*, pages 23–28, 2004.

APÊNDICE A

RESULTADOS DAS FILTRAGENS PARA A ABORDAGEM *BASELINE*

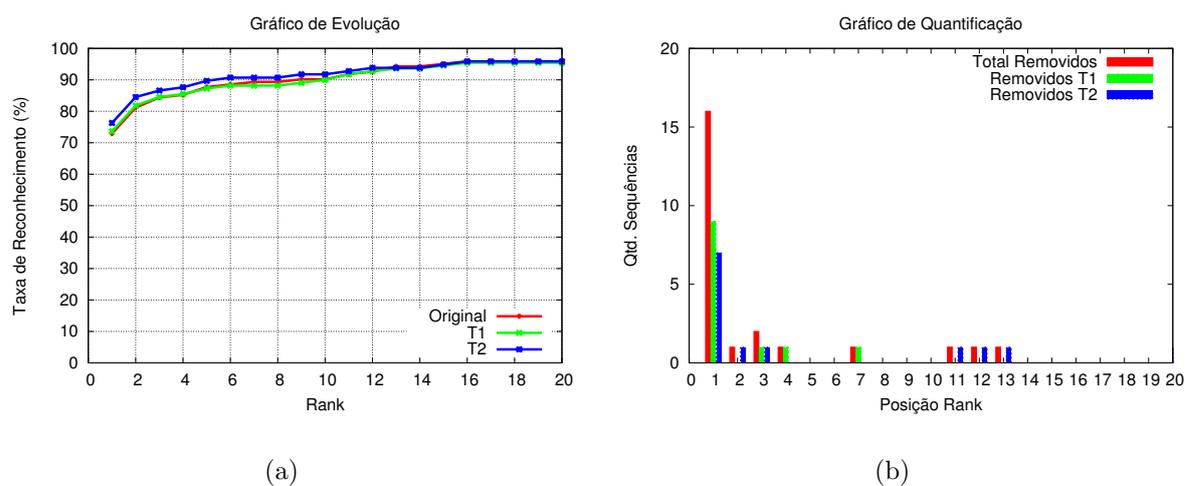


Figura A.1: Taxas de reconhecimento da abordagem *baseline* após filtragem da etiqueta TAG_PERDA_PERNA.

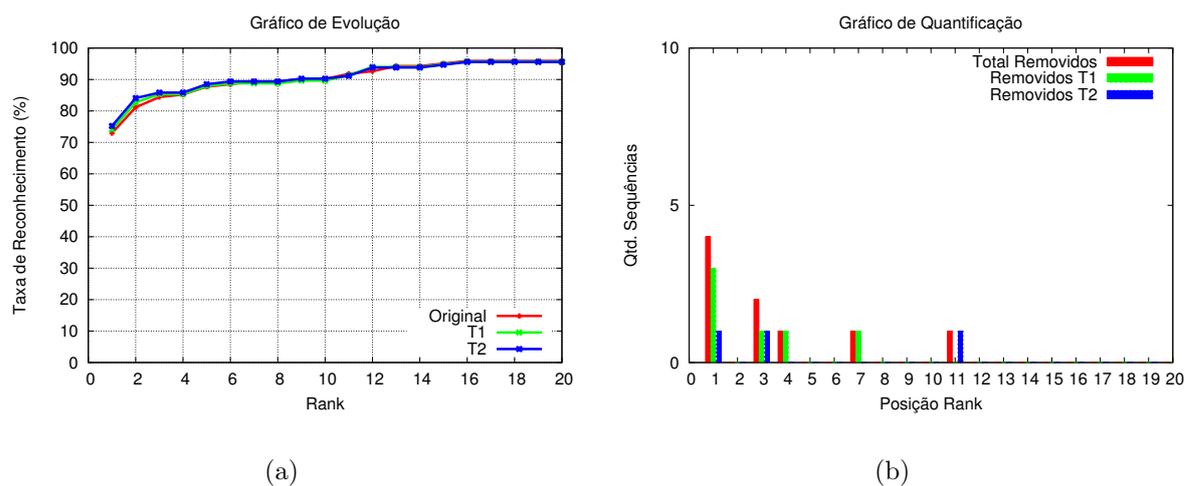
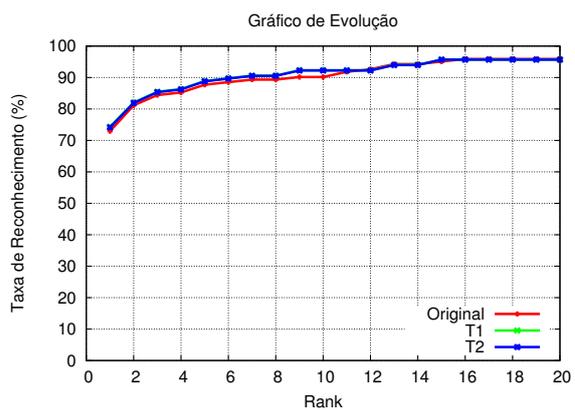
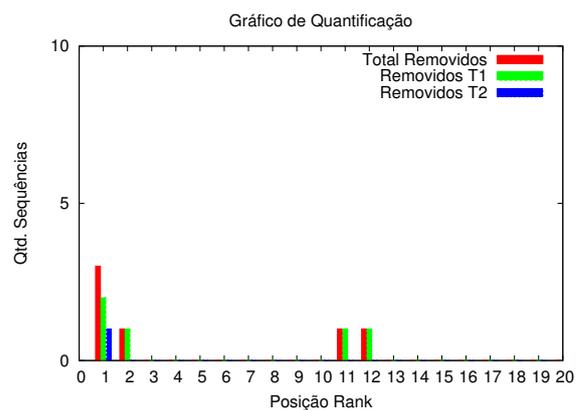


Figura A.2: Taxas de reconhecimento da abordagem *baseline* após filtragem da etiqueta TAG_PERDA_DUAS_PERNAS.

