

UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE SÃO CARLOS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

Milene Arantes

**Método de Reconhecimento da Marcha
Humana por meio da Fusão das
Características do Movimento Global**

São Carlos
2010

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

Milene Arantes

Método de Reconhecimento da Marcha Humana por meio da Fusão das Características do Movimento Global

Tese apresentada à Escola de Engenharia de São Carlos - USP, como parte dos requisitos para obtenção do título de Doutor em Ciências, Programa de Engenharia Elétrica.

Área de Concentração: - Processamento de Sinais e Instrumentação

Orientador: Prof. Dr. Adilson Gonzaga

SÃO CARLOS

2010

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica preparada pela Seção de Tratamento
da Informação do Serviço de Biblioteca – EESC/USP

A662m

Arantes, Milene

Método de reconhecimento da marcha humana por meio de fusão das características do movimento global / Milene Arantes ; orientador Adilson Gonzaga. -- São Carlos, 2010.

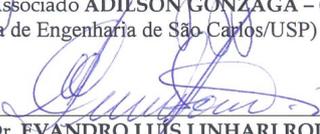
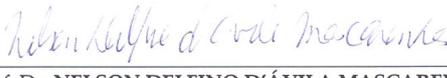
Tese (Doutorado-Programa de Pós-Graduação em Engenharia Elétrica. Área de Concentração em Processamento de Sinais e Instrumentação) -- Escola de Engenharia de São Carlos da Universidade de São Paulo, 2010.

1. Biometria. 2. Marcha humana. 3. Reconhecimento de marcha. 4. Análise da marcha. 5. Fusão de característica. 6. Movimento global. I. Título.

FOLHA DE JULGAMENTO

Candidato(a): Bacharel MILENE ARANTES.

Tese defendida e julgada em 01/04/2010 perante a Comissão Julgadora:

 _____ Prof. Associado ADILSON GONZAGA – (Orientador) (Escola de Engenharia de São Carlos/USP)	<u>APROVADA</u>
 _____ Prof. Dr. EVANDRO LUIS LINHARI RODRIGUES (Escola de Engenharia de São Carlos/USP)	<u>APROVADA</u>
 _____ Prof. Dr. JOSÉ HIROKI SAITO (Universidade Federal de São Carlos/UFSCar)	<u>APROVADA</u>
 _____ Prof. Dr. NELSON DELFINO D'ÁVILA MASCARENHAS (Universidade Federal de São Carlos/UFSCar)	<u>APROVADA</u>
 _____ Prof. Dr. LEONARDO VIDAL BATISTA (Universidade Federal da Paraíba/UFPB)	<u>APROVADA</u>



Prof. Titular **GERALDO ROBERTO MARTINS DA COSTA**
Coordenador do Programa de Pós-Graduação em Engenharia Elétrica e
Presidente da Comissão de Pós-Graduação

*Ao meu filho,
Hugo.*

Agradecimentos

Agradeço a Deus, porque por meio d'Ele tudo é possível.

Há dez anos estudo em São Carlos e, nestes anos conheci muitas pessoas, fiz amizades. Fernando, não há palavras para expressar a minha amizade e nem minha gratidão por tudo que você fez durante estes anos, meu muito obrigado.

Adilson, agradeço não só pela dedicada orientação, mas pela paciência e compreensão com que esteve sempre presente durante toda a execução deste trabalho. Vou confessar que muitas vezes não acompanhava seu raciocínio mas, grande mestre que é, sempre pacientemente explicava tudo novamente.

Agradeço minha irmã Anita pelo constante incentivo e apoio.

Quero também agradecer aos diversos amigos, que muitas vezes ajudaram com seus conhecimentos e experiências e muitas outras com palavras de apoio. Agradecimentos especiais ao Paulo Henrique, Claudia, Cristiane, Cristiano, Luciana, João Marcelo e Carol.

Agradeço aos funcionários e professores do programa de pós-graduação do Departamento de Engenharia Elétrica.

Resumo

ARANTES, M. **Método de Reconhecimento da Marcha Humana por meio da Fusão das Características do Movimento Global.** 2010. 121p. Tese (Doutorado) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2010.

Este trabalho propõe um novo enfoque em visão computacional aplicado a sequências de vídeo, de pessoas em movimento, para reconhecê-las por meio da marcha. O movimento humano carrega diferentes informações, considerando-se diferentes maneiras de analisá-lo. O esqueleto carrega as informações do movimento global de articulações do corpo humano e como se comportam durante a caminhada e a silhueta carrega informações referentes ao comportamento global do contorno do corpo humano. Além disso, imagens binárias e em escala de cinza possuem diferentes informações sobre o movimento humano. O método proposto considera o conjunto de *frames* segmentados de cada indivíduo como uma classe e cada *frame* como um objeto desta classe. A metodologia aplica o Modelo de Mistura de Gaussianas (GMM) para subtração de fundo, redução de escala realizada por meio de técnicas de multiresolução baseadas na Transformada Wavelet (TW) e a extração dos padrões por meio da Análise dos Componentes Principais (PCA). São propostos e ensaiados quatro novos modelos de captura de movimentos globais do corpo humano durante a marcha: o modelo Silhouette-Gray-Wavelet (SGW) captura o movimento baseado nas variações em nível de cinza; o modelo Silhouette-Binary-Wavelet (SBW) captura o movimento baseado nas informações binárias da silhueta; o modelo Silhouette-Edge-Wavelet (SEW) captura o movimento baseado nas informações contidas na borda das silhuetas e o modelo Silhouette-Skeleton-Wavelet (SSW) captura o movimento baseado do esqueleto humano. As taxas de classificações corretas obtidas separadamente a partir destes quatro diferentes modelos são então combinadas utilizando-se uma nova técnica de fusão. Os resultados demonstram excelente desempenho e mostraram a viabilidade para reconhecimento de pessoas.

Palavras-chave: marcha humana, reconhecimento da marcha, análise da marcha, biome-
tria, fusão de característica, movimento global .

Abstract

Arantes, M. **Recognition Method of Human Gait by Fusion of Features of the Global Movement.** 2010. 121p. Thesis (Doctoral) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2010.

This paper proposes a novel computer vision approach that processes video sequences of people walking and then recognises those people by their gait. Human motion carries different information that can be analysed in various ways. The skeleton carries motion information about human joints, and the silhouette carries information about boundary motion of the human body. Moreover, binary and gray-level images contain different information about human movements. Our proposed method considers the set of the segmented frames of each individual as a distinct class and each frame as an object of this class. The methodology applies background extraction using the Gaussian Mixture Model (GMM), a scale reduction based on the Wavelet Transform (WT) and feature extraction by Principal Component Analysis (PCA). We propose four new schemas for motion information capture: the Silhouette-Gray-Wavelet model (SGW) captures motion based on grey level variations; the Silhouette-Binary-Wavelet model (SBW) captures motion based on binary information; the Silhouette-Edge-Binary model (SEW) captures motion based on edge information and the Silhouette-Skeleton-Wavelet model (SSW) captures motion based on skeleton movement. The classification rates obtained separately from these four different models are then merged using a new proposed fusion technique. The results suggest excellent performance in terms of recognising people by their gait.

Keywords: human gait, gait recognition, gait analysis, biometry, fusion of characteristics, global motion.

Lista de Figuras

Figura 1 - Exemplos de cada modelo: (a) SGW; (b) SBW; (c) SEW; (d) SSW.	22
Figura 2 - Representação de apoio e balanço em um ciclo de marcha.	26
Figura 3 - Comparação entre passo e passada. Fonte: (PERRY, 2005)	26
Figura 4 - Divisões do ciclo de marcha. Adaptado de Perry (2005)	27
Figura 5 - As fases da Marcha Humana. a)Contato Inicial; b)Resposta à Carga; c)Apoio Médio; d)Apoio Terminal; e)Pré-Balanço; f)Balanço Inicial; g)Balanço Médio; h)Balanço Terminal. Adaptado de Perry (2005).	28
Figura 6 - Métodos de Reconhecimento da Marcha.	34
Figura 7 - A silhueta de um indivíduo dividido em 7 regiões (a) e elipses (b) são encaixadas em cada região. Fonte: (LEE; GRIMSON, 2002).	35
Figura 8 - Ilustração da representação do formato da silhueta. Adaptado de Wang et al. (2002).	37
Figura 9 - (a) câmera posicionada à esquerda para superfície de concreto, (b) câmera à direita para superfície de concreto, (c) câmera à esquerda para superfície de grama, (d) câmera à direita para superfície de grama . Fonte: (Sarkar et al. (2005)).	40
Figura 10 - Exemplo de quatro sequências externas de uma pessoa. Fonte: (BE- NABDELKAER; CUTLER; DAVIS, 2002)	45
Figura 11 - Segunda base de dados. Sequências com oito variações de ângulo, ambiente interno. Fonte:(BENABDELKAER; CUTLER; DAVIS, 2002).	46
Figura 12 - Exemplo de um <i>frame</i> e seu respectivo MEI, a partir da vista frontal, para cada um dos 18 exercícios aeróbicos utilizados para testar a nova representação do movimento. Fonte: (BOBICK; DAVIS, 2001).	47
Figura 13 - Exemplo de silhuetas pertencentes a duas sequências distintas de marcha. A imagem mais à direita de cada linha corresponde ao GEI obtido. Fonte: (HAN; BHANU, 2006).	50
Figura 14 - Esquema geral do modelo proposto – Global Body Motion (GBM)	57

Figura 15 - (a) sequência original; (b) sequência Binária (SB) obtida a partir da máscara binária gerada pelo GMM; (c) Subtração do fundo (SG) por meio do GMM em escala de cinza.	58
Figura 16 - Algoritmo de decomposição da imagem, por meio da TW de Haar.	59
Figura 17 - (a) Contornos do corpo (SEW); (b) sequência de esqueletos do movimento (SSW).	61
Figura 18 - Exemplo de cada modelo gerado. A sequência SGW é obtida a partir da sequência SG; a sequência SBW é obtida a partir da sequência SB; e as sequências SEW e SSW são obtidas a partir da sequência SBW.	61
Figura 19 - Diferentes níveis de fusão: (a) Fusão no nível de extração de características, (b) Fusão no nível do módulo de combinação e (c) Fusão no nível de decisão. Adaptado de JAIN, ROSS e PRABHAKAR (2004).	65
Figura 20 - Esquema de aquisição da base de imagens própria. A direção de caminhada é perpendicular ao eixo óptico da câmera fixa.	68
Figura 21 - Amostra de quadros com variação de ângulos (0° , 90° e 45°) da Base B. Fonte: (CASIA, 2005).	69
Figura 22 - Amostra de quadros com variação de ângulos da base C. a) Sequência bg; b) Sequência cl; c) Sequência nm – Os frames dispostos no topo representam os 11 diferentes ângulos (0° a 180°), para a sequência carregando bolsa. Fonte: (CASIA, 2005).	70
Figura 23 - Gráfico comparativo das precisões médias de cada modelo, com a variação do sub-espaco PCA para os dois experimentos realizados	80
Figura 24 - Resultado de desempenho para a classe 7 para o método de validação cruzada com 10 partições. O eixo <i>y</i> representa o número da partição utilizada para teste.	85
Figura 25 - Frames defeituosos que prejudicam uma correta classificação (a) Frames do modelo SEW (b) Frames do modelo SBW.	87
Figura 26 - Curvas <i>Recall</i> \times <i>Precision</i>	88
Figura 27 - Curva CMS para o Teste-2 – Base A.	88
Figura 28 - Curvas Recall x Precision - Ângulo 0°	92
Figura 29 - Curvas Recall x Precision - Ângulo 45°	92
Figura 30 - Curvas Recall x Precision - Ângulo 90°	92

Figura 31 - Modelo SGW, considerando a combinação das 4 sequências de caminhada.	93
Figura 32 - Modelo SBW, considerando a combinação das 4 sequências de caminhada.	93
Figura 33 - Modelo SEW, considerando a combinação das 4 sequências de caminhada.	94
Figura 34 - Modelo SSW, considerando a combinação das 4 sequências de caminhada.	94
Figura 35 - Fusão dos 4 modelos, considerando a combinação das 4 sequências de caminhada.	95
Figura 36 - Curvas CMS SGW.	98
Figura 37 - Curvas CMS SBW.	98
Figura 38 - Curvas CMS SEW.	98
Figura 39 - Curvas CMS SSW.	98
Figura 40 - Curvas CMS SGW - cl.	99
Figura 41 - Curvas CMS SBW - cl.	99
Figura 42 - Curvas CMS SEW - cl.	99
Figura 43 - Curvas CMS SSW - cl.	99
Figura 44 - Curvas Recall x Precision SGW - Combinação de duas sequências do tipo bg.	111
Figura 45 - Curvas Recall x Precision SGW - Combinação de duas sequências do tipo cl.	111
Figura 46 - Curvas Recall x Precision SGW- Combinação de seis sequências do tipo nm.	111
Figura 47 - Curvas Recall x Precision SBW - Combinação de duas sequências do tipo bg.	112
Figura 48 - Curvas Recall x Precision SBW - Combinação de duas sequências do tipo cl.	112
Figura 49 - Curvas Recall x Precision SBW- Combinação de seis sequências do tipo nm.	112

Figura 50 - Curvas Recall x Precision SEW - Combinação de duas sequências do tipo bg.	113
Figura 51 - Curvas Recall x Precision SEW - Combinação de duas sequências do tipo cl.	113
Figura 52 - Curvas Recall x Precision SEW- Combinação de seis sequências do tipo nm.	113
Figura 53 - Curvas Recall x Precision SSW - Combinação de duas sequências do tipo bg.	114
Figura 54 - Curvas Recall x Precision SSW - Combinação de duas sequências do tipo cl.	114
Figura 55 - Curvas Recall x Precision SSW- Combinação de seis sequências do tipo nm.	114
Figura 56 - Curvas Recall x Precision SGW: (a) Combinação de sequências bg com sequências cl; (b) combinação de sequências bg com sequências nm; (c) combinação de sequências cl com sequências nm.	115

Lista de Tabelas

Tabela 1 - Taxa de Acertos (CCR). Método utilizando a distância <i>Procrustes</i> com base NLPR.	38
Tabela 2 - Combinação das condições de caminhada para os experimentos de Sundaresan, RoyChowdhury e Chellappa (2003).	39
Tabela 3 - Desempenho do sistema em porcentagem para as três métricas de distância, para o rank 1. Os números entre colchetes denotam o número de pessoas contidas em cada conjunto.	40
Tabela 4 - Pontuação Acumulativa de Acertos para o Rank 1.	42
Tabela 5 - Taxa de classificação correta (CCR) para a base NLPR.	43
Tabela 6 - Taxa de reconhecimento para diferentes números de exemplares de treinamento e teste, para a base SOTON.. . . .	49
Tabela 7 - Taxa de reconhecimento para diferentes números de exemplares de treinamento e teste, para a base SOTON (SHUTLER et al., 2002).	49
Tabela 8 - Comparativo de desempenho do método proposto por Wang et al. (2003) para o Rank 1 - base NLPR, ângulo 0°	53
Tabela 9 - Matrizes de Confusão do Teste-1 para a Base A.	81
Tabela 10 - Porcentagens de Acertos com seus respectivos FR e FA médios, para as 10 amostragens de diferentes frames.	84
Tabela 11 - Matrizes de Confusão do Teste-2 para a Base A.	86
Tabela 12 - Percentuais CMS obtidos para o Teste-2 – Base A.	88
Tabela 13 - Matriz de Confusão: Ângulo 90° - Modelo SGW - sequências 2 e 4.	89
Tabela 14 - Matriz de Confusão: Ângulo 90° – Modelo SBW – sequências 2 e 4.	90
Tabela 15 - Porcentagem de acertos para os modelos: SGW, SBW, SEW e SSW; considerando todas as variações de ângulo e sequências.	91
Tabela 16 - Tabela comparativa do modelo GBM para a vista lateral (0°) – Base B.	96

Tabela 17 - Tabela comparativa para a vista oblíqua, combinação de duas sequências sentido esquerda-direita.	96
Tabela 18 - Tabela Comparativa para variação de ângulos (0° a 90°), considerando-se duas combinações das sequências do tipo bg e duas sequências do tipo cl.	100
Tabela 19 - Tabela Comparativa para variação de ângulos (108° a 180°), considerando-se duas combinações das sequências do tipo bg e duas sequências do tipo cl.	101
Tabela 20 - Variação de ângulos, considerando duas sequências de caminhada com pessoas carregando bolsa e duas sequências com pessoas vestindo casaco, após a fusão.	102
Tabela 21 - Desempenho do sistema para a base B.	105

Sumário

Dedicatória	
1 Introdução	19
1.1 Objetivos	20
1.2 Metodologia	21
1.3 Principais Contribuições	23
1.4 Organização do trabalho	23
2 A Marcha Humana	25
2.1 Considerações Iniciais	25
2.2 Ciclo de marcha	25
2.3 Considerações finais	30
3 A Marcha Humana no Contexto de Técnica Biométrica	33
3.1 Considerações Iniciais	33
3.2 Metodologias baseadas na silhueta	35
3.3 Metodologias baseadas no modelo	43
3.4 Técnicas Híbridas	51
3.5 Considerações Finais	54
4 Metodologia adotada para o Reconhecimento da Marcha Humana	55
4.1 Considerações Iniciais	55
4.2 Modelo do Movimento Global do Corpo - GBM	56
4.3 Consideração Finais	71
5 Resultados	73

5.1	Considerações Iniciais	73
5.2	Métodos de Avaliação dos Resultados	73
5.3	Análise dos Resultados	79
5.4	Considerações Finais	102
6	Conclusão	103
6.1	Publicações	105
	Referências	107
	APÊNDICE A	111

Capítulo 1

Introdução

A marcha pode ser definida por comportamentos motores compostos por movimentos integrados do corpo humano. É um padrão cíclico de movimentos corporais que se repetem indefinidamente a cada ciclo.

A análise da marcha é muito importante na área médica para detecção e tratamento de distúrbios locomotores. Alterações básicas no padrão da marcha servem como indicadores do início de doenças, tais como: mal de Parkinson, Esclerose Múltipla e Hidrocefalia. Durante muito tempo a análise da marcha ficou restrita à área médica, mas agora está se difundindo para outras aplicações, como, por exemplo, a biometria.

Pesquisas comprovaram que os humanos têm formas distintas e especiais de caminhar (WINTER, 1991; SARKAR et al., 2005; HAVASI; ZOLTÁN; SZIRÁNYI, 2007; BOULGOURIS; ZHIWEI, 2007). A partir desta premissa a marcha humana pode ser entendida como uma importante característica biométrica.

Dentre os diversos tipos de sistemas biométricos existentes, o grande desafio está em fazer a identificação de pessoas em imagens de vídeo, em tempo real e não invasiva. Em muitas situações, técnicas biométricas convencionais, como reconhecimento pela íris, face e impressão digital, não são possíveis de serem utilizadas, por falta de cooperação do indivíduo a ser identificado ou devido a distância em que o mesmo se encontra. Por isto, há a necessidade de se avaliar procedimentos não invasivos, como, por exemplo, a marcha

humana.