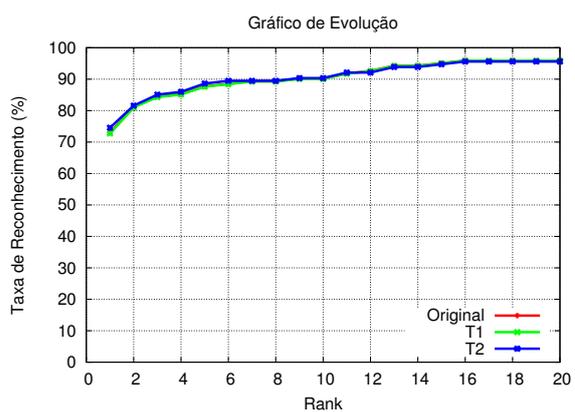


(a)

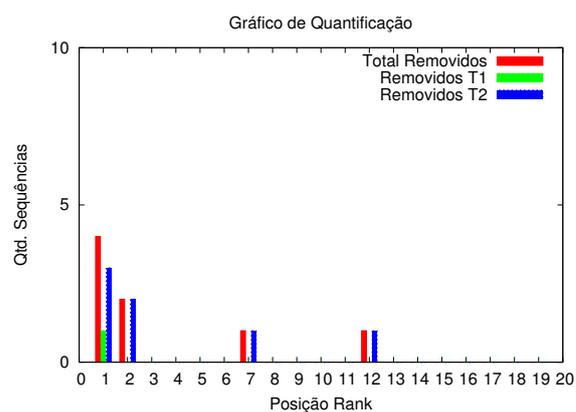


(b)

Figura A.3: Taxas de reconhecimento da abordagem *baseline* após filtragem da etiqueta TAG_PERDA_CABECA.



(a)



(b)

Figura A.4: Taxas de reconhecimento da abordagem *baseline* após filtragem da etiqueta TAG_SUB_VAZIA.

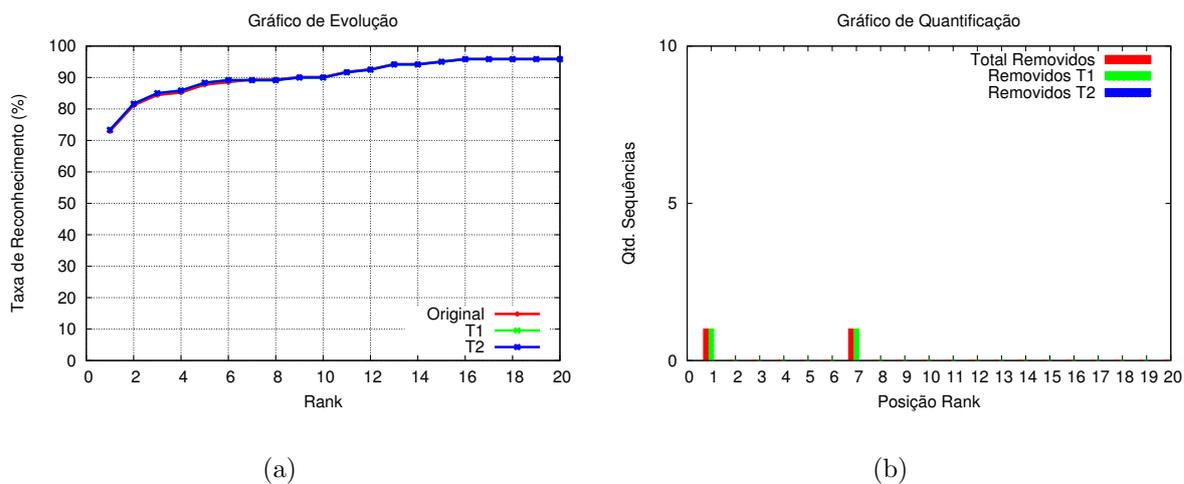


Figura A.5: Taxas de reconhecimento da abordagem *baseline* após filtragem da etiqueta TAG_ERRO_TODO_QUADRO_MOVIMENTO.