A classificação e reconhecimento da marcha têm também, grande potencial de aplicação em sistemas de monitoramento de vídeo, diagnóstico médico e sistemas de vigilância.

É uma técnica que apresenta muitos desafios, pois a correta identificação do indivíduo depende de uma correta segmentação da imagem, onde diversos fatores podem influir no resultado, tais como a roupa do indivíduo, as alterações de estado de humor, a velocidade de caminhar, o tipo de calçado, o tipo de superfície (ex: concreto, grama) as variações do ângulo da câmera ou, até mesmo, o fato da pessoa estar carregando algum objeto. Outro grande desafio também está na extração das características que possam identificar a marcha de forma única.

1.1 Objetivos

Considerando-se que existem diferenças na maneira como cada pessoa caminha e que essas diferenças podem ser significativas para a identificação de um indivíduo, a proposta deste trabalho baseia-se na observação de que em uma sequência de quadros que compõe um vídeo, com apenas uma pessoa caminhando, o movimento desta, mesmo em imagens com fundo complexo, gera dados e estes são altamente correlacionados. Dessa maneira, pode-se supor que uma sequência de vídeo convenientemente segmentada, quadro a quadro de uma pessoa caminhando, forme uma única classe, em que cada quadro é um elemento desta classe. O reconhecimento de uma pessoa através de sua maneira de caminhar depende fortemente de como o formato de sua silhueta muda em relação ao tempo em uma sequência de imagens (MURASE; SAKAI, 1996).

O objetivo deste trabalho, portanto, é estabelecer uma nova metodologia que permita o reconhecimento de uma pessoa em movimento por meio do seu modo de caminhar.

O movimento do corpo humano pode ser obtido de diversas maneiras com o uso de

técnicas de processamento de imagens. Assim, propõe-se obter informações globais do movimento do corpo como um todo, a partir de quatro modelos de imagens segmentadas de vídeo da marcha humana e fundir os resultados em um modelo único que denominamos de GBM (Global Body Motion). Este modelo deverá ser capaz de melhorar os índices de reconhecimento biométrico pela marcha humana, obtidos com os modelos independentes utilizados.

1.2 Metodologia

O método proposto utiliza o Modelo de Mistura de Gaussianas (GMM) para efetuar a segmentação do fundo. O método GMM foi escolhido em função da robustez com relação à variação de iluminação e pelo bom desempenho quando utilizado em sequências de imagens de vídeo (STAUFFER; GRIMSON, 1999; KAWTRAKULLPONG; BOWDEN, 2001; RIBEIRO; GONZAGA, 2006).

Após a segmentação utiliza-se a Transformada Wavelet(TW) de Haar para redução da escala e a Análise dos Componentes Principais (PCA) para extração dos padrões. A TW de Haar aplicada às imagens reduz a dimensionalidade dos dados e conseqüentemente o tempo total de processamento, mantendo as informações necessárias sobre o movimento.

A técnica PCA é aplicada para a extração dos padrões principais que ocorrem dentro de cada classe, separando as variações inter-classes. Diferentemente dos métodos encontrados na literatura, a proposta deste estudo considera o conjunto dos quadros segmentados de cada indivíduo como uma classe distinta e cada quadro como um objeto dessa classe. Além disso, todo o objeto segmentado é utilizado para gerar o vetor de características, enquanto na maioria das abordagens, o reconhecimento é feito por meio da segmentação de partes da pessoa (BOBICK; DAVIS, 2001; LIU; SARKAR, 2006; WANG et al., 2003).

São propostos e avaliados quatro modelos de captura do movimento global do corpo humano. Finalmente é proposta a fusão dos quatro modelos distintos:

- Modelo SGW - Silhouette-Gray-Wavelet: neste modelo, cada classe é representada por uma sequência de silhuetas, em nível de cinza, geradas por meio da TW de Haar aplicada aos objetos em movimento segmentados através do Modelo de Mistura de Gaussianas (GMM). O modelo SGW possui informação do movimento global tridimensional da marcha humana nas variações de cinza durante a caminhada;
- Modelo SBW - Silhouette-Binary-Wavelet: neste modelo, cada classe é representada por uma sequência de silhuetas binárias, geradas pela Transformada Wavelet (TW) de Haar, aplicada aos objetos em movimento segmentados por GMM. O modelo SBW fornece informação bidimensional sobre o movimento global bidimensional da silhueta do corpo humano durante a caminhada;
- Modelo SEW - Silhouette-Edge-Wavelet: neste modelo, cada classe é representada por uma sequência de silhuetas de imagens de bordas, obtidas a partir do método SBW. O modelo SEW carrega informação referente ao comportamento global do contorno do corpo humano durante a caminhada;
- Modelo SSW - Silhouette-Skeleton-Wavelet: neste modelo, cada classe é representada por uma sequência de silhuetas de imagens esqueletizadas, obtidas a partir do método SBW. O modelo SSW possui informações sobre o movimento global de articulações do corpo humano e como estas se comportam durante a caminhada.

É ilustrado na Figura 1 um exemplar de cada modelo.

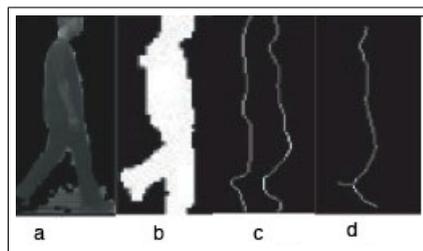


Figura 1: Exemplos de cada modelo: (a) SGW; (b) SBW; (c) SEW; (d) SSW.

1.3 Principais Contribuições

Comparando-se com o estado da arte, pode-se citar como contribuições deste trabalho:

- Nova metodologia para representação da marcha - o modelo proposto, GBM (Global Body Motion), que considera as diferentes informações de movimento do corpo humano, quando em marcha, capturadas por diferentes segmentações das imagens de seqüências de vídeos. Como as pessoas possuem características individuais quando caminham e, em geral, utilizam diferentes maneiras de se vestir, o GBM busca capturar informações globais do movimento para reconhecimento dos indivíduos;
- Fusão das características - o reconhecimento individual é realizado pela fusão de quatro maneiras distintas de representação do movimento humano. Cada modelo possui informações distintas - contidas nas imagens em nível de cinza (3D), nas imagens binarizadas (2D), no contorno da silhueta do corpo humano em movimento e no movimento das articulações (esqueleto). O processo de fusão destas características, integrando pontuações de similaridade, considera como peso o resultado das classificações independentes de cada um dos modelos.

1.4 Organização do trabalho

Este trabalho está organizado em seis capítulos. Neste capítulo, foi descrito brevemente o objetivo do trabalho e a metodologia proposta.

Capítulo 2 - A Marcha Humana

Nesse capítulo, são descritos alguns conceitos sobre a Marcha Humana.

Capítulo 3 - A Marcha Humana no Contexto e Técnica Biométrica

São descritos aspectos relacionados ao reconhecimento automático da Marcha Humana como, por exemplo, as metodologias usadas para esse fim.

Capítulo 4 -O modelo Global Body Motion - GBM

Nesse capítulo são descritas as técnicas utilizadas no desenvolvimento do modelo GBM e a metodologia envolvida.

Capítulo 5 - Resultados

São descritos os resultados obtidos com o método GBM.

Capítulo 6 - Conclusão

São apresentadas as considerações finais sobre o modelo GBM.

Capítulo 2

A Marcha Humana

2.1 Considerações Iniciais

A marcha humana pode ser definida como a *maneira de uma pessoa caminhar*. Além disso, a marcha retém características peculiares de uma pessoa. Com a utilização dessas características as pessoas podem se distinguir umas das outras. Este capítulo tem por objetivo descrever alguns conceitos relacionados à marcha humana.

2.2 Ciclo de marcha

A marcha utiliza uma sequência de repetições de movimento para mover o corpo para a frente enquanto, simultaneamente, mantém a postura estável.

Conforme o corpo move-se para frente, um membro serve como fonte móvel de apoio, enquanto o outro membro avança para a nova posição. Em seguida, os membros invertem seus papéis. Essa sequência de movimentos é denominado de ciclo de marcha (PERRY, 2005), ou seja, é o período que decorre entre o contato do calcanhar de um pé, e o contato seguinte do mesmo calcanhar. Este ciclo divide-se em duas fases - apoio e balanço -

frequentemente chamados de fases da marcha. Quando o pé está tocando a superfície é o período de apoio. Esse período começa com o contato inicial do pé na superfície, conforme ilustrado na Figura 2.

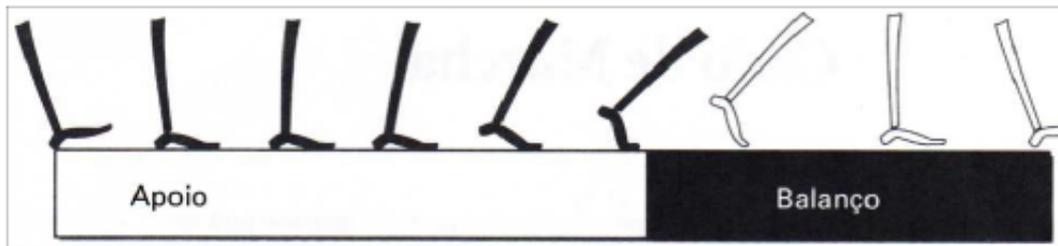


Figura 2: Representação de apoio e balanço em um ciclo de marcha.

O período de balanço é o tempo em que o pé está no ar - para o avanço do membro. A distribuição normal dos períodos de contato com o solo é de 60% para o apoio e 40% para o balanço (MURRAY; DROUGHT; KORY, 1964).

A duração das fases de apoio e balanço (duração do ciclo de marcha) varia de acordo com a velocidade da marcha do indivíduo (ANDRIACCHI; OGLE; GALANE, 1977). Na velocidade de marcha habitual de 80 m/min, os períodos de apoio e balanço representam respectivamente 62% e 38% do ciclo de marcha. Ele também tem sido identificado pelo termo de passada (MURRAY; DROUGHT; KORY, 1964). Ocasionalmente a palavra passo é utilizada, mas é inapropriada. Passada é o equivalente a um ciclo de marcha. É baseada nas ações de um membro. A duração de uma passada é o intervalo entre dois contatos iniciais sequenciais entre o solo e o mesmo membro (isto é, contato inicial direito e o próximo contato inicial direito). O passo refere-se ao intervalo entre os dois membros. Existem dois passos em cada passada (ou ciclo de marcha), conforme ilustrado na Figura 3.

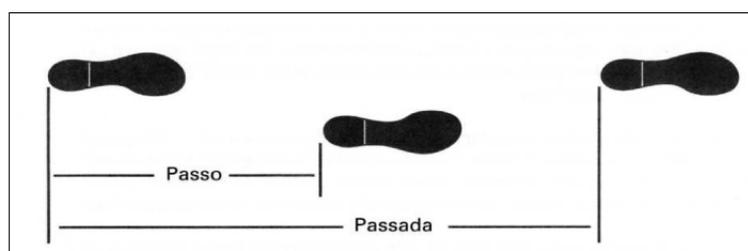


Figura 3: Comparação entre passo e passada. Fonte: (PERRY, 2005)

O intervalo entre um contato inicial por um dos pés é um passo. A mesma alternância no tempo será repetida de modo recíproco por toda a marcha.

2.2.1 Fases do ciclo de marcha

Para promover as funções básicas necessárias para caminhar, cada passada envolve uma constante mudança no alinhamento entre o corpo e o pé de suporte durante o apoio e avanço seletivo dos segmentos do membro no balanço. Essas reações resultam em uma série de padrões de movimento realizados pelo quadril, joelho e tornozelo. Pesquisas anteriores reconheceram que cada padrão do movimento relacionava-se a uma diferente exigência funcional e designaram esses padrões de movimento como fases da marcha. A marcha é dividida em oito fases e cada uma delas tem um objetivo funcional e um padrão crítico de movimento para realizar essa meta (PERRY, 2005). As divisões do ciclo da marcha podem ser visualizadas na Figura 4.

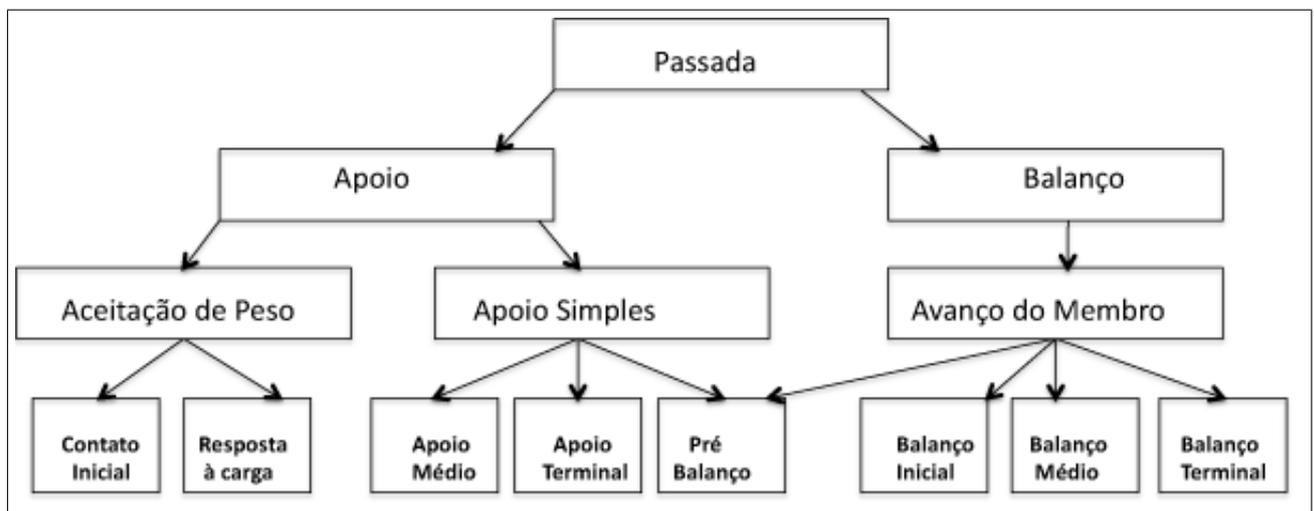


Figura 4: Divisões do ciclo de marcha. Adaptado de Perry (2005)

A análise do padrão de marcha de uma pessoa por meio das fases identifica mais diretamente a importância funcional dos diferentes movimentos que ocorrem nas articulações individuais. As fases da marcha também fornecem um meio para correlacionar as ações simultâneas das articulações individuais no contexto dos padrões da função total

do membro.

A importância do movimento de uma articulação de um indivíduo comparada ao de outra varia durante as fases da marcha. Estas fases da marcha estão ilustradas na figura 5.

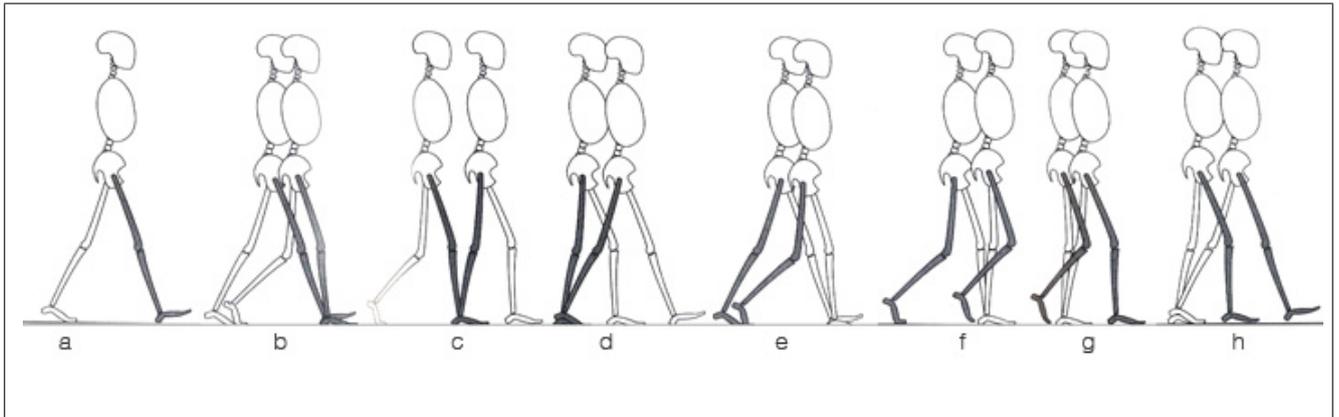


Figura 5: As fases da Marcha Humana. a)Contato Inicial; b)Resposta à Carga; c)Apoio Médio; d)Apoio Terminal; e)Pré-Balço; f)Balço Inicial; g)Balço Médio; h)Balço Terminal. Adaptado de Perry (2005).

Tarefas e objetivos relacionados a cada fase:

1. Tarefa A: Aceitação do peso

Esta é a tarefa de maior exigência no ciclo de marcha. Três padrões funcionais são necessários: absorção de choque, estabilidade inicial do membro e preservação da progressão. O desafio é a transferência abrupta do peso do corpo para um membro que acabou de terminar o balanço para frente e tem um alinhamento estável. Fases da marcha envolvidas:

- Fase 1 - Contato inicial

Nesta fase, o quadril está fletido e um dos membros está com o joelho estendido e o tornozelo dorsifletido em neutro. O contato com o solo é realizado pelo calcanhar. O outro membro está no final do apoio terminal. O objetivo desta fase é posicionar o membro para iniciar o apoio com o rolamento do calcanhar;

- Fase 2 - Resposta à carga

O peso do corpo é transferido sobre o membro anterior. Usando o calcanhar como um rolamento, o joelho é fletido para absorção de choque. A flexão da plantar dos tornozelos limita o rolamento do calcanhar pelo contato do antepé com o solo. O membro oposto está na fase de pré-balanço. Tem por objetivo a absorção do choque, a estabilidade para recepção de peso e a preservação da progressão.

2. Tarefa B: Apoio Simples

A elevação do outro pé para o balanço inicia o intervalo do apoio simples para o membro de apoio. Essa fase continua até que o pé oposto toque novamente o solo. Duas fases estão envolvidas no apoio simples:

- Fase 3 - Apoio médio

Na primeira metade do apoio simples, um dos membros avança sobre o pé estacionário por meio da dorsiflexão do tornozelo (rolamento do tornozelo), enquanto o joelho e o quadril estendem-se. O membro oposto está avançando na sua fase de balanço médio. Esta fase tem por objetivo a progressão sobre o pé estacionário e a estabilidade do tronco e do membro.

- Fase 4 - Apoio Terminal

Durante a segunda metade do apoio simples, o calcanhar eleva-se e o membro avança sobre o rolamento do antepé. O joelho aumenta sua extensão e então começa a fletir levemente. O aumento da extensão do quadril coloca o membro em uma maior posição de queda. O outro membro está no balanço terminal. O apoio terminal tem como objetivo a progressão do corpo além do pé de sustentação.

3. Tarefa C: Avanço do membro

Para suprir as altas exigências do avanço do membro, uma postura preparatória inicia-se no apoio. Então, este membro oscila através de três posturas, conforme ele eleva-se, avança e prepara-se para o próximo intervalo de apoio. São quatro as fases envolvidas nesta tarefa:

- Fase 5 - Pré-balanço

O contato com o solo por um dos membros dá início ao duplo apoio terminal. O membro responde com o aumento da flexão plantar do tornozelo, maior flexão do joelho e perda da extensão do quadril. O outro membro está na posição de carga. O objetivo desta fase é posicionar o membro para balanço.

- Fase 6 - Balanço inicial

O pé é elevado e o membro avança pela flexão do quadril e o aumento da flexão do joelho. O joelho fica parcialmente dorsiflexo. O outro membro está no início do apoio médio. Os objetivos são: liberar o pé do solo e avançar o membro a partir de sua posição de queda.

- Fase 7 - Balanço médio

Avanço do membro anterior à linha do peso do corpo é obtido por uma flexão adicional do quadril. O joelho estende em resposta à gravidade, enquanto o tornozelo continua dorsifletindo em neutro. O outro membro está no final do apoio. Os objetivos são: avanço do membro e a liberação do pé do solo.

- Fase 8 - Balanço terminal

O avanço do membro é completado pela extensão do joelho. O quadril mantém sua flexão inicial e o tornozelo permanece dorsifletido em neutro. O outro membro está no apoio terminal. Os objetivos desta fase são: completar o avanço do membro e preparar o membro para o apoio.

2.3 Considerações finais

Neste capítulo, foram descritos alguns conceitos sobre a marcha humana, visto que este é o objeto principal de estudo deste trabalho. Por meio da análise da marcha humana é possível detectar características psicológicas, comportamentais, disfunções patológicas e, como será tratada neste trabalho, a marcha como uma importante característica bio-

métrica.

Como pode ser visto, as fases da marcha humana obedecem a um padrão de comportamento, mas cada indivíduo que o realiza produz imagens diferentes durante o movimento. Este fato corrobora com a hipótese que durante um ciclo de marcha as silhuetas obtidas por segmentação formam classes distintas de *frames* que podem ser agrupadas em base de vídeos.

Capítulo 3

A Marcha Humana no Contexto de Técnica Biométrica

3.1 Considerações Iniciais

Os primeiros estudos sobre a marcha humana foram apresentados por Johansson (1975). Ele apresentou a um grupo de observadores imagens reduzidas a um ponto de luz, na forma de um display. Seus experimentos sugerem que os seres humanos têm implícita alguma noção do movimento humano e pode-se reconhecer dados temporais dentro deste contexto. Estudos posteriores demonstraram que não somente a forma humana pode ser distinguida a partir de pontos de luz, mas também movimentos, como: saltar, caminhar e inclinar (DITTRICH, 1993). Em um dos estudos e experimentos conduzidos por Cutting e Proffi (1981), concluiu-se que o sexo de um indivíduo pode ser determinado a partir dos movimentos humanos e suas características inerentes. Cutting e Kozlowski (1977) demonstraram que observadores podiam reconhecer-se com segurança, assim como seus amigos, a partir de um display de pontos de luz dinâmicos.

Basicamente as técnicas propostas para o reconhecimento da marcha são divididas em dois grupos principais: metodologias baseadas na silhueta (*Silhouette-Based*) e meto-

dologias baseadas no modelo (ou na dinâmica) (*Model-Based*), conforme é ilustrado na figura 6. Na abordagem baseada na silhueta, a forma desta e seus atributos são as características mais importantes. Assim sendo, o reconhecimento do indivíduo é realizado com medições que refletem a sua forma. As abordagens baseadas no modelo têm por objetivo obter o movimento do tronco e/ou das pernas. Ao contrário de abordagens baseadas em silhuetas, esta geralmente se concentra na dinâmica do movimento, omitindo sua forma corporal.

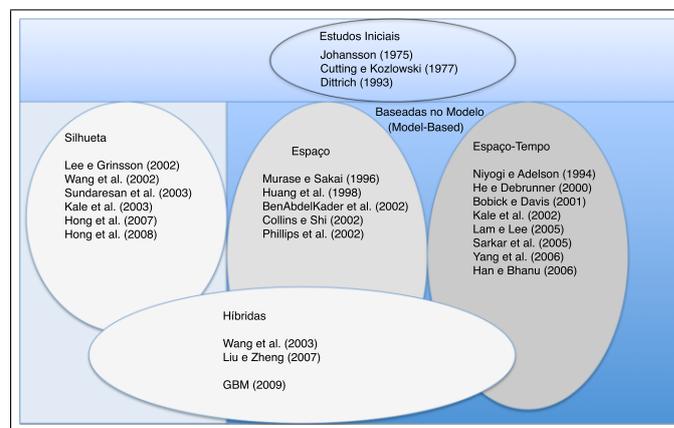


Figura 6: Métodos de Reconhecimento da Marcha.

A maioria das técnicas baseadas no modelo pode ser dividida em duas classes principais (BENABDELKAER; CUTLER; DAVIS, 2002): métodos baseados no espaço e métodos baseados no espaço-temporal.

As técnicas baseadas no espaço consideram o movimento da marcha como sendo composto por uma sequência de poses estáticas do corpo e o reconhecimento é realizado por meio de observações de variações temporais com respeito a estas poses.

Os métodos baseados no espaço-temporal visam capturar os padrões espaço-temporais discriminantes em uma sequência de marcha para identificação humana.

Existem técnicas que agregam os dois modelos principais, proporcionando uma metodologia híbrida.

Nas seções seguintes são descritos os principais trabalhos de cada metodologia.

3.2 Metodologias baseadas na silhueta

As metodologias baseadas na silhueta tendem a oferecer mais velocidade e simplicidade, mas são apenas indiretamente ligadas à marcha. Além disso, são mais suscetíveis ao ruído e demais variações na mesma (vestimenta, porte de objetos e outros).

Um dos estudos para reconhecimento da marcha baseado na silhueta é conduzido por Lee e Grimson (2002). Em sua análise, a silhueta é dividida em sete regiões e cálculos estatísticos são realizados com objetivo de extrair características da forma do indivíduo, em cada uma destas sete regiões. Os autores utilizam as características físicas agregadas a características da aparência, tais como: proporção do dorso, vestimenta, balanço dos braços e ritmo de caminhar. A intenção é que o vetor de padrões da marcha tenha as seguintes propriedades: habilidade de descrever a silhueta em um nível mais compacto que a descrição do corpo inteiro sem a necessidade de segmentar os membros individualmente; robustez ao ruído e simplicidade de representação. Uma elipse é encaixada em cada uma das sete regiões da silhueta, conforme é ilustrado na figura 7.

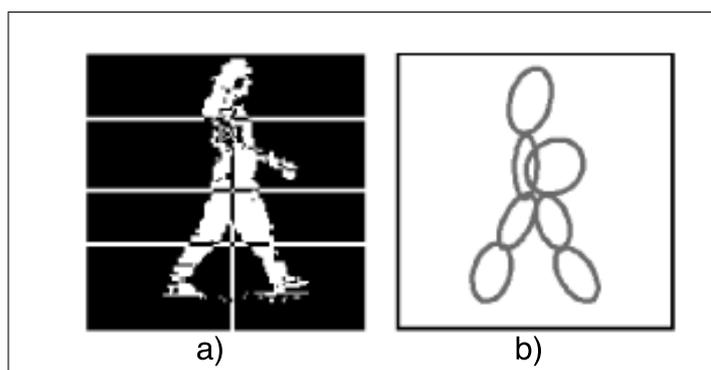


Figura 7: A silhueta de um indivíduo dividido em 7 regiões (a) e elipses (b) são encaixadas em cada região. Fonte: (LEE; GRIMSON, 2002).

Os padrões extraídos de cada *frame* de uma sequência em movimento consistem dos padrões extraídos de cada uma destas sete regiões. No entanto, além das características do formato das silhuetas, é necessária uma representação concisa em relação ao tempo. Assim sendo, são computados dois tipos de vetores de padrões:

- um vetor de padrões contendo a média e desvio padrão das regiões em relação ao tempo e;
- magnitude e fase de cada região relatada, para a frequência dominante de caminhada.

A média dos padrões descreve a média das características contidas em cada elipse em cada uma das sete regiões do corpo. Juntas estas sete elipses descrevem a média do formato do corpo. O desvio padrão descreve, de modo geral, as mudanças causadas pelo movimento. Os componentes dos padrões da *magnitude* medem os valores das mudanças em cada uma das sete regiões devido ao movimento de caminhada e os componentes da *fase* medem o atraso de tempo de diferentes regiões da silhueta. Estes dois vetores de padrões são utilizados para identificação de pessoas e também para classificação do gênero masculino ou feminino. Para validar os padrões obtidos nas sequências de marchas são utilizadas duas bases de imagens de marcha. Sendo, uma delas, uma base própria, obtida em ambiente interno, com diferentes fundos de cenas, capturadas em quatro dias diferentes no intervalo de dois meses. Esta base consiste de 24 pessoas, sendo 10 mulheres e 14 homens caminhando em velocidade normal e em passos largos. Para cada pessoa, são obtidas duas sequências de caminhada: ida e volta. As sequências de caminhada de direções opostas são rotacionadas, assim, todas as sequências possuem a mesma direção. Ao todo, 194 sequências foram obtidas, com uma média de 8 sequências por pessoa. A outra base de sequências de marcha pertence ao CMU - *Gait Database* (GROSS; SHI, 2001), sendo utilizadas 370 sequências.

O desempenho estatístico do método é medido por meio da pontuação cumulativa de acertos (CMS) (PHILIPS et al., 2000). Para os 41 melhores padrões (obtidos por meio do ANOVA¹), obtêm-se 100% de identificações corretas para a primeira busca. Quando são utilizados todos os padrões dos vetores, são alcançados 97% de classificações corretas para a primeira busca. Os autores confrontaram sequências de treinamento obtidas em um dia (dias A, B, C ou D) contra todas as sequências obtidas em outro dia ((B,C,D), (A,C,D),

¹Análise de variância entre grupos.

(A,B,D) ou (A,B,C)). Os desempenhos de classificação mais baixos são aqueles nos quais as sequências de testes são de pessoas vestindo um tipo de roupa substancialmente diferente das sequências treinadas. Os resultados obtidos para o vetor de padrões dos componentes espectrais (magnitude e fase) são inferiores aos resultados obtidos pelo vetor de padrões baseados na aparência - quando as sequências de busca e treinamento são aquelas obtidas em um mesmo dia; mas, são significativamente melhores para sequências de busca e treinamento obtidas em dias distintos. Este comportamento resulta das características dos próprios vetores de padrões. Vetores de características baseados na aparência e nas características físicas são muito sensíveis às possíveis alterações do comportamento e aparência do indivíduo. Apesar destes padrões proporcionarem em alguns casos alta taxa de desempenho, seria interessante a combinação dos dois vetores de padrões. O algoritmo descrito pode ser utilizado em conjunto com outros algoritmos biométricos (face, íris e outros), com o intuito de obter melhor desempenho de reconhecimento.

Wang et al. (2002) propõem um método para reconhecimento da marcha baseado na análise estática da silhueta. Com o objetivo de reduzir as informações redundantes contidas na silhueta, apenas o contorno desta é utilizado para representar os padrões temporais da marcha. A borda é obtida por meio de um algoritmo de bordas baseado em conectividade. Após a extração da borda, o próximo passo é a obtenção do centro da silhueta, que será considerado a origem da silhueta no espaço 2D. Estes formatos de silhuetas são representados como um conjunto de pontos de pixels, obtidos no sentido anti-horário, pertencentes à borda em uma coordenada complexa comum. É ilustrado na figura 8, a obtenção dos pontos da borda.

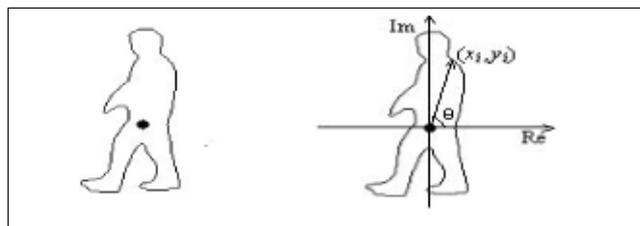


Figura 8: Ilustração da representação do formato da silhueta. Adaptado de Wang et al. (2002).

Cada silhueta pode ser descrita como um vetor de números complexos com N pontos

da borda. As representações únicas do formato de uma sequência de marcha são utilizadas para encontrar um formato médio, sendo a assinatura da marcha. Os experimentos para avaliação do algoritmo proposto são realizados com a base NLPR (CASIA, 2005). São utilizados três classificadores: classificador do vizinho mais próximo (NN), classificador dos k -vizinhos mais próximo (KNN, $k=3$) e o classificador do vizinho mais próximo com exemplar de classe (ENN). Para medir a similaridade entre duas sequências de marcha, utiliza-se a métrica *Procrustes* de Distância Média da Forma (MSD) (KENT, 1992). As taxas de classificações corretas (CCR) para os ângulos 0° , 45° e 90° estão dispostas na tabela 1.

Tabela 1: Taxa de Acertos (CCR). Método utilizando a distância *Procrustes* com base NLPR.

Classificador	0°	45°	90°
k=1 (NN)	71,25%	72,50%	81,25%
k=3 (3NN)	72,5%	73,75%	80,0%
ENN	88,75%	88,75%	90,0%

O método possui algumas propriedades interessantes, tais como o baixo custo computacional, pois utiliza apenas o contorno da silhueta. A avaliação do método foi realizada em uma base de dados grande (20 pessoas, 240 sequências), com diferentes ângulos, enquanto que alguns trabalhos realizados anteriormente (MURASE; SAKAI, 1996; HUANG; HARRIS; NIXON, 1998; NIYOGI; ADELSON, 1994; J.LITTLE; J.BOYD, 1998; SHUTLER; NIXON; HARRIS, 2000) utilizam-se de uma base de sequências de marchas bem menor. Apesar das taxas de reconhecimento serem relativamente baixas, em relação a estes trabalhos citados, os resultados obtidos em uma base de dados maior são mais convincentes. No entanto, o método baseia-se apenas no formato da silhueta, sendo suscetível às variações comportamentais da pessoa. Além disso, os autores não especificam se foi utilizado algum método para avaliar a qualidade dos resultados.

Sundaresan, RoyChowdhury e Chellappa (2003) propõem um sistema genérico para reconhecimento da marcha baseado no *Hidden Markov Model* HMM. O HMM torna-se uma ferramenta possível de ser utilizada, devido principalmente ao fato de a marcha de

uma pessoa poder ser entendida como um conjunto de poses, em uma sequência que possui uma base de natureza probabilística. As poses que um indivíduo adota durante o movimento são únicas para cada e fornecem um meio de discriminação. A natureza probabilística do HMM provê robustez ao modelo. Para avaliação do algoritmo é utilizada a base de marcha NIST / USF ². Esta base contém sequências de vídeo de 75 pessoas, caminhando em uma trajetória elíptica. Estas sequências são coletadas com as seguintes variações: superfície (grama/concreto), direção da câmera (direita/esquerda) e tipo de sapato(I/II). Assim, há oito possíveis combinações para cada pessoa, formando-se assim oito diferentes subconjuntos, conforme é mostrado na tabela 2.

Tabela 2: Combinação das condições de caminhada para os experimentos de Sundaresan, RoyChowdhury e Chellappa (2003).

Conjunto	Combinação	Descrição
A	G,A,L	Grama, sapato A e câmera à esquerda
B	G,B,R	Grama, sapato B e câmera à direita
C	G,B,L	Grama, sapato B e câmera à esquerda
D	C,A,R	Concreto, sapato A e câmera à direita
E	C,B,R	Concreto, sapato B e câmera à direita
F	C,A,L	Concreto, sapato A e câmera à esquerda
G	C,B,L	Concreto, sapato B e câmera à esquerda

É ilustrado na figura 9 um exemplo de cada *frame* para a base USF.

Os parâmetros do HMM são treinados para um conjunto com as combinações: superfície grama, sapato A e câmera à direita (G,A,R). Em cada experimento, tentou-se identificar sequências em cada um dos sete subconjuntos a partir dos modelos obtidos do conjunto de treinamento. Os experimentos são realizados para três diferentes medidas de distância de similaridade: distância Euclidiana, distância do produto interno (IP) e somatório das distâncias das diferenças absolutas (SAP).

Os resultados dos experimentos estão dispostos na tabela 3, na forma da Pontuação Cumulativa de Acertos (CMS) (PHILIPS et al., 2000) para o rank 1.

Como pode ser observado, por meio da tabela 3, os melhores desempenhos são para

²<http://marathon.csee.usf.edu>

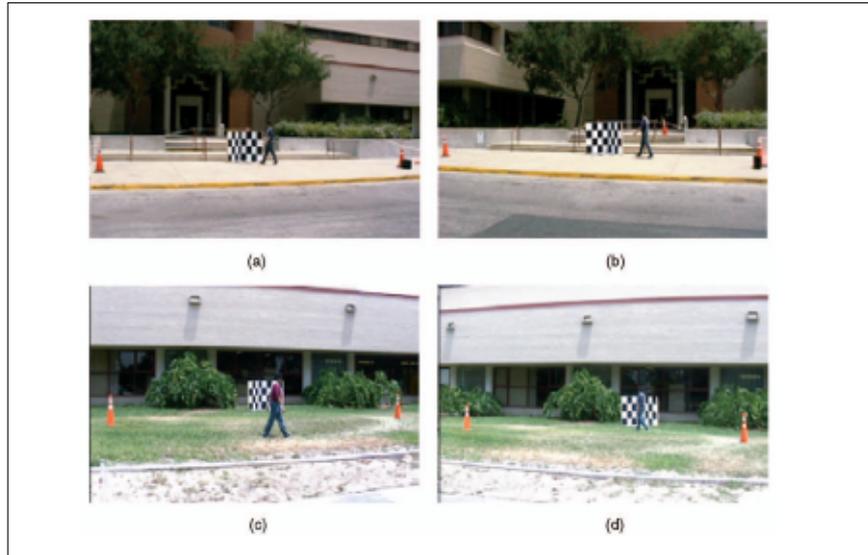


Figura 9: (a) câmera posicionada à esquerda para superfície de concreto, (b) câmera à direita para superfície de concreto, (c) câmera à esquerda para superfície de grama, (d) câmera à direita para superfície de grama . Fonte: (Sarkar et al. (2005)).

os conjuntos A e B, independente da métrica de distância utilizada.

Tabela 3: Desempenho do sistema em porcentagem para as três métricas de distância, para o rank 1. Os números entre colchetes denotam o número de pessoas contidas em cada conjunto.

Conjunto	IP%	Euclidiana%	SAP%
A[66]	99	99	98
B[37]	89	89	89
C[37]	78	78	75
D[62]	36	29	23
E[39]	29	28	21
F[62]	24	19	16
G[38]	18	14	15

Assim, percebe-se que o HMM não conseguiu capturar as variações contidas nos demais subconjuntos. No conjunto de treinamento, a superfície de caminhada é a grama, com ângulo da câmera à direita. Alterando-se apenas o ângulo da câmera, há 99% de acertos para a primeira busca. Pode-se observar que o fator que mais influência tem durante a caminhada, além do tipo de sapato, é a superfície de caminhada. Comparando o trabalho de Sundaresan, RoyChowdhury e Chellappa (2003) com o de Wang et al. (2003), que utiliza a mesma base de imagens, este o supera, em termos de acertos pertencentes

às primeiras buscas, para os conjuntos A, B e C. Outros trabalhos também comprovam a utilização do HMM para análise da marcha (HE; DEBRUNNER, 2000; KALE et al., 2002).