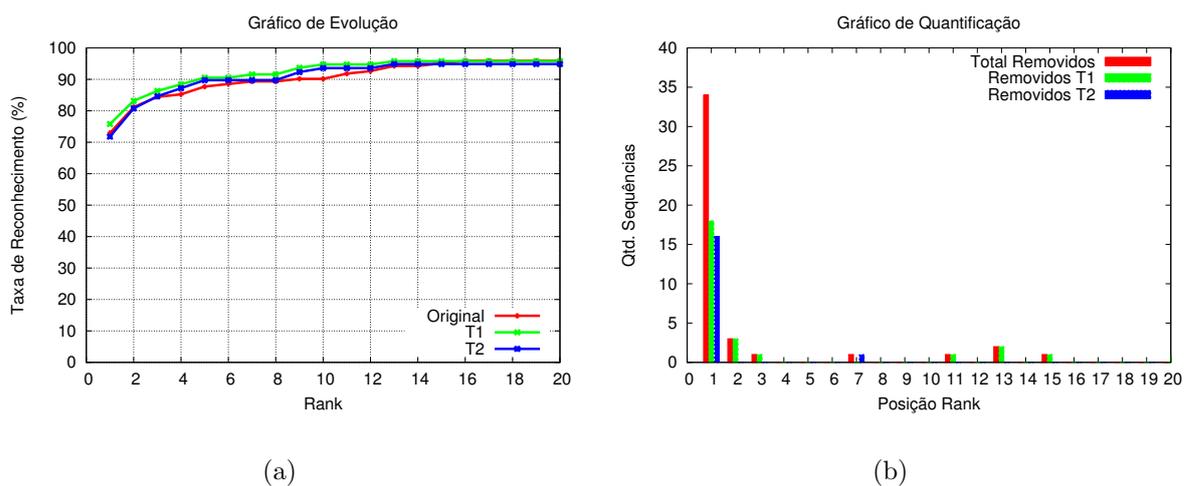


Figura A.6: Taxas de reconhecimento da abordagem *baseline* após filtragem da etiqueta TAG_DEGRADACAO_1.

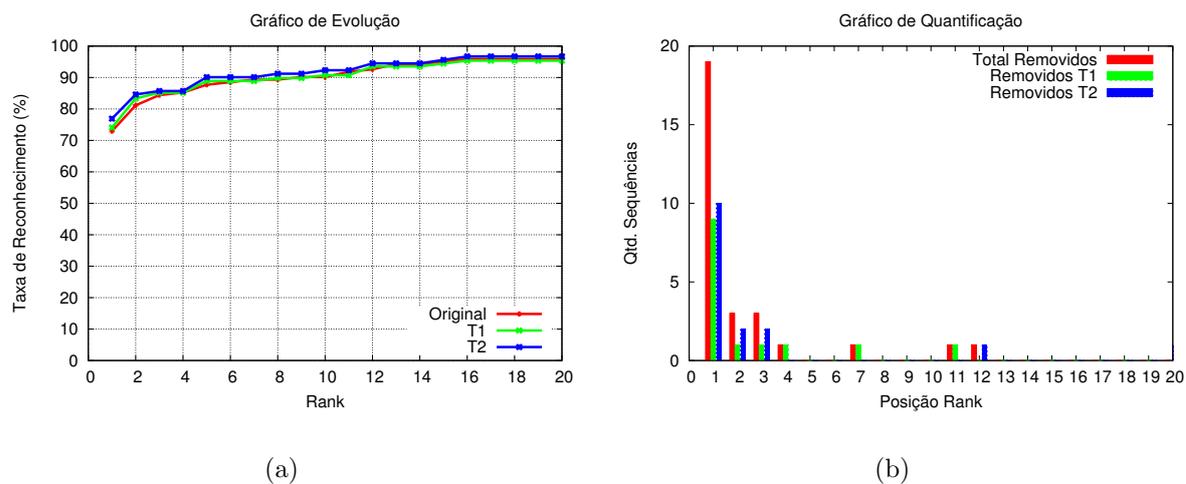


Figura A.7: Taxas de reconhecimento *baseline algorithm* após filtragem da etiqueta TAG_DEGRADACAO_2.

APÊNDICE B

RESULTADOS DAS FILTRAGENS PARA A ABORDAGEM DE *SILHUETAS MÉDIAS*

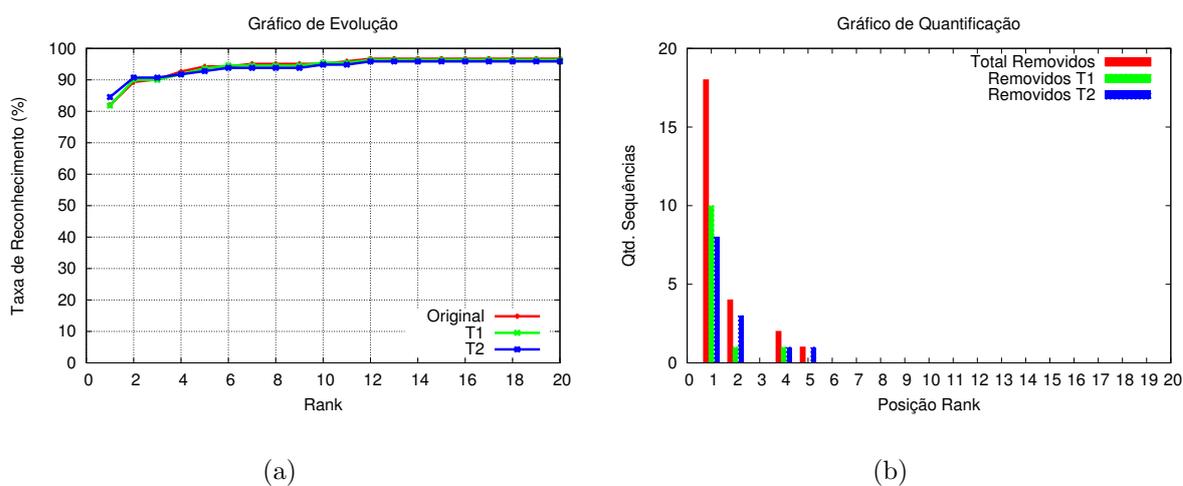


Figura B.1: Taxas de reconhecimento da abordagem de silhuetas médias após filtragem da etiqueta TAG_PERDA_PERNA.