No trabalho proposto por Kale et al. (2003), considera-se como representação da marcha a largura do contorno externo de uma silhueta binarizada da pessoa em movimento. Para se obter o vetor de padrões, que caracteriza a assinatura da marcha, inicialmente computa-se o tamanho de todas as linhas que compõe a silhueta. Este tamanho é obtido efetuando-se a diferença das localizações dos pixels mais a esquerda e mais a direita da respectiva borda. Desta maneira, é formado um vetor para cada *frame* da sequência de marcha. Os vetores são superpostos em um único vetor para cada indivíduo. A superposição dos vetores não captura os aspectos temporais. A variação de cada componente do vetor é considerada como a assinatura da marcha para um indivíduo. Para imagens de baixa qualidade e baixa resolução, a extração do contorno da silhueta pode não ser confiável. Então, nestes casos, o vetor de largura pode também não ser confiável como sendo a única representação da marcha. A avaliação do método é realizada com três bases de marcha: CMU (GROSS; SHI, 2001), USF³ e UMD⁴. Para a base UMD, o melhor resultado obtido é de 79,07% de acertos. Para a base CMU, o melhor resultado é de 95,8% de acertos, considerando a primeira busca (Rank =1). A base USF fornece resultados de desempenho inferiores as duas outras bases. Isto se deve ao fato de que a superfície tem maior influência na classificação do que na velocidade de caminhada e no ângulo da câmera. Posteriormente, o método foi avaliado com a base NLPR (CASIA, 2005), para os ângulos de 0° 45° e 90°. Os resultados com esta base não diferem muito das outras três avaliadas. O melhor resultado é para o ângulo de 45°, com 92,5% de acertos, para a primeira busca. Com estes resultados, pode-se observar que o método é razoavelmente robusto a variações de velocidade e ângulo e as informações contidas no vetor de padrões possui informações discriminativas da marcha humana.

Tendo como base o trabalho de Kale et al. (2003), Hong et al. (2007) propõe um novo método para reconhecimento da marcha. A diferença principal está no cômputo do vetor

³<http://marathon.csee.usf.edu>

⁴<http://degas.umiacs.umd.edu/hid>

de padrões, denominado de vetor de massa (*Mass Vector*). O vetor de massa é definido como o número de pixels de valores não zero de cada linha da silhueta binarizada de um indivíduo em movimento. Sequências dos vetores ordenados, com relação ao tempo, são usados para representar a marcha de cada indivíduo. Para efetuar a verificação de um exemplar de teste com a base treinada, utiliza-se o método DTW - *Dynamic Time-Warping*. O algoritmo proposto é avaliado com a base NLPR (CASIA, 2005) e demonstra um desempenho superior ao método proposto por Kale et al. (2003), conforme pode ser observado na tabela 4.

Tabela 4: Pontuação Acumulativa de Acertos para o Rank 1.

Ângulo	Vetor	Rank =1(%)
0°	Largura (KALE et al., 2003)	82,5
	Massa (HONG et al., 2007)	96,25
45°	Largura (KALE et al., 2003)	92,5
	Massa (HONG et al., 2007)	96,25
90°	Largura (KALE et al., 2003)	77,5
	Massa (HONG et al., 2007)	88,75

Com estes resultados, pode-se concluir que o vetor de massa possui um potencial discriminativo superior ao método de Kale et al. (2003), para a base NLPR. Apesar de seu desempenho ser superior e de que padrões extraídos das silhuetas forneceram um certo potencial discriminativo, as informações dinâmicas da marcha são também fundamentais para o reconhecimento.

Hong et al. (2008) prosseguem com a pesquisa de reconhecimento da marcha, propondo desta vez um método de fusão de características, método este denominado de MBCM (*Multi-Bipolarized Contour Mean*). A técnica consiste na fusão de quatro componentes da silhueta humana binarizada: a média do contorno vertical positivo, a média do contorno vertical negativo, a média do contorno horizontal positivo e a média do contorno horizontal negativo. Para se obter estes padrões, primeiro é encontrado o pixel localizado no centro da silhueta. A partir daí, calcula-se a diferença entre os pixels localizados: mais à direita da borda, mais à esquerda da borda, mais ao topo da borda superior e mais abaixo da borda inferior, e o pixel central. Este processo é realizado para cada *frame* pertencente a

uma dada sequência. Calcula-se a média destes pontos e estes são utilizados como o vetor de padrões da marcha humana.

A técnica PCA é aplicada nestes vetores para reduzir a dimensionalidade dos dados. A classificação é realizada pelo classificador do vizinho mais próximo (NN). Experimentos são realizados com a base NLPR (CASIA, 2005), para validação do método. A tabela 5 mostra a taxa de classificação correta para a fusão dos vetores verticais (vertical negativo com vertical positivo) - C_v , horizontais (horizontal positivo com horizontal negativo) - C_h e verticais com horizontais - C_{vh} .

Tabela 5: Taxa de classificação correta (CCR) para a base NLPR.

Vetor de padrões	Ângulo		
	0°	45°	90°
C_v	0,91	0,97	0,83
C_h	0,72	0,91	0,81
C_{vh}	0,86	0,96	0,85

Quando os testes são realizados nos quatro vetores separadamente, as taxas de classificações corretas indicam um baixo desempenho, especialmente para os vetores de contornos horizontais. No entanto, quando realizada a fusão, o melhor desempenho permanece para os vetores de contorno vertical. O método de fusão otimiza alguns resultados, mas esperava-se um melhor desempenho para a fusão dos quatro vetores.

3.3 Metodologias baseadas no modelo

As metodologias baseadas no modelo tendem a capturar a forma e a dinâmica da marcha, para se extrair o vetor de padrões. As metodologias deste grupo podem ser divididas em duas classes.

3.3.1 Técnicas baseadas no espaço

Utilizando o formato das silhuetas e suas alterações temporais, Murase e Sakai (1996) apresenta um método de verificação, com base na representação do auto-espaço (EST) para distinguir diferentes marchas. As imagens, devidamente segmentadas, são projetadas no auto-espaço. O reconhecimento realizado em uma base de dez sequências de sete pessoas mostrou uma taxa de classificações corretas de 100% para 16 auto-vetores e de 88% para oito auto-vetores. O método foi principalmente projetado para a caracterização genérica de objetos em movimento e utilizou-se da marcha humana como um exemplo para seu trabalho.

Nesta linha de pesquisa, Huang, Harris e Nixon (1998) ampliaram a pesquisa de Murase e Sakai (1996) adicionando a Análise Canônica (CST) com transformação do auto-espaço (EST) para a extração dos padrões, método utilizado para reduzir a dimensionalidade dos dados e para garantir a separabilidade das classes das diferentes sequências de marcha. Cada imagem da sequência, pertencente ao espaço de alta dimensão, é projetada em um único ponto do espaço canônico de baixa dimensão. O reconhecimento da marcha é realizado neste espaço. Neste caso, o reconhecimento da marcha torna-se mais simples.

Os experimentos são realizados em uma pequena base de dados, composta por vinte e cinco sequências de vídeo de cinco pessoas, sendo que cada uma possui cinco sequências. Para treinamento são separados cinco sequências, uma de cada pessoa. Os resultados dos experimentos demonstram que as cinco sequências de treinamento são transformadas em cinco agrupamentos separados no novo espaço. A taxa de reconhecimento para as vinte sequências de teste atingem 100% de acertos, no entanto, a base de testes é pequena. É interessante a realização de testes em uma base maior e com diferentes variações de condições, tais como: ângulo da câmera, vestimenta, superfície de caminhada, iluminação e outros, para uma comprovação mais efetiva do método.

Baseado na dinâmica do movimento das pessoas, o Eigengait (BENABDELKAER; CUTLER; DAVIS, 2002) mapeia uma sequência de imagens - com fundo previamente extraído - de pessoas caminhando, em pontos de similaridade (SP). Estes pontos são obtidos por meio da correlação de cada par de imagens pertencentes à sequência, formando uma matriz de auto-similaridades. Estes padrões obtidos de duas dimensões codificam a projeção planar da dinâmica da marcha e conseqüentemente a assinatura da mesma. O método possui a vantagem de ser robusto para erros de rastreamento e segmentação da pessoa em movimento e para variações de vestuário e fundo de cena. Para avaliação do método foram utilizadas duas bases de vídeo de marcha. A primeira base consiste de sequências de vídeo de 44 pessoas caminhando em ambiente externo. Quatro sequências de cada pessoa foram obtidas em dois dias diferentes. A figura 10 mostra 4 exemplares desta base de uma mesma pessoa.



Figura 10: Exemplo de quatro sequências externas de uma pessoa. Fonte: (BENABDELKAER; CUTLER; DAVIS, 2002) .

A segunda base consiste de sequências de vídeo de 7 pessoas caminhando em uma esteira, obtidas com 8 diferentes ângulos e em 7 dias diferentes. A figura 11 mostra um exemplar de cada ângulo desta base.

Para a primeira base, obtida em ambiente externo, a taxa de classificação correta é de 77%. Considerando que há grande variação de iluminação e as cenas possuem um fundo complexo, este é um bom resultado. Para a segunda base, o melhor resultado é de 78% para as sequências frontais-paralelas. Para as demais vistas, a taxa de classificação correta fica em torno de 65%, demonstrando a suscetibilidade do método em relação ao

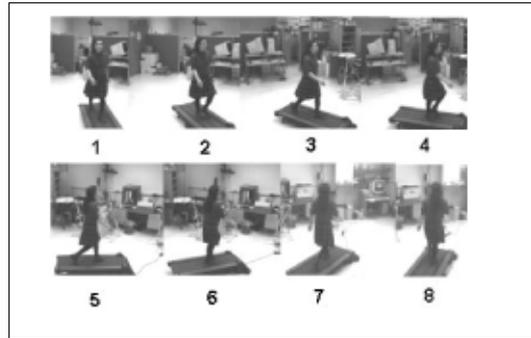


Figura 11: Segunda base de dados. Sequências com oito variações de ângulo, ambiente interno. Fonte:(BENABDELKAER; CUTLER; DAVIS, 2002).

ângulo e à velocidade de caminhada.

Posteriormente, o desempenho deste método foi avaliado com a base NLPR (CASIA, 2005), com a métrica da pontuação cumulativa de acertos (CMS) e apenas 72,5% das classificações corretas pertencem à primeira busca (rank =1). Este resultado é inferior aos métodos propostos por Collins e Shi (2002), Lee e Grimson (2002), Phillips et al. (2002), Wang et al. (2003), Kale et al. (2003), Hong et al. (2007).

3.3.2 Técnicas espaço-temporais

Uma das primeiras técnicas para reconhecimento de figuras em movimento foi proposta por Niyogi e Adelson (1994). Por meio de sua técnica, é possível distinguir caminhantes por meio da extração dos padrões espaço-temporais da marcha, a partir de uma curva-combinada denominada "snake". Os diferentes padrões dos movimentos das pernas e braços no espaço XT (Translação e Tempo) são processados para determinar o movimento do contorno do corpo. A classificação obtida ficou em torno de 60% e no melhor caso em 80%. Apesar do número de amostras ser pequeno e do sistema apresentar algumas características indesejáveis, a técnica contribuiu para novas pesquisas na área, pois provou que é possível extrair características próprias da marcha, podendo-se distinguir indivíduos.

Mais recentemente, o modelo temporal foi vastamente estudado por Bobick e Davis (2001), como uma representação compacta de seqüências de imagens binárias que descrevem o movimento humano. Seus estudos comprovam a viabilidade da representação temporal por meio de dois componentes. O primeiro componente é um valor binário que indica a presença de movimento e o segundo é uma função de atualização deste movimento. A partir destes componentes, os autores propõe a construção de dois modelos: MEI (*motion-energy image*) e MHI (*motion-history image*). O MEI registra onde o movimento ocorreu em uma imagem sintetizada, gerada a partir de uma seqüência de imagens de marcha em movimento; o próximo passo é gerar o MHI. O MHI assume que o valor da intensidade do pixel é em função do histórico temporal do movimento, ou seja, é um vetor de valores escalares, onde os pixels com maior intensidade representam o movimento mais recente. Para avaliar o método, são utilizados vídeos de pessoas executando exercícios aeróbicos. A figura 12 mostra um *frame*, com vista frontal, para cada um dos movimentos aeróbicos juntamente com o MEI.

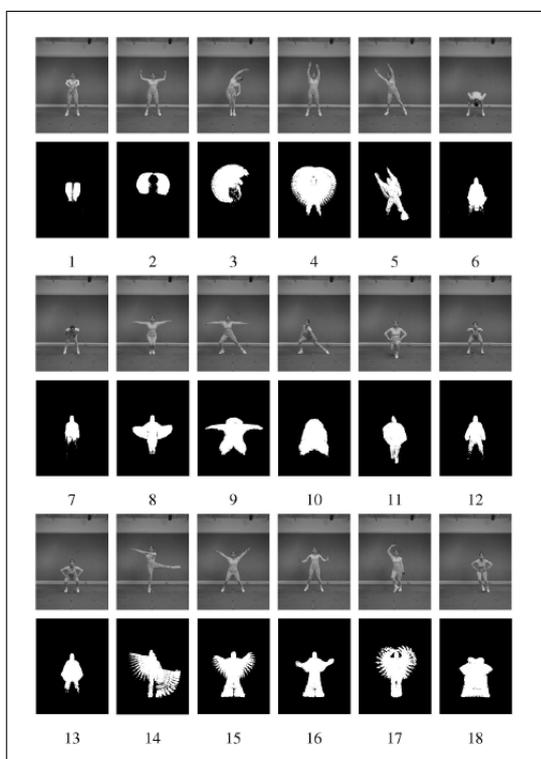


Figura 12: Exemplo de um *frame* e seu respectivo MEI, a partir da vista frontal, para cada um dos 18 exercícios aeróbicos utilizados para testar a nova representação do movimento. Fonte: (BOBICK; DAVIS, 2001).

O método automaticamente desempenha a segmentação temporal, é invariante à mudanças lineares de velocidade e sua execução é em tempo real. Esta pesquisa contribui para o desenvolvimento de muitas outras na análise da marcha humana, na tentativa de recuperar padrões para o reconhecimento desta (LAM; LEE, 2005; HAN; BHANU, 2006; LIU; ZHENG, 2007).

Laptev (2005) estuda a noção de pontos de interesse espacial para o domínio espaço-temporal e mostra como os padrões resultantes frequentemente refletem eventos interessantes. Estes, por sua vez, podem ser usados para uma representação compacta dos dados no vídeo, assim, como, para a interpretação de eventos espaço-temporais. Estes aspectos podem ser pesquisas relacionadas à marcha humana.

Inspirado no modelo MHI descrito por Bobick e Davis (2001), é proposto por Lam e Lee (2005) um método denominado *Motion Silhouette Images* - MSI. O MSI é uma imagem em nível de cinza, na qual estão embutidas as informações espaciais e temporais do movimento. Assim como no MHI, a intensidade do pixel na imagem é uma função do histórico temporal do movimento daquele pixel. A Análise dos Componentes Principais é aplicada para reduzir a dimensionalidade dos dados das sequências MSI e otimizar a separabilidade entre as classes. A base SOTON (SHUTLER et al., 2002), desenvolvida pela Universidade de Southampton, é utilizada para a realização dos experimentos. Esta base contém sequências de vídeo de 115 pessoas, totalizando 2128 sequências de marcha. Para avaliar os efeitos sobre o reconhecimento com diferentes números de exemplares para treinamento são conduzidos três testes:

- (A) 90% das sequências de imagens em cada classe são utilizadas para treinamento e 10% para teste;
- (B) 75% das sequências de imagens para treinamento e 50% para teste;
- (C) 50% das sequências de imagens para treinamento e 15% para teste.

A tabela 6 mostra o *ranking* estatístico para os três testes realizados. A melhor taxa de classificação correta é de 87,15% para o treinamento (A).

Tabela 6: Taxa de reconhecimento para diferentes números de exemplares de treinamento e teste, para a base SOTON..

	TOP 1	TOP 5	TOP10
(A) 90% treinamento 10% teste	87,15%	95,18%	95,98%
(B) 75% treinamento 15% teste	84,85%	93,76%	95,37%
(C) 50% treinamento 50% teste	82,10%	92,31%	94,94%

Os resultados demonstram que o MSI possui potencial discriminativo, como padrão de característica, para reconhecimento da marcha. Além disso, possui a grande vantagem na redução de espaço de armazenamento das sequências de marcha. A base SOTON (SHUTLER et al., 2002), originalmente com 1, 92GB, foi reduzida a 1, 92KB após ser transformada em sequências MSI. Os autores implementaram o algoritmo de Wang et al. (2003) e o comparam com o método MSI. Os resultados obtidos estão dispostos na tabela 7. Como pode ser observado, o MSI possui um desempenho inferior ao método comparado, para o primeiro *rank*. Isto deve-se principalmente ao fato de o método implementado por Wang et al. (2003) ser um método híbrido, ou seja, combina características estáticas e dinâmicas da marcha.

Tabela 7: Taxa de reconhecimento para diferentes números de exemplares de treinamento e teste, para a base SOTON (SHUTLER et al., 2002).

Método Rank	MSI			Wang et al. (2003)		
	TOP 1	TOP 5	TOP10	TOP 1	TOP 5	TOP10
50% treinamento 50% teste	82,10%	92,31%	94,94%	93,16%	97,56%	98,31%
75% treinamento 15% teste	84,85%	93,76%	95,37%	94,47%	97,33%	98,04%
90% treinamento 10% teste	87,15%	95,18%	95,98%	94,38%	97,56%	98,39%

Continuando nesta linha de pesquisa e utilizando-se dos conceitos estudados por Bobick e Davis (2001), uma nova representação temporal para análise e reconhecimento da marcha é proposto por Han e Bhanu (2006). Sua representação denominada *Gait Energy Image* - GEI - é uma imagem em nível de cinza que armazena as informações espaço-temporais da marcha contidas em um ciclo. Dada uma silhueta binária pre-processada $B_t(x, y)$ no tempo t , em uma sequência de marcha, a imagem GEI em nível de cinza é dada por: $G(x, y) = \frac{1}{N} \sum_{t=1}^N B_t(x, y)$. Sendo N o número de *frames* de um ciclo completo

em uma sequência de marcha; t o respectivo *frame* da sequência e x e y são os valores das coordenadas $2D$ da imagem. Desta forma, o GEI é a imagem de energia acumulativa, normalizada no tempo, de uma sequência completa de caminhada. O pixel com maior intensidade no GEI indica que o movimento humano ocorreu com maior frequência naquela posição. Uma vez obtida uma série de imagens GEI, é necessário extrair os padrões discriminantes para a fase de treinamento. As técnicas estatísticas PCA (Principal Component Analysis) combinada com o MDA (Multiple Discriminant Analyses), são utilizadas para obter a projeção dos melhores representantes dos dados e otimizar a separabilidade entre os dados. A figura 13 mostra *frames* de silhuetas normalizadas e alinhadas pertencentes à diferentes sequências de marcha e o respectivo GEI obtido para cada sequência.