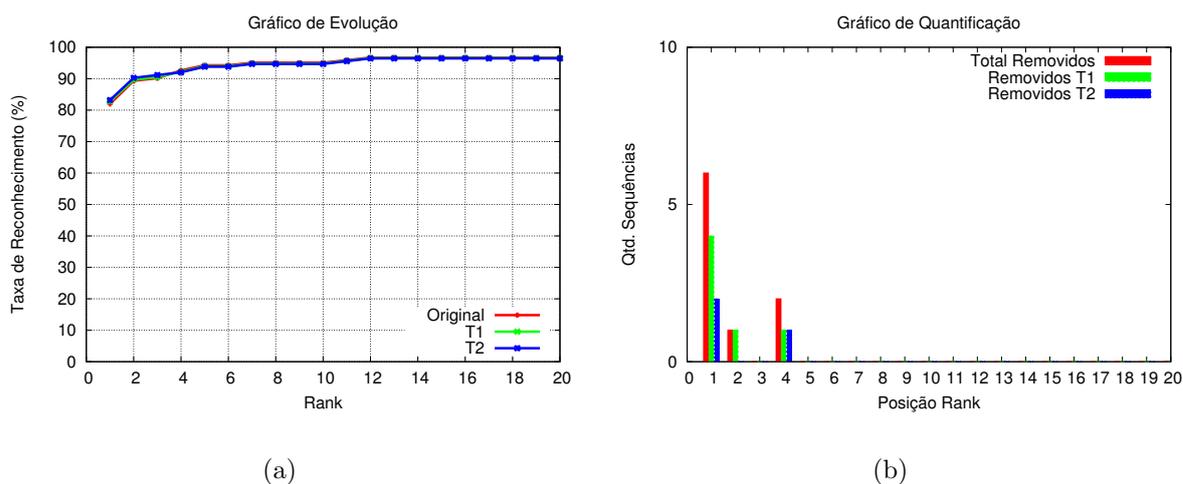
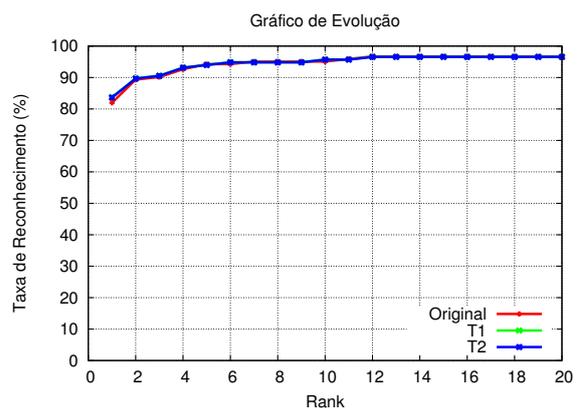
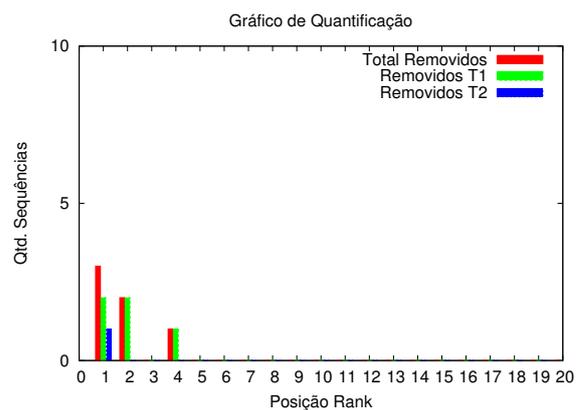


Figura B.2: Taxas de reconhecimento da abordagem de silhuetas médias após filtragem da etiqueta TAG_PERDA_DUAS_PERNAS.

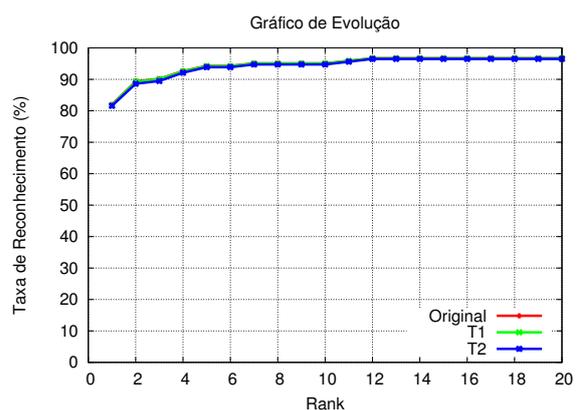


(a)

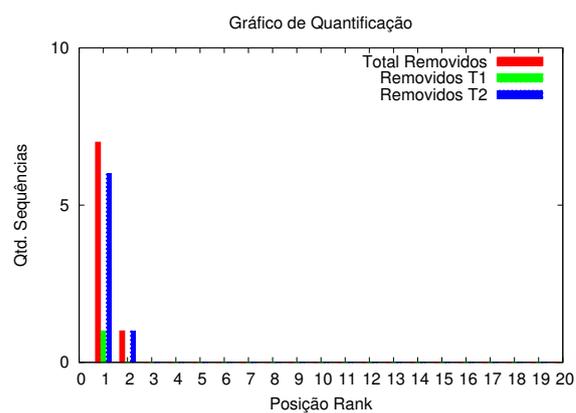


(b)

Figura B.3: Taxas de reconhecimento da abordagem de silhuetas médias após filtragem da etiqueta TAG_PERDA_CABECA.