Figura 13: Exemplo de silhuetas pertencentes a duas sequências distintas de marcha. A imagem mais à direita de cada linha corresponde ao GEI obtido. Fonte: (HAN; BHANU, 2006).

O método possui algumas características muito interessantes, tais como: a representação de uma sequência de marcha humana em uma única imagem em escala de cinza, preservando as informações temporais; o método é menos sensível ao ruído contido nas silhuetas e experimentos realizados com a base USF apresentam bons resultados de classificação e são superiores aos métodos propostos por Sarkar et al. (2005), Collins e Shi (2002), Sundaresan, RoyChowdhury e Chellappa (2003), para esta base.

A abordagem de Sarkar realiza o reconhecimento utilizando-se da correlação temporal das silhuetas (SARKAR et al., 2005). O objetivo é desenvolver uma técnica na qual o desempenho de futuros algoritmos possam ser avaliados. Este algoritmo é composto por quatro módulos. O primeiro identifica o caminhante na cena; o segundo extrai as silhuetas; o terceiro computa o período da marcha das silhuetas - o período da marcha é utilizado para particionar a sequência para efetuar a correlação espaço-temporal - e o quarto módulo

realiza a correlação espaço-tempo para computar a similaridade entre duas sequências da marcha. Para avaliação do método é utilizada a base MOBO (GROSS; SHI, 2001) e as bases NIST / USF e comparadas com outras técnicas selecionadas. Isso permitiu determinar as co-variáveis de maior impacto durante o reconhecimento, que em um dos testes realizados revelaram ser a superfície sobre a qual indivíduos caminharam e o intervalo de tempo entre a gravação de dados. A influência da superfície foi detectada durante os experimentos realizados com grama, que é mais irregular que o concreto, como já demonstraram as pesquisas de Sundaresan, RoyChowdhury e Chellappa (2003).

O método proposto por Yang, Wu e Peng (2006) extrai fatias de diferença do movimento. O método da Diferença Temporal é utilizado para detectar indivíduos que se movimentam na cena. A partir dessa subtração de imagens extrai-se somente o objeto em movimento. A área extraída, cuja dimensão e valores de cinza são normalizados é chamada de Fatia de Movimento da Diferença (*Difference Motion Slice*). Esta fatia contém informações tanto da estrutura do corpo como da dinâmica do movimento. As diferentes fatias de cada indivíduo são então somadas, obtendo uma imagem denominada Imagem da Diferença da Marcha. Quando o conjunto de imagens é grande, as imagens da Diferença da Marcha do mesmo indivíduo terão padrões semelhantes, e padrões diferentes entre as diferentes pessoas do conjunto. A taxa de classificação de acertos (CCR) é calculada utilizando-se a regra *leave-one-out*. Os experimentos são realizados para base NLPR (CASIA, 2005), para quatro diferentes ângulos: 0° , 45° , 135° e 180° . Entre os melhores resultados obtidos, pode-se citar o ângulo de 0° com 98,33% de acertos e o ângulo de 180° , com 100% de acertos utilizando-se o classificador do vizinho mais próximo (NN).

3.4 Técnicas Híbridas

Um método para reconhecimento automático da marcha baseado na análise espaço-temporal da silhueta (WANG et al., 2003) utiliza Análise de Componentes Principais (PCA)

para capturar as características estruturais e de transição (dinâmica) da marcha. As silhuetas de duas dimensões (2D) são mapeadas em um sinal de distância de uma dimensão (1D). As alterações no formato das silhuetas em relação ao tempo são então transformadas em sequências de sinais de distância unidimensional (1D) para aproximar as mudanças temporais do padrão da marcha. Estes sinais de distância consistem da diferença entre as localizações do pixel central da silhueta e os pixels das bordas superiores e inferiores da mesma. A técnica PCA é aplicada nestes sinais, para extrair os componentes principais da assinatura da marcha (fase de treinamento) ou determinar a identidade do indivíduo por meio de técnicas não paramétricas de classificação de padrões no auto-espaço de baixa dimensão (fase de classificação).

Este método possui as seguintes propriedades positivas: visa capturar as características estruturais (aparência) e dinâmicas contidas no movimento humano, é pouco sensível às alterações de cores e texturas da roupa, padrões relacionados ao ritmo da marcha melhoram o índice de classificações corretas e as bases de sequências de marcha utilizadas nos experimentos são amplamente utilizadas na literatura. No entanto, possui as seguintes características não desejáveis: o método é sensível ao ângulo da câmera (é eficiente para vista lateral) e não é eficiente para trajetórias de caminhada não lineares.

O desempenho do método é avaliado com a base USF e NLPR (CASIA, 2005). Para a base USF, seu desempenho é relativamente inferior ao algoritmo base de Sarkar et al. (2005). Como fatores determinantes para a redução de desempenho, pode-se citar o ruído presente na segmentação da imagem e o fato da trajetória de caminhada ser na forma de uma elipse. Com relação à base NLPR, para o ângulo 0° , o método é comparado com os trabalhos de BenAbdelkaer, Cutler e Davis (2002), Collins e Shi (2002), Lee e Grimson (2002), Phillips et al. (2002). Seu desempenho supera os métodos de BenAbdelkaer, Cutler e Davis (2002), Collins e Shi (2002), Phillips et al. (2002). O custo computacional destes algoritmos também é calculado, sendo o de Wang et al. (2003) o melhor de todos. A tabela 8 mostra a taxa de classificações corretas para os ranks: 1, 5 e 10, utilizando-se o protocolo FERET (PHILIPS et al., 2000).

Tabela 8: Comparativo de desempenho do método proposto por Wang et al. (2003) para o Rank 1 - base NLPR, ângulo 0°

Método	TOP 1	Custo Computacional (min/seg)
BenAbelKader (BENABDELKAER; CUTLER; DAVIS, 2002)	72,50	Médio(8,446)
Collins (COLLINS; SHI, 2002)	71,25	Alto(17,807)
Lee (LEE; GRIMSON, 2002)	87,50	Baixo(2,367)
Phillips (PHILLIPS et al., 2002)	78,75	Alto(200)
Wang (WANG et al., 2003)	82,50	Baixo(2,054)

O método foi re-implementado por Lam e Lee (2005) e comparado com seu método, como pode ser observado na tabela 7.

Inspirando-se no MHI (BOBICK; DAVIS, 2001) e no GEI (HAN; BHANU, 2006), Liu e Zheng (2007) propõe um método híbrido para representação e reconhecimento da marcha denominado: Imagem do Histórico da Marcha (GHI). O GHI modela as características estáticas e dinâmicas da marcha, assim como suas variações espaciais e temporais. Sendo que o MEI e o MHI registram apenas os aspectos dinâmicos do movimento; o GEI registra os aspectos estáticos e dinâmicos e não as variações temporais. Diferentemente do GEI, o tempo de duração do GHI corresponde a 1/4 de um ciclo de marcha. Métodos estatísticos tais como: LDA e distância euclidiana são utilizados para aprender as características distintas da marcha para posterior reconhecimento. O GHI herda os conceitos do MHI de que a informação temporal do movimento adicionadas às informações espaciais podem ser registradas.

Os experimentos são inicialmente realizados com a base NLPR (CASIA, 2005), mas sobre um conjunto menor. São utilizadas apenas duas sequências de caminhada da esquerda para a direita com ângulo de 0°. A taxa de classificação correta (CCR) com LDA e Distância Euclidiana fica acima de 80%, sendo este um bom resultado. No entanto, seria interessante a realização de mais experimentos para uma análise ainda mais eficiente sobre o método.

3.5 Considerações Finais

O objetivo principal deste capítulo foi fornecer uma análise geral do estado da arte para identificação e reconhecimento de pessoas por meio da marcha.

Um grande número de pesquisas realizadas comprovaram ser possível a identificação e reconhecimento humano por meio da marcha. Em comparação com técnicas biométricas tradicionais, tais como: impressão digital e face, as pesquisas sobre a marcha estão apenas começando. É uma área interessante e com muitos desafios a serem superados.

Grande parte das pesquisas baseia-se na análise da silhueta, considerando em sua grande maioria aspectos estáticos das mesmas. No entanto, pode-se comprovar, pelas pesquisas realizadas, que os aspectos estáticos não são suficientes para uma boa identificação do indivíduo. O modelo espaço-temporal é o que mais se aproxima de uma classificação e reconhecimento mais confiável, pois trata os aspectos dinâmicos da marcha. Metodologias híbridas, ainda com poucos trabalhos na área, consideram ambos aspectos pertencentes à marcha: dinâmicos e estáticos.

Diferentemente dos trabalhos encontrados na literatura, o método proposto neste trabalho, GBM, considera diferentes informações globais do corpo em movimento. Estas informações são distintas, devido ao fato de serem capturadas por diferentes segmentações da imagem.

A metodologia do trabalho proposto pode ser classificada como uma técnica híbrida.

Capítulo 4

Metodologia adotada para o Reconhecimento da Marcha Humana

4.1 Considerações Iniciais

Neste trabalho, como já mencionado, são propostos quatro novos esquemas para captura das informações do movimento humano:

1. O modelo SGW possui informações globais das variações dos níveis de cinza durante a marcha. As informações globais do movimento 3-D são capturadas por este modelo, no entanto, é bastante sensível a variações na iluminação;
2. O modelo SBW reduz a sensibilidade à variação da iluminação, mas também perde as informações globais 3-D do movimento, retendo apenas informações sobre o movimento da silhueta (2-D). O vestuário também é uma variável que interfere na qualidade das informações necessárias ao reconhecimento neste caso;
3. O modelo SEW é ainda mais imune às variações de iluminação, trazendo informações sobre o comportamento do contorno do corpo em movimento. No entanto, as informações contidas apenas no contorno não são suficientes para um reconhecimento satisfatório;

4. Finalmente, o modelo SSW busca extrair os movimentos das principais articulações envolvidas na marcha, como as relativas ao movimento das pernas e da coluna vertebral.

Cada modelo traz intrinsecamente informações sobre características específicas do movimento global e são por sua vez, mais vulneráveis ou menos vulneráveis às condições de iluminação, vestuário, ângulo de caminhada e outros.

O sistema proposto deve ser capaz de obter as informações globais do movimento humano como um todo, a partir destes quatro modelos. Estes são comparados separadamente e posteriormente são então combinados, formando-se então um novo modelo de representação da marcha, denominado Movimento Global do Corpo (Global Body Motion) - GBM.

4.2 Modelo do Movimento Global do Corpo - GBM

O esquema ilustrado na Figura 14 apresenta o modelo GBM proposto neste trabalho. A partir dos padrões extraídos, são gerados os quatro novos modelos propostos (SGW, SBW, SEW and SSW) do movimento global do corpo humano.

Diferentemente dos métodos encontrados na literatura, o GBM considera cada indivíduo como uma classe distinta e cada *frame* como um objeto desta classe. Mais ainda, todo o objeto segmentado é utilizado para gerar o vetor de característica.

A etapa inicial do método consiste na subtração do fundo, das imagens que compõem as sequências de marcha. Para a subtração do fundo e conseqüente segmentação do movimento, utiliza-se o algoritmo baseado em Mistura de Gaussianas (GMM), proposto originalmente por Stauffer e Grimson (STAUFFER; GRIMSON, 1999) com as modificações de KaewTrakullpong e Bowden (KAWTRAKULLPONG; BOWDEN, 2001).

Com esta segmentação, têm-se dois tipos de imagens:

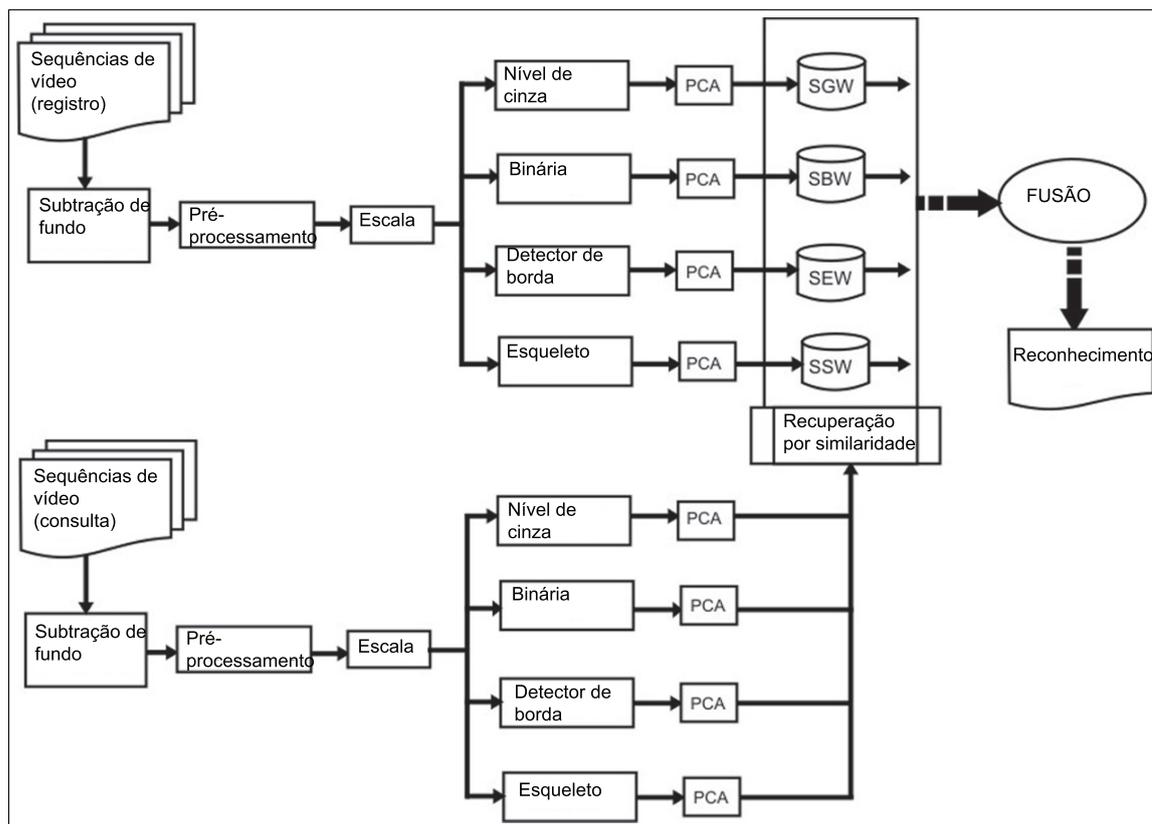


Figura 14: Esquema geral do modelo proposto – Global Body Motion (GBM)

- O primeiro corresponde à imagem em movimento segmentada, em tons de cinza, cuja sequência será denominada de Silhouette-Gray (SG);
- A segunda imagem é obtida a partir da máscara binária gerada pelo GMM, cuja sequência será denominada de Silhouette-Binary (SB).

É ilustrado na Figura 15 o resultado da segmentação realizada com o GMM.

Para remover ruídos e preencher pequenos "buracos" é realizado em cada frame segmentado um pré-processamento. Este pré-processamento consiste inicialmente na aplicação de operadores morfológicos de abertura e fechamento, disponíveis na biblioteca OpenCV (INTEL, 2006). Com isto, é obtida uma imagem com ruído atenuado. Após este processo, os objetos segmentados em cada frame (pessoa caminhando) são centralizados, considerando seu centro de massa.

Para enquadrar os frames, a partir da imagem, é inicialmente encontrado o centro de massa da figura conforme mostrado pela equação 4.1, cuja dimensão da imagem F é de

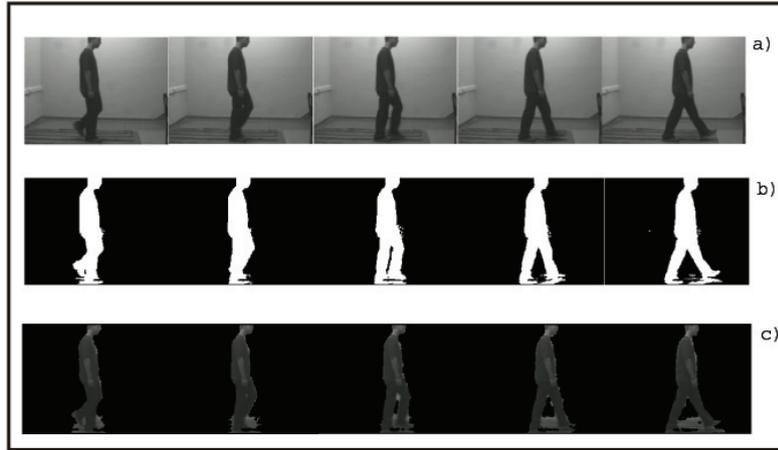


Figura 15: (a) sequência original; (b) sequência Binária (SB) obtida a partir da máscara binária gerada pelo GMM; (c) Subtração do fundo (SG) por meio do GMM em escala de cinza.

$n \times m$ (n linhas e m colunas), as coordenadas (x, y) fornecem o centro de massa e $F(i, j)$ assume valor de intensidade da imagem. Com o centro de massa calculado o janelamento é enquadrado com as dimensões desejadas, tomando como centro do janelamento o centro de massa da imagem, assim as imagens ficam com dimensão de 124×240 pixels.

$$(x, y) = \frac{1}{n \times m} \sum_{i=0}^n \sum_{j=0}^m \{(i, j) \times F(i, j)\} \quad (4.1)$$

4.2.1 Redução de escala - TW de Haar

Para sintetizar dados e fazer compressões, usam-se as funções wavelet ortogonais, que representam os sinais de forma mais completa. Dentre as funções wavelet ortogonais, a mais simples é a wavelet de Haar.

A imagem original é decomposta em quatro sub-bandas com diferentes informações de conteúdo de detalhes. Para cada nível de decomposição, quatro novos componentes são gerados, com metade da resolução espacial e escala de entrada. Cada nível contém

um componente resultante da filtragem passa-baixa e três componentes resultantes da filtragem passa-alta. Com o filtro passa-baixa obtém-se a imagem aproximação de baixa resolução e, a partir dos filtros passa-alta, são gerados os componentes com detalhes verticais, horizontais e diagonais.

A Figura 16(a) mostra a imagem original de dimensão 124×240 pixels. A Figura 16(b) mostra o primeiro nível de decomposição wavelet, decomposta nos quatro componentes: coeficiente de baixa frequência (imagem de aproximação) e coeficientes com os detalhes horizontais, verticais e diagonais, respectivamente. Nesta fase a dimensão da imagem gerada é de 62×120 pixels. A Figura 16(c) mostra o segundo nível de decomposição wavelet, cuja dimensão atual é 31×60 pixels.