(a)



(b)

Figura B.4: Taxas de reconhecimento da abordagem de silhuetas médias após filtragem da etiqueta TAG_SUB_VAZIA.

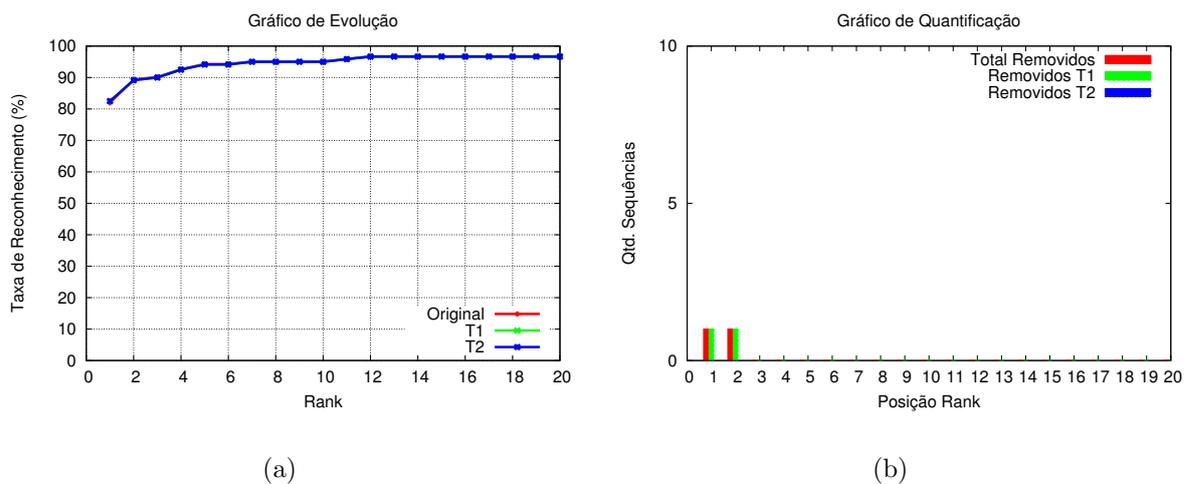


Figura B.5: Taxas de reconhecimento da abordagem de silhuetas médias após filtragem da etiqueta TAG_ERRO_TODO_QUADRO_MOVIMENTO.

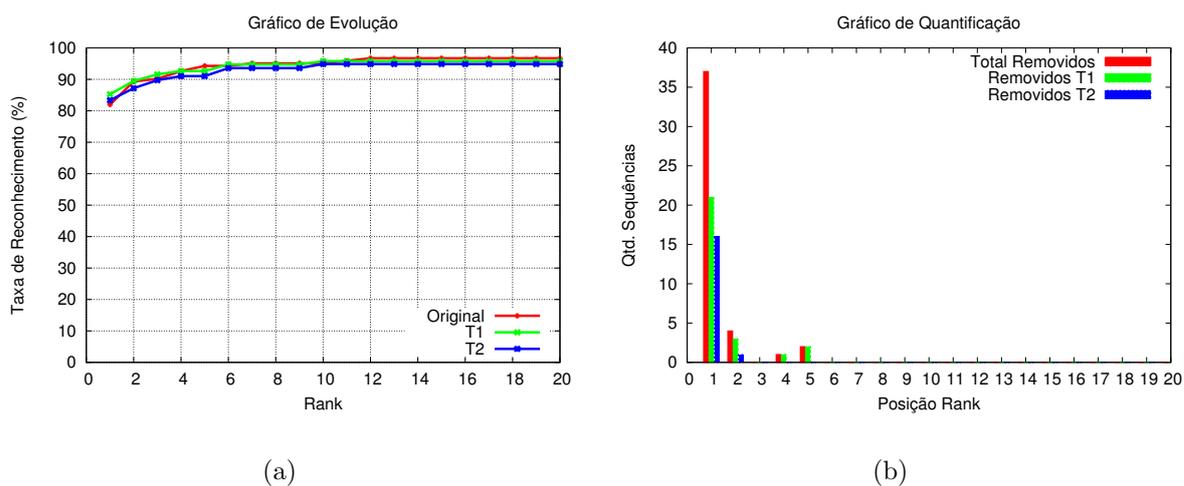


Figura B.6: Taxas de reconhecimento da abordagem de silhuetas médias após filtragem da etiqueta TAG_DEGRADACAO_1.

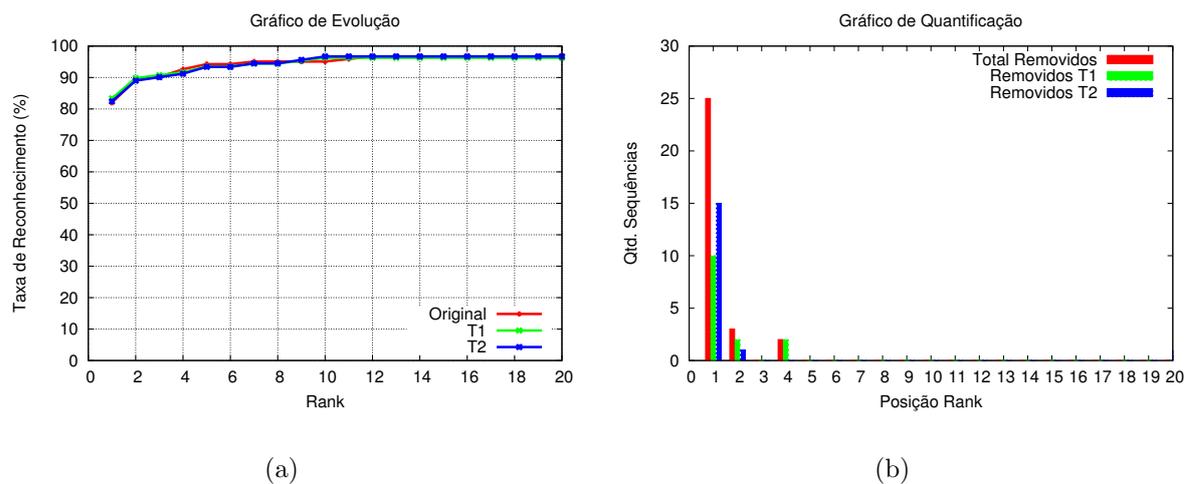


Figura B.7: Taxas de reconhecimento da abordagem de silhuetas médias após filtragem da etiqueta TAG_DEGRADACAO_2.

APÊNDICE C

RESULTADOS DAS FILTRAGENS PARA A ABORDAGEM DE VETORES DE LARGURA

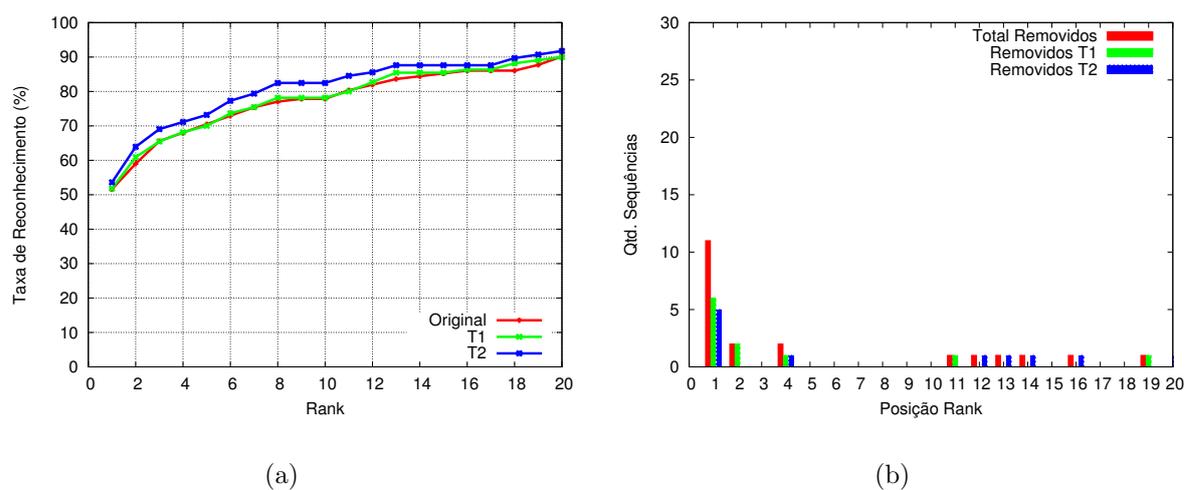


Figura C.1: Taxas de reconhecimento da abordagem de vetores de largura após filtragem da etiqueta TAG_PERDA_PERNA.

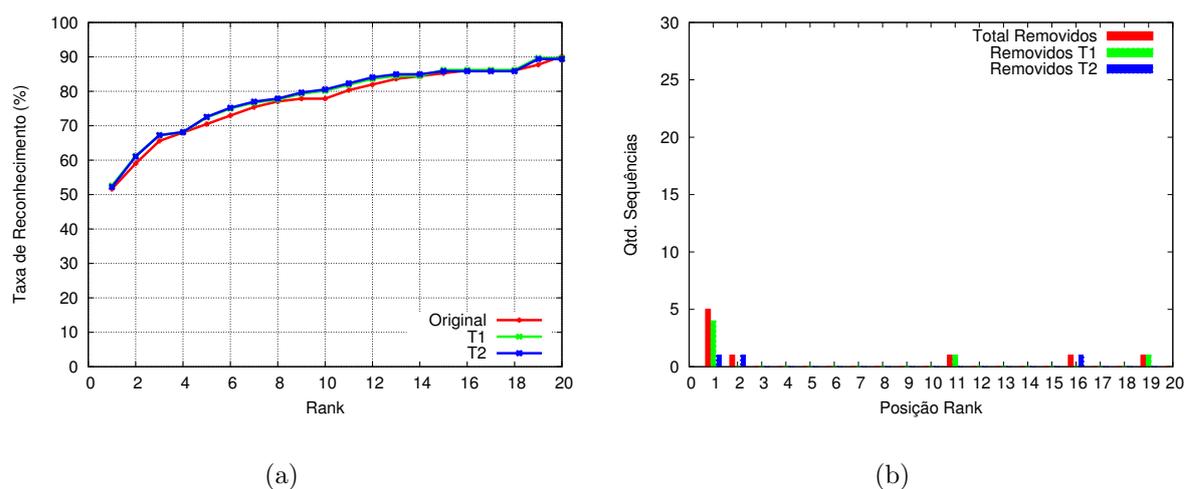


Figura C.2: Taxas de reconhecimento da abordagem de vetores de largura após filtragem da etiqueta TAG_PERDA_DUAS_PERNAS.