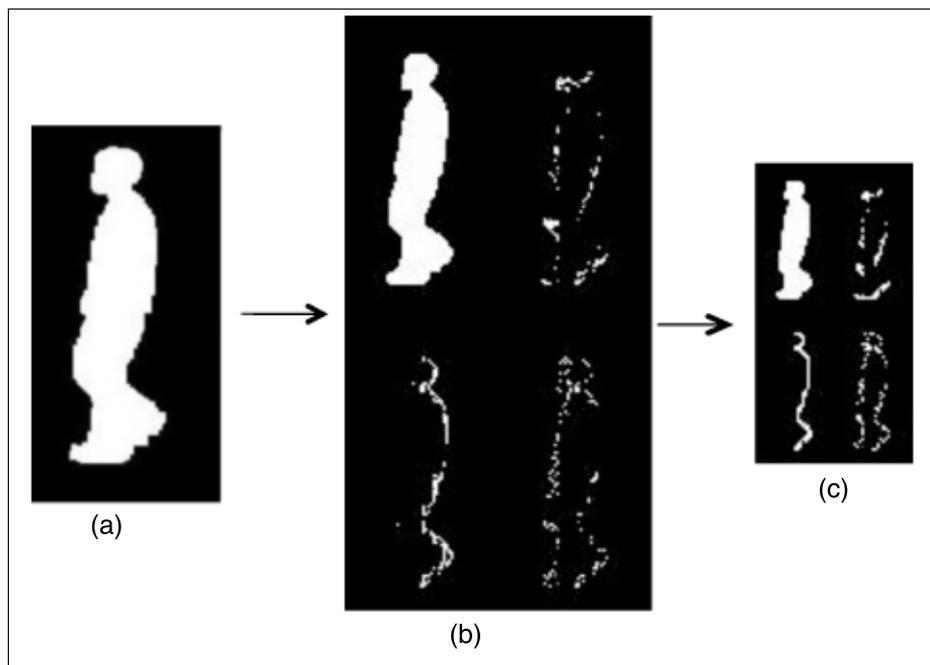


Figura 16: Algoritmo de decomposição da imagem, por meio da TW de Haar.

Considerando-se que a imagem segmentada original contém todas as informações sobre o movimento global do corpo humano caminhando e que estas informações não se alteram significativamente com a escala, aplica-se a TW de Haar em dois níveis a cada uma das seqüências segmentadas, utilizando-se da imagem de aproximação. Desta maneira, a partir da seqüência SG gera-se a seqüência SGW e da seqüência SB gera-se a seqüência SBW.

As seqüências segmentadas, que constituem cada classe de indivíduos, são então for-

madras por imagens de 31×60 pixels com a pessoa centralizada em cada frame.

Esta redução de escala é fundamental para a redução da quantidade de dados sem redução da quantidade de informação global contida no movimento, otimizando assim o esforço computacional do reconhecimento.

Experimentos foram realizados com redução de escala por filtro gaussiano e por filtro da média. As perdas em baixa frequência destes dois filtros reduziram o desempenho do sistema, concluindo ser necessário preservar estas informações. A imagem aproximação de Haar preserva as informações de baixa frequência, sendo que das famílias wavelets, a Haar é a menor complexidade computacional. Sua complexidade é linear - $O(n)$ – sendo que esta é a melhor situação possível para um algoritmo que tem que processar n elementos de entrada ou produzir n elementos de saída.

4.2.2 Movimento do contorno e do esqueleto

Visando capturar as variações globais do movimento do corpo humano contidas apenas no contorno da silhueta, aplica-se o detector de bordas de Canny nas sequências SBW, gerando-se as sequências SEW, formando classes do movimento de contorno, conforme pode ser visto na Figura 17(a).

As sequências SEW são esqueletizadas, utilizando-se operadores matemáticos morfológicos, que estão disponíveis na biblioteca OpenCV (INTEL, 2006). Assim, são geradas as classes de sequências de movimentos globais do esqueleto, denominadas de SSW e mostradas na Figura 17(b). A vantagem destas duas abordagens é reduzir a quantidade da informação redundante sem perder significativamente a qualidade da informação do movimento global do corpo humano em movimento.

Assim sendo, são gerados os quatro novos modelos: SGW, SBW, SEW e SSW. A figura 18 mostra um exemplo de cada modelo gerado.

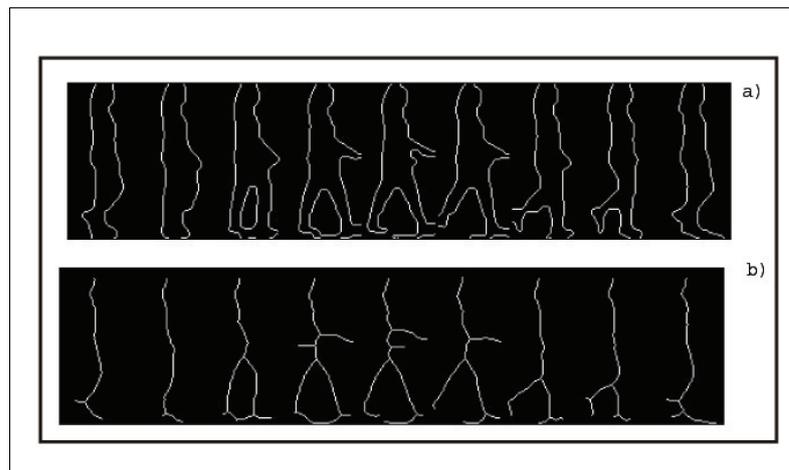


Figura 17: (a) Contornos do corpo (SEW); (b) sequência de esqueletos do movimento (SSW).

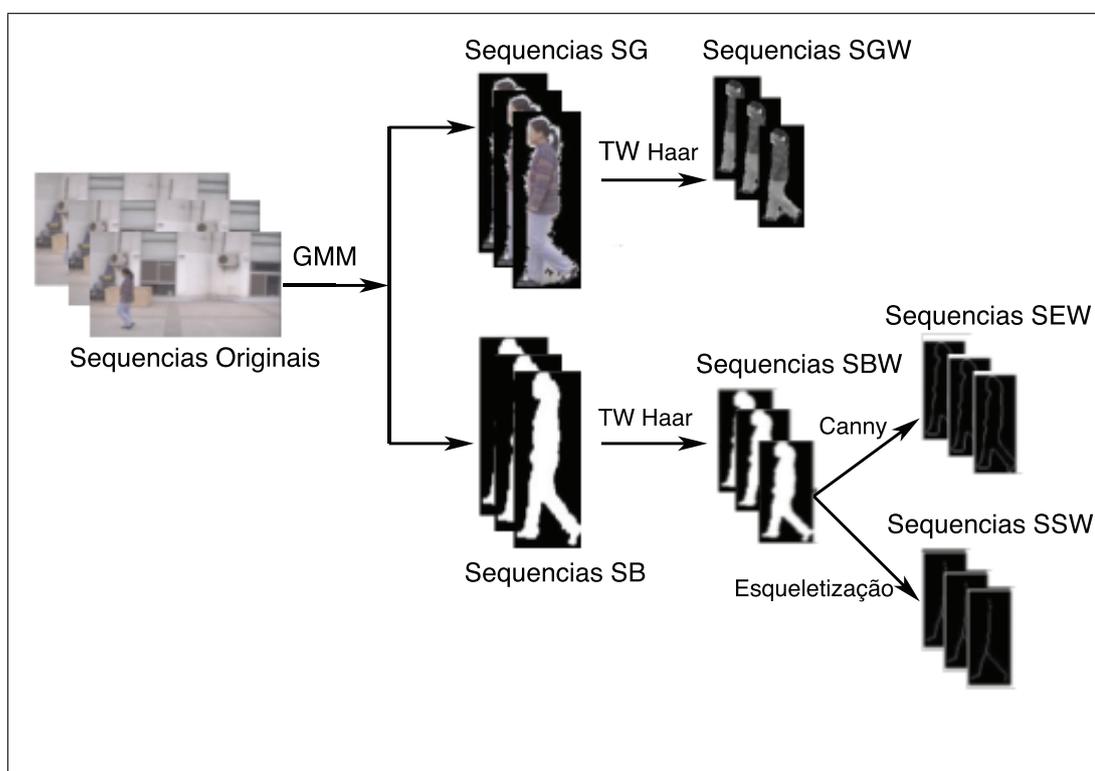


Figura 18: Exemplo de cada modelo gerado. A sequência SGW é obtida a partir da sequência SG; a sequência SBW é obtida a partir da sequência SB; e as sequências SEW e SSW são obtidas a partir da sequência SBW.

4.2.3 Extração de características - EigenGait

Para processar dados correlacionados, é necessário aplicar métodos para decorrelação dos mesmos. Isto é necessário para imagens da marcha humana, porque eles contêm

uma substancial correlação entre amostras consecutivas. Um método que é amplamente utilizado para descorrelacionar dados é a Análise dos Componentes Principais (PCA).

Ao descorrelacionar os dados, a informação redundante é eliminada em cada dimensão. Assim sendo, o propósito do PCA é encontrar uma transformação que seja mais representativa e mais compacta dos dados observados. Normalmente, a grande maioria dos dados pode ser explicada por um número reduzido de componentes, sendo possível descartar outros componentes sem perda relevante dos dados. No entanto, para minimizar a informação perdida durante a aplicação do PCA, os auto-vetores gerados durante o processo são analisados para não comprometer o desempenho do sistema.

Como cada sequência de frames representa uma classe correspondente a um indivíduo caminhando, com a imagem segmentada centralizada em cada frame e em cada uma das quatro sequências (SGW, SBW, SEW e SSW) processadas, a variância intra-classe é pequena e a variância inter-classes é maior. A técnica PCA é aplicada para que se possam extrair as características importantes para o reconhecimento. Com isso, a dimensionalidade dos dados é também reduzida em relação às variáveis originais, preservando-se a informação relevante. As características principais extraídas formam o vetor de características que será utilizado para classificação das silhuetas em suas respectivas classes.

Cada classe, considerando-se formada por uma sequência de frames de um mesmo indivíduo em cada um dos tipos (SGW, SBW, SEW e SSW), é projetada no sub-espço PCA. Para se chegar a uma dimensão ótima, vários testes foram realizados, iniciando-se com dimensão 32 e aumentando-se gradativamente, obtendo-se os melhores resultados de classificação conforme pode ser visto na seção 5.3.1.

O EigenGait é gerado por meio da aplicação do PCA, no conjunto de sequências SGW, SBW, SEW e SSW. O EigenGait pode ser considerado como o conjunto de "padrões de marcha" e são usados para se avaliar a similaridade entre uma "sequência de busca" e os padrões de marcha .

Para se obter um conjunto de EigenGait, são realizados os seguintes passos:

1. Formar os conjuntos de treinamento. As imagens tomadas para o treinamento são aquelas pertencentes aos conjuntos SGW, SBW, SEW e SSW. Cada imagem de uma sequência de vídeo é tratada como um vetor. Para isto, concatena-se as linhas dos pixels da imagem original. Por exemplo, uma matriz M de $n \times m$ ($n = 31$ linhas e $m = 60$ colunas) é tratada como um vetor S de $(nm) \times 1$, ou seja um vetor S de 1860×1 . Todas as imagens do conjunto estão dispostas em uma matriz T de dimensão $n \times m$, sendo que cada coluna representa uma imagem;
2. Calcular a média das imagens do conjunto de treinamento T ;
3. Subtrair a média calculada de cada imagem original em T , obtendo-se assim uma nova matriz T_1 . A partir de T_1 calcula-se a matriz de co-variância B ;
4. Calcular os auto-vetores e os auto-valores da matriz B . Cada auto-vetor tem a mesma dimensionalidade dos dados originais. O auto-vetor com maior auto-valor corresponde a maior porcentagem da variabilidade total presente e assim sucessivamente. Os auto-vetores são ordenados pelos auto-valores, em ordem de significância, do maior para o menor;
5. Escolher os componentes do vetor de características, reduzindo a dimensionalidade dos dados. São escolhidos para formar o vetor, os auto-vetores de maior significância. Uma matriz de padrões T é construída com os auto-vetores escolhidos em sua coluna;
6. Derivar o novo conjunto de dados. Uma vez construída a matriz de padrões, obtém-se a matriz transposta $T1$. Agora, os auto-vetores estão nas linhas, com o auto-vetor mais significativo no topo.
7. Calcular a média das imagens de cada classe da matriz $T1$. A média das imagens ou protótipos de cada classe são denominados Eigengaits.

A similaridade entre as classes e o EigenGait é medida, neste trabalho, com o Classificador do Vizinheiro mais Próximo (NN), usando a métrica Euclidiana.

4.2.4 Fusão de Características

Hong, Jain e Pankanti (HONG; JAIN; PANKANT, 1999) demonstraram que a integração de múltiplas técnicas biométricas resultam em uma melhora significativa no desempenho global de um sistema, com a diminuição das Taxas de Falsa Aceitação e de Falsa Rejeição.

Apesar de se utilizar uma única técnica biométrica, também se utiliza diferentes formas de representação desta e cada representação carrega informações bem distintas do movimento do corpo humano e da própria silhueta, além de serem vulneráveis em diferentes situações (presença de sombra, alteração de iluminação, alteração na vestimenta, pouca informação contida no esqueleto, etc). Sendo assim, espera-se que a fusão agregue as características estáticas da silhueta humana – presentes nos modelos SGW, SBW e SEW - e as características dinâmicas do movimento – presentes no SSWs.

A integração das informações apresentadas pelos diversos indicadores biométricos pode ser realizada de três maneiras possíveis: no módulo de extração de características, no módulo de verificação ou no módulo de decisão. Quando a integração ocorre no nível de extração de características, esta assume uma forte integração dentro das medidas de entrada e tais esquemas são conhecidos como integração fortemente acoplada (CLARK; YUILLE, 1990). A integração fracamente acoplada, assume muito pouco ou nenhuma interação dentro das medidas de entrada, nestes casos, a integração ocorre no módulo de decisão. A figura 19 mostra os três esquemas de fusão.

1. Integração no módulo de extração de características

Se as características extraídas de uma modalidade biométrica são independentes das outras modalidades (por exemplo, a face e a impressão digital) é, então, sugerido por JAIN, ROSS e PRABHAKAR (2004) que os vetores de características sejam concatenados em um único vetor composto por diferentes indicadores biométricos. Portanto, o novo vetor de características terá uma maior dimensionalidade e representará a identificação do indivíduo num espaço de características diferente

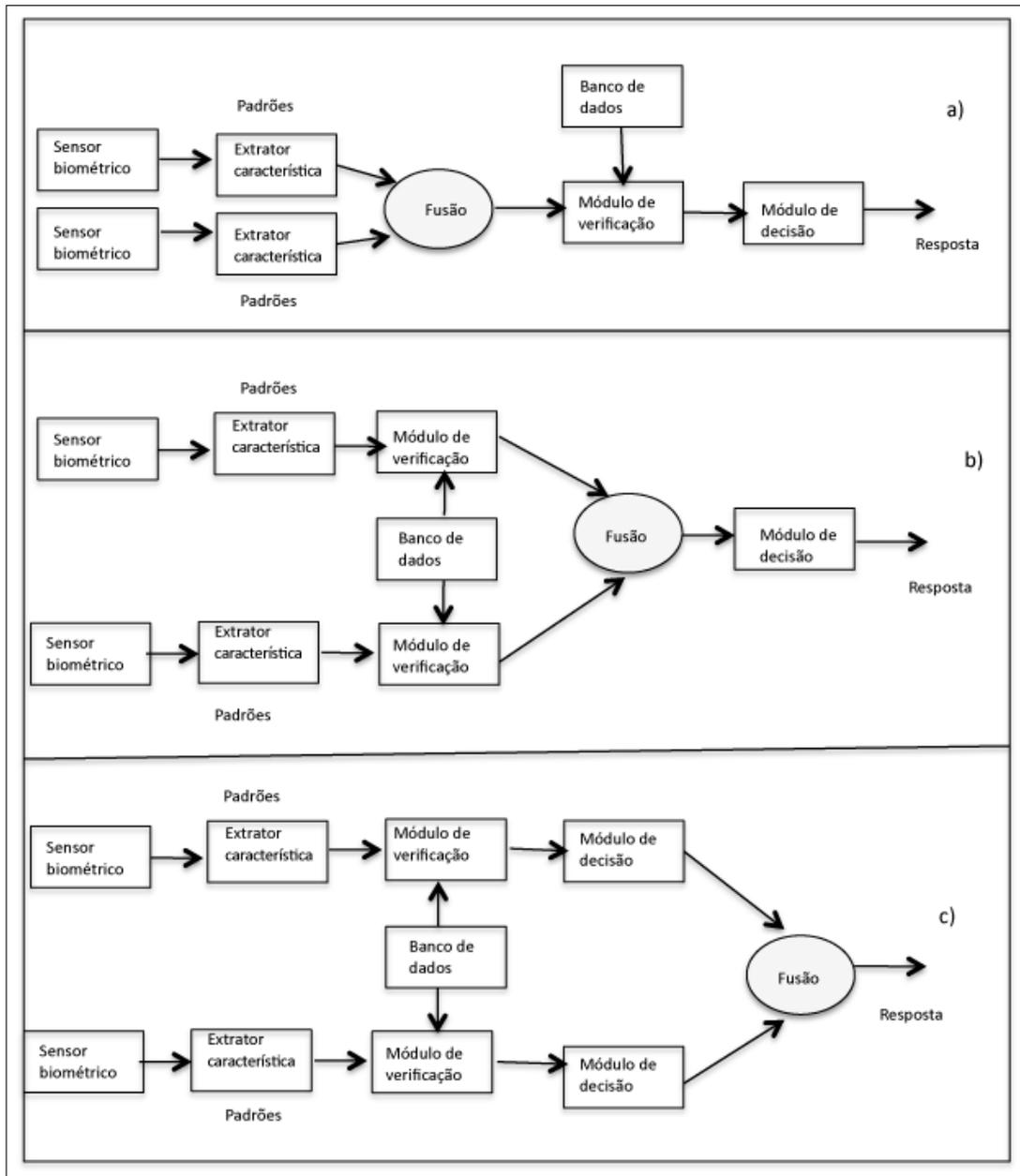


Figura 19: Diferentes níveis de fusão: (a) Fusão no nível de extração de características, (b) Fusão no nível do módulo de combinação e (c) Fusão no nível de decisão. Adaptado de JAIN, ROSS e PRABHAKAR (2004).

e, espera-se, mais discriminante. Neste modo de fusão, as técnicas de redução de características podem ser aplicadas nos respectivos módulos de extração, de cada modalidade biométrica, para que esses módulos possam extrair um número menor de características pertinentes ao invés de um número muito grande, o que ocasionará um grande aumento da dimensionalidade do vetor de características resultante do módulo de fusão.

2. Integração no módulo de verificação

Cada identificador biométrico gera valores de similaridade indicando o grau de igualdade entre o vetor de característica extraído da modalidade biométrica com o vetor de características do modelo armazenado no banco de dados do sistema de reconhecimento biométrico. Estes valores podem ser combinados para aumentar a confiabilidade da identificação solicitada. Segundo JAIN, ROSS e PRABHAKAR (2004), a técnica que efetua essa fusão pode ser a média ponderada dos valores de similaridade de cada modalidade biométrica.

3. Integração no módulo de decisão

Cada modalidade biométrica efetua a própria decisão de reconhecimento baseada em seus próprios vetores de características.

Neste esquema de integração a tomada de decisão do sistema pode ser realizada por meio de regras de inferência do tipo *AND* e *OR*. Na regra *OR*, se a saída de um dos módulos de decisão do sistema rejeitar um usuário, é dada uma nova chance ao usuário verificar sua identidade em outro módulo de decisão do sistema.