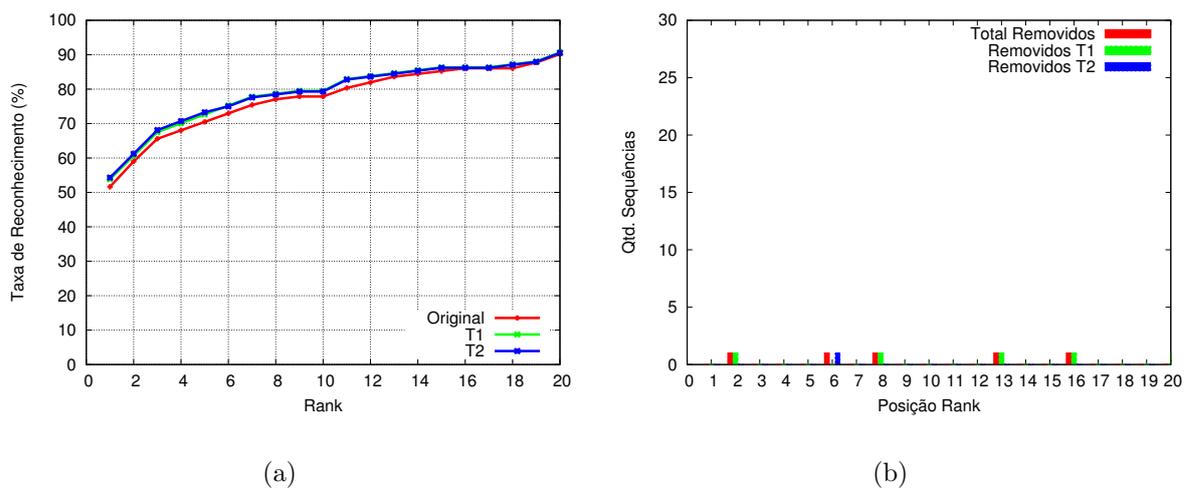


Figura C.3: Taxas de reconhecimento da abordagem de vetores de largura após filtragem da etiqueta TAG_PERDA_CABECA.

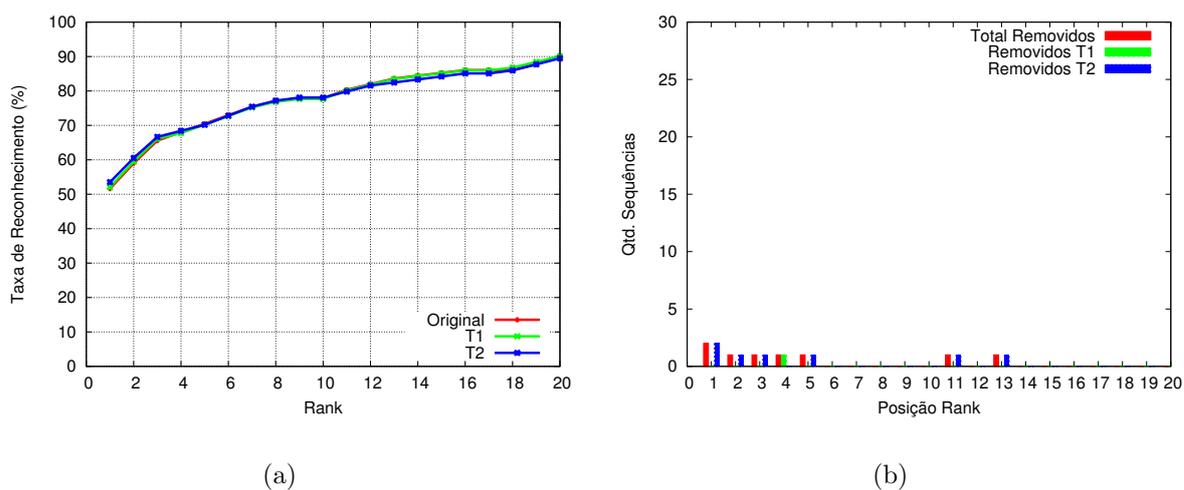


Figura C.4: Taxas de reconhecimento da abordagem de vetores de largura após filtragem da etiqueta TAG_SUB_VAZIA.

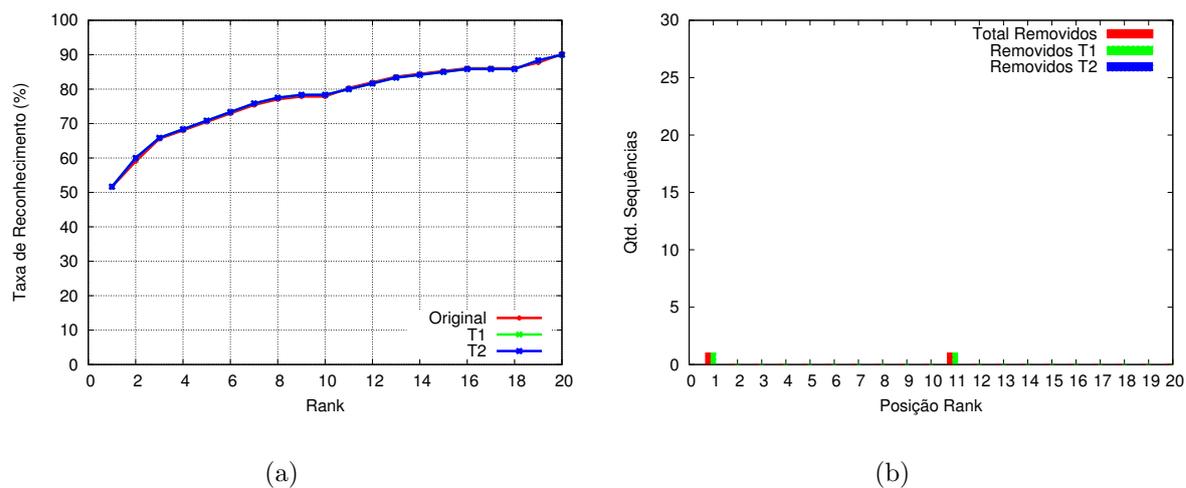


Figura C.5: Taxas de reconhecimento da abordagem de vetores de largura após filtragem da etiqueta TAG_ERRO_TODO_QUADRO_MOVIMENTO.

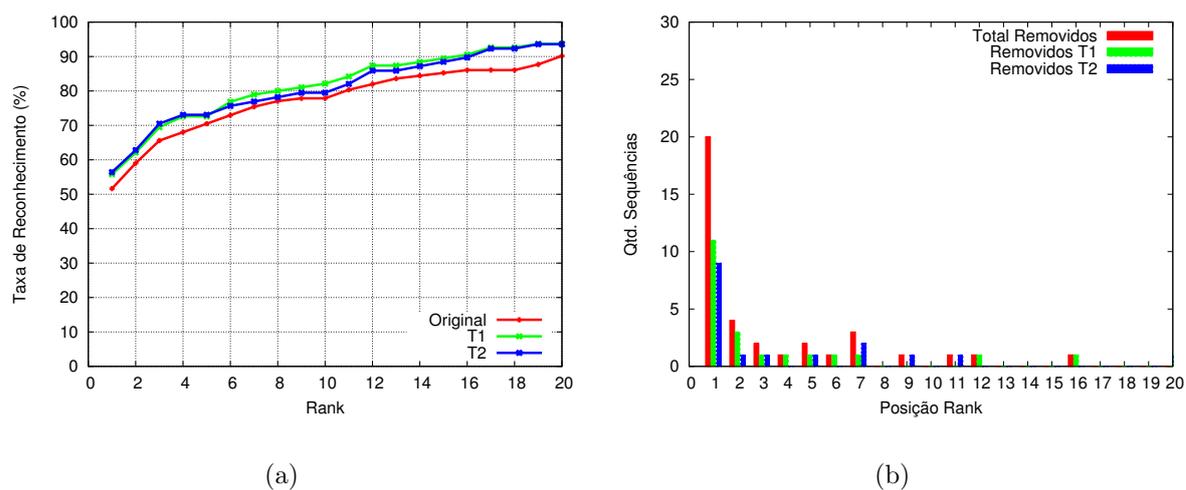


Figura C.6: Taxas de reconhecimento da abordagem de vetores de largura após filtragem da etiqueta TAG_DEGRADACAO_1.

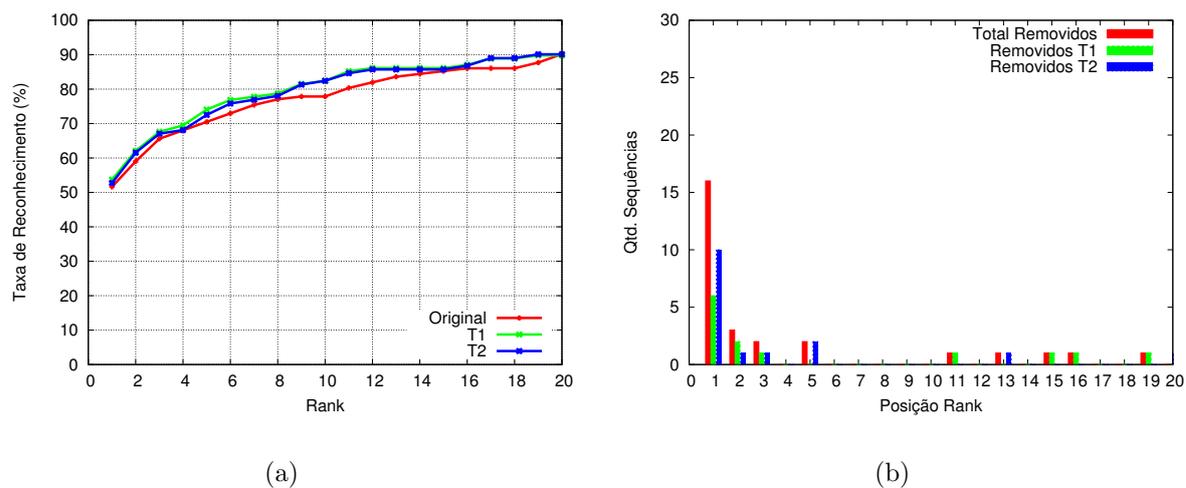


Figura C.7: Taxas de reconhecimento da abordagem de vetores de largura após filtragem da etiqueta TAG_DEGRADACAO_2.