Para o caso de se utilizar a regra *AND*, para um usuário ser aceito, sua identificação deve ser positivo, em todos os módulos de saída.

Zuev e Ivanov (1999) propuseram um esquema de votação da maioria para realizar o processo de fusão da decisão de reconhecimento.

A presente proposta, apesar de utilizar apenas uma técnica biométrica, propõe diferentes representações de modelos da marcha humana com informações distintas sobre o movimento global humano. No processo de fusão, as características presentes nos modelos SGW, SBW, SEW e SSW, são agregadas para otimizar o desempenho do sistema.

A abordagem de fusão proposta nesta tese considera que a saída de cada modelo (SGW, SBW, SEW e SSW), treinado individualmente, fornece uma pontuação de similaridade entre cada frame e as respectivas classes em que serão classificados.

A similaridade entre cada imagem projetada no sub-espço PCA e o EigenGait de cada

classe é obtido por meio do classificador do vizinho mais próximo (NN), utilizando-se a métrica Euclidiana. A menor distância obtida é aquela em que o *frame* será classificado. Assim, é obtido o percentual de acertos para cada modelo individualmente.

O modelo de fusão proposto utiliza a média ponderada das pontuações de similaridade. O modelo de representação da marcha que individualmente obteve o melhor desempenho, tem maior peso no processo de decisão da classificação.

Segue abaixo a descrição do algoritmo:

1. Calcular a medida de similaridade entre o j^{th} frame e o Eigengait de todas as classes c do modelo i ($S_{j,c,i}$) fornecido pela equação 4.2:

$$S_{(j,c,i)} = MIN(||frame_j - Eigengait_c||) \quad (4.2)$$

Sendo $S_{(j,c,i)}$ a menor distância euclidiana entre o frame j e o Eigengait de cada classe c , utilizando o modelo i . Neste caso, temos i variando de 1 até 4.

2. Calcular a precisão média de acertos (ρ_i) para cada modelo i , fornecido pela equação 4.3

$$\rho_i = \left(\frac{TP_i}{TG_i} \right) \quad (4.3)$$

Sendo TP_i o total de verdadeiros positivos do modelo i ; TG_i é o total de amostras do modelo i .

3. Calcular a medida de similaridade entre o j^{th} frame e a classe c , fornecida pela equação 4.4:

$$\phi_i = \frac{\sum_{i=1}^4 \rho_i S_{(j,c,i)}}{\sum_{i=1}^4 \rho_i} \quad (4.4)$$

A equação 4.4 é denominada **fusão**, pois representa a média ponderada entre as pontuações de similaridade de todos os modelos. Sendo utilizado como peso a média de acertos obtida no item 2.

4.2.5 Materiais

Para a implementação dos algoritmos propostos neste trabalho, foram desenvolvidos programas na linguagem C, utilizando-se a biblioteca OpenCV (Open Source Computer Vision Library) (INTEL, 2006). São utilizadas três bases independentes de imagens para validar a metodologia proposta neste trabalho.

- Base A

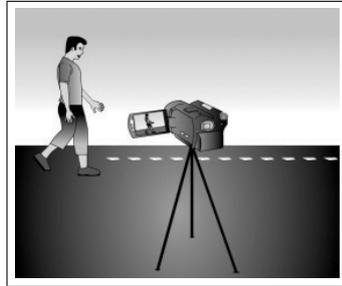


Figura 20: Esquema de aquisição da base de imagens própria. A direção de caminhada é perpendicular ao eixo óptico da câmera fixa.

É uma base de vídeo própria, gravados no formato AVI, de cenas de pessoas caminhando em ambiente interno, com iluminação controlada. A câmera utilizada possui resolução de 320×240 pixels. A digitalização das imagens capturadas pela câmera foi feita por meio de uma placa digitalizadora conectada a um micro-computador do tipo PC a uma taxa de 15 FPS. Para formar esta base, foram digitalizadas imagens de 10 pessoas, ou seja, 10 classes. Cada classe é composta por 3 sequências de vídeo, obtidas em um mesmo dia. Essas sequências foram concatenadas, formando uma única sequência, normalizada em 100 frames. Em cada vídeo, as pessoas caminham em uma única direção, numa superfície plana e perpendicular ao eixo óptico da câmera fixa. Apenas uma pessoa está presente por vídeo. A Figura 20 mostra o esquema utilizado para captura das imagens.

- Base B

Essa base é formada por imagens do “*Gait Database*” do *National Laboratory of*



Figura 21: Amostra de quadros com variação de ângulos (0° , 90° e 45°) da Base B. Fonte: (CASIA, 2005).

Pattern recognition - NLPR (CASIA, 2005), pertencente ao *Institute of Automation, Chinese Academy of Science* disponibilizadas em formato png. Estas imagens foram geradas em ambiente externo, com luz natural. As imagens incluem três vistas: lateral (0°), oblíqua (45°) e frontal (90°). Cada classe possui três vistas, com quatro sequências por ângulo.

As sequências são assim denominadas:

- Sequência 1: direção da caminhada é da direita para esquerda e, na vista frontal, de frente para a câmera;
- Sequência 2: direção da caminhada é esquerda para a direita e, na vista frontal, de costas para a câmera;
- Sequência 3: direção da caminhada é da direita para esquerda e, na vista frontal, de frente para a câmera;
- Sequência 4: direção da caminhada é da esquerda para a direita e, na vista frontal, de costas para a câmera.

São imagens de 24 bits coloridas e capturadas a uma taxa de 25 FPS com resolução de 352×240 pixels. O comprimento de cada sequência varia de acordo com a velocidade do passo do indivíduo, sendo a média de 90 *frames* por sequência. É ilustrado na Figura 21 um quadro de cada variação de ângulo em cada uma das quatro sequências. Para que a aplicação de nossa abordagem pudesse ser realizada, foram montadas as sequências de vídeos a partir das imagens disponibilizadas. Foram obtidas ao todo 20 classes, com 240 sequências de vídeo (20 classes, 3 vistas, 4 sequências por vista) e 8640 *frames*, sendo que cada sequência contém 36 *frames*.

- Base C

Essa base é formada por dez classes, também disponibilizada pelo NLPR no formato AVI. Cada classe ou indivíduo possui três tipos de sequências denominadas: bg (carregando bolsa), cl (vestindo casaco) e nm (sem bolsa e casaco).

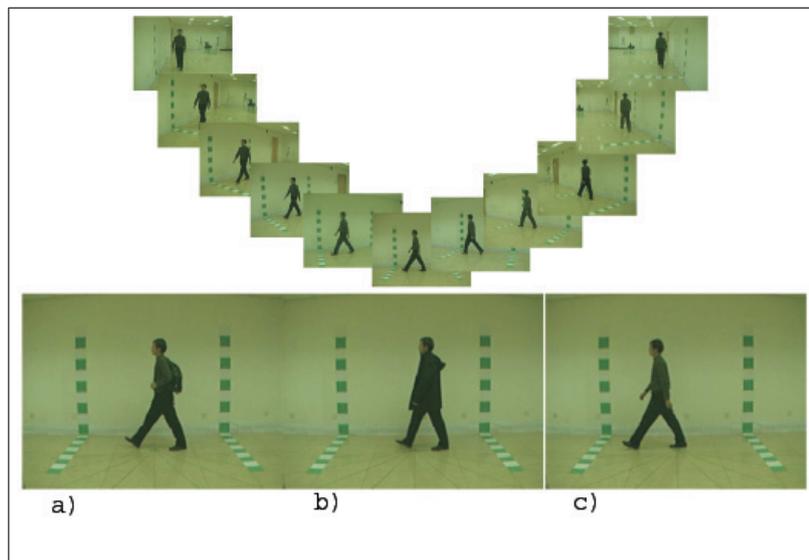


Figura 22: Amostra de quadros com variação de ângulos da base C. a) Sequência bg; b) Sequência cl; c) Sequência nm – Os frames dispostos no topo representam os 11 diferentes ângulos (0° a 180°), para a sequência carregando bolsa. Fonte: (CASIA, 2005).

- bg: sequência com o indivíduo caminhando em uma superfície plana carregando uma bolsa no ombro ou na mão;
- cl: sequência com o indivíduo caminhando, sem bolsa e com um casaco sobre a roupa e;
- nm: sequência com o indivíduo caminhando em uma superfície plana sem bolsa e com a mesma roupa da sequência bg.

É ilustrado na Figura 22 um exemplar de cada tipo de sequência (com bolsa, com casaco e sem bolsa e casaco) para o ângulo de 90° . A sequência de imagens situadas no topo da figura representam a variação de ângulo (de 0° até 180°) de uma sequência carregando bolsa.

Os vídeos foram capturados por 11 câmeras em diferentes direções. O ângulo Θ entre a direção da câmera e o indivíduo tem os valores de 0° até 180° , variando de 18° em 18° . Cada pessoa caminhou 10 vezes ao longo de uma linha demarcada (seis vezes caminhando normalmente, duas vezes caminhando com casaco e duas vezes caminhando sem casaco). Foram registradas 110 sequências, ou seja, 1100 sequências de vídeo ao todo.

4.3 Consideração Finais

O objetivo deste capítulo foi descrever o método GBM e, as técnicas envolvidas para o desenvolvimento deste.

No GBM é proposta a representação da marcha humana em quatro modelos, sendo o módulo final do GBM a fusão destes quatro modelos.

Capítulo 5

Resultados

5.1 Considerações Iniciais

Neste trabalho é proposta a geração de quatro modelos distintos de representação da marcha humana: SGW, SBW, SEW e SSW e a fusão das características destes modelos para reconhecimento biométrico. São utilizadas três bases independentes de imagens de pessoas caminhando para validar a metodologia proposta: Base A, Base B e Base C.

O objetivo deste capítulo é descrever os resultados dos experimentos realizados para cada modelo proposto (SGW, SBW, SEW e SSW) em cada base de dados. Os resultados são avaliados individualmente para cada modelo e posteriormente avaliados com o processo de fusão.

5.2 Métodos de Avaliação dos Resultados

O método proposto é avaliado com: matrizes de confusão, curvas *recall* \times *precision* e curvas CMS.

1. Matrizes de Confusão

Para as três bases de dados utilizadas, cada imagem de cada frame é projetada no sub-espço PCA e comparada com o EingenGait protótipo de cada classe. Matrizes de confusão mostram o número de acertos em cada classe para as bases A e B. Para cada experimento são geradas Matrizes de Confusão e a partir das mesmas são calculadas a Falsa Aceitação (FA) - o elemento de entrada é impostor (Negativo) e o classificador o classifica como Positivo e; Falsa Rejeição (FR) - o elemento de entrada é genuíno (Positivo) e o classificador o analisa como Negativo dentro de cada classe. Os valores de FA e FR, anotados nas Matrizes de Confusão são obtidos com as Equações 5.1 e 5.2.

$$FA = \frac{FP}{TI} \quad (5.1)$$

$$FR = \frac{FN}{TG} \quad (5.2)$$

Sendo TP o número de Verdadeiros Positivos - o elemento de entrada é Genuíno (Positivo) e o classificador o analisa como Positivo; TI , o número de frames não verdadeiros; FN , o número de frames erroneamente Rejeitados (em cada linha da Matriz) e TG o total de elementos de uma classe.

A partir dos valores de FA e FR, é calculada a acurácia do classificador. A acurácia (exatidão) é a proporção de classificações corretas para o total de elementos classificados (genuínos e impostores). A acurácia é obtida pela equação 5.3.

$$A = \frac{TP + TN}{TG + TI} \quad (5.3)$$

Sendo TN o Verdadeiro Negativo - o elemento de entrada é Impostor (Negativo) e o classificador o classifica como Negativo.

Para a Base C, não são utilizadas as Matrizes de Confusão, devido ao número elevado de experimentos que se realiza com esta base, pois existe uma grande variação de

vestimenta e ângulo.

2. Curvas $Recall \times Precision$ - Avaliação do classificador baseado na Relevância

Como a abordagem proposta é a mesma empregada em recuperação da informação em bases de dados, são geradas curvas $Recall \times Precision$ para avaliar a qualidade dos resultados das consultas por similaridade. Recall é a medida do número de objetos retornados e Precision, a medida de quanto dos objetos retornados está correto. Para gerar a curva $Recall \times Precision$, cada imagem projetada é um objeto de consulta. Desta forma, cada imagem projetada é comparada com ela mesma e com todas as demais. Os valores de $Recall \times Precision$ são calculados de acordo com as Equações 5.4 e 5.5

$$Recall = \frac{|RA|}{R} \quad (5.4)$$

$$Precision = \frac{|RA|}{A} \quad (5.5)$$

Sendo $|RA|$ o número de elementos relevantes retornados pela pesquisa; R corresponde ao número de elementos no conjunto em que está localizado o objeto de consulta e; A corresponde ao número de elementos retornados da pesquisa.

3. Curvas CMS - Cumulative Match Score

O protocolo FERET (PHILIPS et al., 2000), com a validação cruzada *leave-one-out*, é também utilizado para avaliar os resultados obtidos. Este protocolo foi desenvolvido inicialmente para avaliar os algoritmos de reconhecimento de face, mas pode ser usado para outros modelos de sistemas biométricos. O desempenho estatístico deste método é relatado como CMS (Pontuação Cumulativa de Acertos), que é definido como a probabilidade acumulativa de uma classe verdadeira de um objeto de teste pertencer ao topo de k acertos. O CMS relata para que fração de exemplares de testes as respostas corretas estão no topo de k buscas. Isto permite avaliar quantas imagens devem ser examinadas para se obter um nível desejável de desempenho.

Após computar a diferença de similaridade entre o exemplar de teste e o conjunto de treinamento, o classificador do vizinho mais próximo (NN) é aplicado para a classificação.

5.2.1 Base A

Para a avaliação do modelo proposto (GBM) operando sobre a base A, são realizados dois testes independentes, para cada tipo de sequência (SGW, SBW, SEW e SSW). Para o primeiro teste são separados 10 frames de cada classe, perfazendo um total de 100 frames. No segundo teste, utilizam-se todos os frames das classes.

Experimentos realizados:

- Teste-1: 10 classes, 10 frames de cada classe, sub-espço PCA igual a 64;
- Teste-2: 10 classes, 100 frames de cada classe, sub-espço PCA igual a 256.

5.2.2 Base B

Para se avaliar o desempenho do modelo GBM aplicado na Base B, foram realizados vinte e um experimentos independentes para cada tipo de sequência: SBW, SGW, SEW e SSW. Em cada uma destas sequências, a direção do movimento do caminhante é realizada em ângulos de 0° , 45° e 90° . Para cada direção do movimento, esta base possui quatro sequências de frames (1, 2, 3 e 4) e 20 classes, ou melhor, 20 diferentes indivíduos. Os vinte e um testes podem ser divididos em duas partes.

1. Parte-I: Sequências individuais:

- Ângulo 0° , sequência 1;

- Ângulo 0, sequência 2;
- Ângulo 0, sequência 3;
- Ângulo 0, sequência 4;
- Ângulo 45, sequência 1;
- Ângulo 45, sequência 2;
- Ângulo 45, sequência 3;
- Ângulo 45, sequência 4;
- Ângulo 90, sequência 1;
- Ângulo 90, sequência 2;
- Ângulo 90, sequência 3;
- Ângulo 90, sequência 4.

2. Parte-II: Combinações de sequências:

- Ângulo 0, sequências 1 e 3;
- Ângulo 0, sequências 2 e 4;
- Ângulo 0, sequências 1, 2, 3 e 4;
- Ângulo 45, sequências 1 e 3;
- Ângulo 45, sequências 2 e 4;
- Ângulo 45, sequências 1, 2, 3 e 4;
- Ângulo 90, sequências 1 e 3;
- Ângulo 90, sequências 1, 2, 3 e 4;
- Ângulo 90, sequências 2 e 4.

São utilizados para cada um dos testes da Parte-I 36 frames de cada classe, com 20 classes, totalizando 720 frames.

Para os experimentos da Parte-II, o número de elementos de cada classe é a soma dos frames das sequências individuais, assim, para a combinação de duas sequências (1 – 3,

2 – 4), são utilizados 72 elementos de classe. Para a combinação de quatro sequências são utilizados 144 frames em cada classe.

5.2.3 Base C

Para avaliação do modelo GBM, com esta base de dados, são realizados cinco experimentos independentes para cada tipo de sequência gerada: SGW, SBW, SEW e SSW. Os resultados obtidos independentemente em cada modelo não passaram pelo processo de fusão.

Esta base é formada por 10 classes, cada classe com 3 sequências (com bolsa, com casaco e sem bolsa e casaco) e cada sequência é gerada nas direções variando de 0° a 180° em ângulos de 18° .

O primeiro experimento foi realizado com todos os ângulos em cada sequência: com bolsa, com casaco e sem bolsa e casaco. São utilizados 35 frames de cada sequência, perfazendo um total de 350 frames.

Para o segundo experimento são utilizadas duas sequências com bolsa e com casaco, para cada ângulo de 0° a 180° . Nesse experimento são considerados 70 frames de cada classe, totalizando 700 frames.

No terceiro experimento são consideradas 3 sequências sem bolsa e casaco, para todos os ângulos. Tem-se assim 105 frames em cada classe, totalizando 1050 frames.

No quarto experimento são consideradas 6 sequências sem bolsa e casaco, para todos os ângulos. Perfazendo 210 frames de cada classe, totalizando 2100 frames.

E finalmente, para o quinto experimento, consideram-se as combinações de sequências com bolsa com sequências com casaco; sequências com bolsa com sequências sem bolsa e casaco e sequências com casaco com sequências sem bolsa e casaco. Para cada teste deste experimento são utilizados duas sequências de cada tipo. Assim sendo, tem-se 70 frames

de cada classe, totalizando 700 frames.

5.3 Análise dos Resultados

5.3.1 Base A

Para minimizar a quantidade de informação perdida durante a aplicação do PCA, os auto-vetores gerados durante o processo são analisados, a fim de não comprometer o desempenho do sistema.

Assim sendo, experimentos são realizados com a dimensão do subespaço PCA variando de 32 à 256, a precisão média de cada modelo pode ser observada na Figura 23.

Com o aumento da dimensão do subespaço PCA, observa-se uma melhora no número de acertos principalmente para as imagens da sequência SSW e SEW. Para o Teste-1, há um aumento na porcentagem de classificações corretas quando aumenta-se de 32 para 64 auto-vetores. Continuando o incremento de auto-vetores, não há variação no índice de classificações corretas. Dessa forma, 64 auto-vetores são suficientes para representar as características principais. Para o Teste-2, os melhores resultados são obtidos quando o sub-espaço PCA é igual a 256, conforme ilustrado na Figura 23(b)

Quando o número de auto-vetores é 256, há um aumento significativo no índice de classificações corretas para as sequências SSW e SEW, no entanto, não há alteração para as sequências SGW e SBW. Assim sendo, para o Teste-1 utiliza-se a dimensão do sub-espaço PCA igual a 64 e para o Teste- 2 a dimensão do sub-espaço PCA igual a 256. Valores acima de 256 para o sub-espaço PCA aumentam o custo computacional e não aumentam a Precisão Média significativamente. Nas subseções 5.3.1.1 e 5.3.1.2, a seguir, estão dispostos os resultados obtidos em cada teste realizado com a base A.

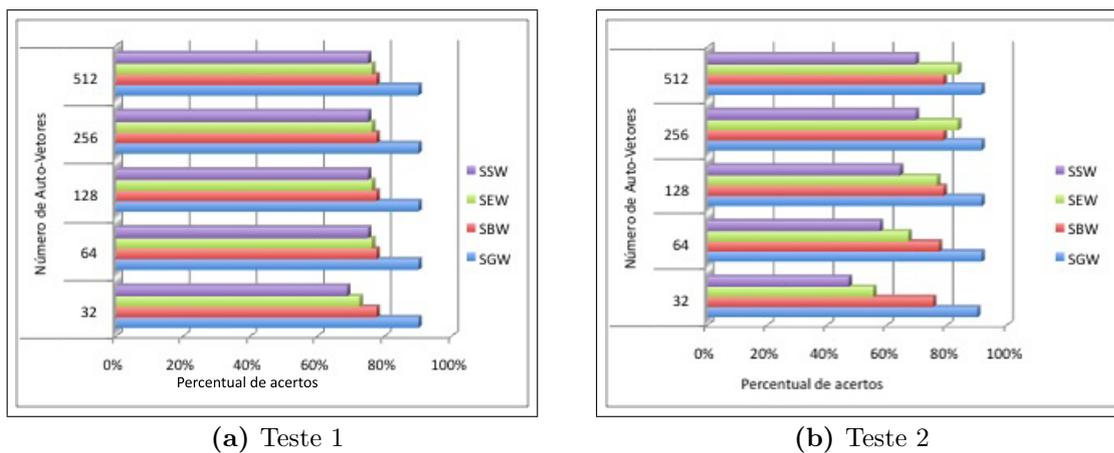


Figura 23: Gráfico comparativo das precisões médias de cada modelo, com a variação do sub-espaco PCA para os dois experimentos realizados

5.3.1.1 TESTE 1

As Matrizes de Confusão, ilustradas na Tabela 9 mostram o resultado da classificação para o Teste-1 realizados sobre as sequências: SGW, SBW, SEW e SSW.

Tabela 9: Matrizes de Confusão do Teste-1 para a Base A.

Classe	1	2	3	4	5	6	7	8	9	10	FR %
1	8	0	0	0	1	0	1	0	0	0	20
2	0	10	0	0	0	0	0	0	0	0	0
3	0	0	10	0	0	0	0	0	0	0	0
4	0	1	0	9	0	0	0	0	0	0	10
5	0	0	0	0	10	0	0	0	0	0	0
6	0	0	1	0	0	9	0	0	0	0	10
7	0	0	0	3	0	1	6	0	0	0	40
8	0	0	0	0	0	0	0	10	0	0	0
9	0	0	0	1	0	0	0	0	9	0	10
10	0	0	1	0	0	0	0	0	0	9	10
FA%	0	1,1	2,2	4,4	1,1	1,1	1,1	0	0	0	90

SGW											
Classe	1	2	3	4	5	6	7	8	9	10	FR %
1	7	0	3	0	0	0	0	0	0	0	30
2	2	8	0	0	0	0	0	0	0	0	20
3	2	0	8	0	0	0	0	0	0	0	20
4	3	0	0	7	0	0	0	0	0	0	30
5	1	0	0	0	9	0	0	0	0	0	10
6	0	0	2	0	0	8	0	0	0	0	20
7	0	1	0	2	0	0	7	0	0	0	30
8	0	0	1	0	0	1	1	7	0	0	30
9	2	0	0	0	1	0	0	0	7	0	30
10	0	0	0	0	0	0	2	0	0	8	20
FA%	11,1	1,1	6,6	2,2	1,1	1,1	3,3	0	0	0	76

SBW											
Classe	1	2	3	4	5	6	7	8	9	10	FR %
1	7	0	2	0	0	0	0	0	0	0	30
2	2	8	0	0	0	0	0	0	0	0	20
3	0	0	7	1	1	1	0	0	0	0	30
4	0	1	0	7	0	1	1	0	0	0	30
5	0	1	0	0	9	0	0	0	0	0	10
6	0	0	1	1	0	8	0	0	0	0	20
7	0	2	0	1	0	0	7	0	0	0	30
8	0	0	0	1	1	1	1	7	0	0	30
9	0	0	0	0	0	0	0	0	8	0	20
10	0	0	1	0	0	2	0	0	0	7	30
FA%	2,2	4,4	4,4	4,4	2,2	7,7	1,1	1,1	0	0	75

SEW											
Classe	1	2	3	4	5	6	7	8	9	10	FR %
1	9	0	0	1	0	0	0	0	0	0	10
2	0	10	0	0	0	0	0	0	0	0	0
3	0	0	10	0	0	0	0	0	0	0	0
4	1	0	0	9	0	0	0	0	0	0	10
5	0	0	0	0	10	0	0	0	0	0	0
6	0	2	0	0	0	8	0	0	0	0	20
7	0	0	0	1	0	0	9	0	0	0	10
8	0	0	0	1	0	1	0	9	0	0	10
9	0	0	0	0	0	0	0	0	10	0	0
10	0	0	0	0	0	0	0	0	0	10	0
FA%	1,1	2,2	0	3,3	0	0	0	0	0	0	94

Fusão											
-------	--	--	--	--	--	--	--	--	--	--	--

Cada frame de busca é comparado com a média, ou vetor protótipo de cada classe. O resultado visto no canto inferior direito de cada matriz é a porcentagem de classificações corretas de cada método ou Precisão Média (razão entre a soma dos elementos da diagonal principal e a soma dos elementos da matriz).

Avaliando-se os resultados pode-se afirmar que mesmo quando o número de elementos de cada classe é pequeno (10 frames/classe), é possível extrair características que consigam distinguir os objetos entre as classes. Isto pode ser verificado por meio dos resultados obtidos para a sequência SGW que mostra um índice de classificação correta acima de 90% para as classes 2 (10 acertos em 10 frames - 100%), 3(100%), 4 (9 acertos em 10 frames - 90%), 5(100%), 6(90%), 8(100%), 9(90%) e 10(90%) e uma taxa de Falsa Aceitação (FA) igual a 0% nas classes 1, 6, 8, 9 e 10. Apenas duas classes têm índice de classificação correta inferior a 90%: a classe 1 e a classe 7.

A classe 1 tem uma porcentagem de classificação correta de 80% (8 acertos em 10 frames) com uma Falsa Rejeição (FR) de 20%, no entanto, a sua Falsa Aceitação é igual a zero, ou seja, nenhum elemento de outra classe foi confundido com a classe 1. Para os modelos SGW e SBW, o desempenho mais baixo é em relação à classe 7, que obteve apenas 60% das classificações corretas com uma alta taxa de Falsa Rejeição (FR=40%) no SGW e 40% de acertos no SBW. No entanto, esta classe tem uma taxa de Falsa Aceitação baixa (FA = 1,1% no SGW e 0% no SBW), ou seja, apenas 1 frame da classe 1 é confundido, ou falsamente aceito como um elemento da classe 7. Analisando-se mais detalhadamente as imagens desta classe, pode-se observar que o indivíduo não completa pelo menos dois ciclos de marcha. Dessa forma, é difícil estabelecer um padrão que a caracterize de forma única. Há também o fato do indivíduo pertencente à classe 7 possuir disfunção locomotora, fato este que aumenta a variância dos dados, visto que os ciclos de sua marcha não são uniformes.

No entanto, o indivíduo 7 é melhor reconhecido com as sequências SEW e SSW (ambas com 70% de acertos (7 acertos em 10 frames)), o que leva a crer que as bordas e o esqueleto possuem maior informação global para identificação da marcha deste indivíduo do que as

imagens em nível de cinza ou binária.

O modelo GBM produz 100% de classificações corretas para as classes 1, 5 e 8 na sequência SBW e 90% de acertos para a classe 5 na sequência SEW. A taxa de classificação correta da sequência SEW é ligeiramente menor em algumas classes devido ao fato do volume de dados ser bastante reduzido por representar apenas o contorno da imagem. A sequência SSW é o modelo que fornece o menor índice de classificações corretas. Considerando que essa sequência é formada pelo esqueleto das silhuetas obtidas, fornecendo um número muito menor de informações, ainda assim é possível extrair características únicas de um indivíduo, capturando informações globais de movimento da estrutura do corpo humano e do movimento das articulações. É interessante ressaltar que a taxa de Falsa Aceitação para as classes 8, 9 e 10 é bastante baixa ($<1,4\%$) para qualquer das sequências avaliadas.

Todavia, como são utilizados os 10 primeiros frames da sequência, este resultado pode não ser o mesmo para os demais frames. Assim sendo, o método de re-amostragem *cross-validation* é utilizado. O conjunto de *frames* é dividido em 10 partições diferentes. A cada iteração deste método, uma partição diferente é utilizada para testar o sistema e todas as outras $k - 1$ partições são utilizadas para treinamento. Assim, são realizadas 10 iterações, a cada iteração uma partição é escolhida para testar e as outras 9 partições restantes são utilizadas para treinar o sistema. A média de acertos e o valor médio das taxas de FA e FR, são resultados da média obtida nas 10 iterações e são mostrado na tabela 10.

É ilustrado na figura 24 o desempenho da classe 7, durante a execução das 10 iterações da validação cruzada. Os números dentro de cada cilindro do gráfico equivalem à quantidade de *frames* corretamente classificados de cada 10 utilizados.

O modelo SSW, para a classe 7, possui a melhor média de acertos para as 10 iterações realizadas. A precisão média de acertos para o modelo SGW é de 68%, enquanto que no modelo SSW é de 71%. Pode-se verificar que para alguns subconjuntos, o modelo SGW supera o SSW, como, por exemplo, para os experimentos 4 e 7.

As classes 8, 9 e 10 são as que possuem menor taxa de Falsa Aceitação para os modelos

Tabela 10: Porcentagens de Acertos com seus respectivos FR e FA médios, para as 10 amostragens de diferentes frames.

Classe	Média Acertos	FA %	FR %	Classe	Média Acertos	FA %	FR %
1	78	2,3	22	1	68	3,3	32
2	100	1,0	0	2	66	3,2	34
3	91	1,1	9	3	77	3,9	22
4	83	1,2	17	4	69	7,2	31
5	97	0,8	3	5	90	1,7	10
6	84	3,4	16	6	62	3,0	40
7	68	2,0	32	7	59	4,8	41
8	96	0,2	4	8	87	0,8	13
9	91	0	9	9	85	0,4	17
10	93	0	7	10	84	0,3	17
Total	88,1	1,2	11,9	Total	74,7	2,8	25,7
SGW				SBW			
Classe	Média Acertos	FA %	FR %	Classe	Média Acertos	FA %	FR %
1	75	2,3	25	1	71	1,3	29
2	77	3,3	23	2	77	3,0	23
3	80	2,8	20	3	76	4,0	24
4	81	3,3	19	4	76	4,9	24
5	87	1,8	13	5	77	1,1	23
6	73	2,2	27	6	70	2,1	30
7	73	3,0	27	7	71	3,2	29
8	83	1,3	17	8	78	1,4	22
9	83	1,1	17	9	78	0,8	22
10	86	1,0	14	10	77	0,6	23
Total	79,8	2,2	20,2	Total	75,1	2,2	24,9
SEW				SSW			

SGW, SBW e SEW.

Ao aplicar o algoritmo de fusão, há um aumento de 4% no índice de classificações corretas, com referência ao melhor índice de desempenho obtido anteriormente pelo modelo SGW. As taxas de Falsa Aceitação e Falsa Rejeição tiveram seus valores bastante reduzidos se comparados ao modelo SGW. A taxa de FA é de apenas 0,66% e a taxa de FR de 6%.

5.3.1.2 Teste 2

Os resultados do Teste-2, dispostos na tabela 11, mostram o melhor desempenho de classificação para indivíduos da classe 2 e 5 da sequência SGW, com 100% de acertos. A classe 5 é a classe que possui melhor resultado para as 4 sequências, conseqüentemente sua

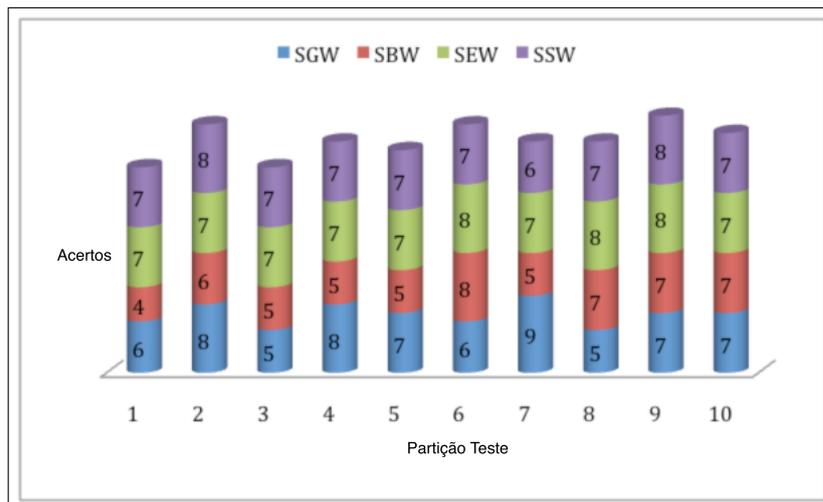


Figura 24: Resultado de desempenho para a classe 7 para o método de validação cruzada com 10 partições. O eixo y representa o número da partição utilizada para teste.

Falsa Rejeição é muito baixa. As classes 9 e 10 não possuem uma FR baixa, no entanto suas taxas de FA são praticamente iguais a zero. Para o conjunto de 100 exemplares o desempenho da classe 7 é superior ao Teste-1 pois o número de ciclos completados na marcha é significativamente maior que no Teste-1, assim, o número de características obtidas consegue identificar, com maior exatidão, o indivíduo.

Os índices de classificações corretas mais baixos pertencem às sequências SBW e SSW. Este resultado demonstra que sendo a quantidade de informação do movimento global do modelo SGW maior do que a dos modelos SBW, SEW e SSW, o GBM reconhece com maior eficiência uma pessoa por meio da marcha nos dois primeiros casos. No entanto, isto não se aplica para todas as classes. Dependendo da característica do movimento da marcha do indivíduo, o SSW captura estas informações com maior precisão do que os demais modelos.

Quando o algoritmo de fusão é realizado, o índice de classificações corretas aumenta. O maior índice FR é de 20% para a classe, sendo que antes da fusão o maior índice é 41% para a classe 7 pertencente ao modelo SGW. A taxa FR diminui sensivelmente em todas as classes, assim como a taxa de Falsa Aceitação.

A Tabela 11 ilustra as Matrizes de Confusão para o Teste-2 com 100 frames/classe.

Tabela 11: Matrizes de Confusão do Teste-2 para a Base A.

Classe	1	2	3	4	5	6	7	8	9	10	FR %	Classe	1	2	3	4	5	6	7	8	9	10	FR %	
1	88	0	0	0	12	0	0	0	0	0	12	1	84	2	0	2	2	10	0	0	0	0	16	
2	0	100	0	0	0	0	0	0	0	0	0	2	9	67	0	9	7	8	0	0	0	0	33	
3	0	4	92	0	0	4	0	0	0	0	8	3	3	2	88	0	0	7	0	0	0	0	12	
4	0	0	0	92	8	0	0	0	0	0	8	4	0	6	0	80	0	0	13	1	0	0	20	
5	0	0	0	0	100	0	0	0	0	0	0	5	5	0	1	0	94	0	0	0	0	0	6	
6	12	4	4	0	5	75	0	0	0	0	25	6	6	10	6	6	0	71	0	1	0	0	29	
7	0	0	0	0	0	10	90	0	0	0	10	7	4	13	6	14	0	0	59	0	2	2	41	
8	0	2	0	0	0	3	0	95	0	0	5	8	0	0	4	0	0	0	8	88	0	0	12	
9	0	5	0	2	0	3	0	0	90	0	10	9	5	2	3	0	0	5	0	0	85	0	15	
10	0	0	0	0	0	0	0	5	0	95	5	10	5	5	2	0	10	3	0	0	0	75	25	
FA%	1,3	1,6	0,4	0,2	2,7	2,2	0	0,5	0	0	91,7	FA%	4,1	4,4	3,4	3,4	2,1	3,6	2,3	0,2	0,2	0,2	79,1	
SGW													SBW											
Classe	1	2	3	4	5	6	7	8	9	10	FR %	Classe	1	2	3	4	5	6	7	8	9	10	FR %	
1	79	4	7	0	7	1	2	0	0	0	21	1	61	12	12	15	0	0	0	0	0	0	39	
2	14	80	4	1	0	0	0	1	0	0	20	2	3	69	6	10	4	2	4	2	0	0	31	
3	1	4	83	0	3	8	1	0	0	0	17	3	8	7	69	11	0	1	1	1	1	1	31	
4	2	4	4	82	0	2	4	1	0	1	18	4	10	15	13	59	0	0	0	0	2	1	41	
5	0	2	0	0	97	0	1	0	0	0	3	5	5	10	0	1	84	0	0	0	0	0	16	
6	2	1	9	6	4	77	1	0	0	0	23	6	3	3	6	6	2	74	3	3	0	0	26	
7	5	2	0	4	0	0	89	0	0	0	11	7	4	10	11	6	2	2	63	2	0	0	37	
8	2	2	4	2	2	2	4	82	0	0	18	8	0	8	1	3	0	6	5	77	0	0	23	
9	5	0	5	0	10	0	5	0	75	0	25	9	0	10	10	0	6	0	0	4	70	0	30	
10	0	0	8	0	0	0	0	0	0	92	8	10	10	0	5	0	5	0	10	0	0	70	30	
FA%	3,4	2,1	4,5	1,2	2,8	1,4	2,0	0,2	0	0,1	83,6	FA%	4,7	8,3	7,1	5,6	2,1	1,2	2,5	1,3	0,3	0,2	69,6	
SGW													SBW											
Classe	1	2	3	4	5	6	7	8	9	10	FR %													
1	80	5	5	0	2	4	0	2	2	0	20													
2	0	100	0	0	0	0	0	0	0	0	0													
3	3	0	94	0	0	0	3	0	0	0	6													
4	0	2	0	98	0	0	0	0	0	0	2													
5	1	0	0	0	98	0	0	1	0	0	2													
6	2	4	0	0	0	86	0	2	3	3	14													
7	2	2	0	0	2	0	94	0	0	0	6													
8	0	1	1	0	0	0	0	93	0	5	7													
9	3	0	2	0	0	1	0	0	93	1	7													
10	3	2	0	0	0	0	0	0	0	95	5													
FA%	1,5	1,7	0,8	0	0,4	0,5	0,3	0,5	0,5	0,1	93,1													
Fusão																								

Alguns fatores contribuem para o aumento nas taxas de FR e FA, tais como:

- A existência de sombras na imagem;
- Falhas na extração dos contornos e contornos falsos;
- Áreas internas da imagem não preenchidas (pequenos “buracos”).

Esses fatores contribuem negativamente para uma correta atribuição de classe conforme é apresentado na Tabela 11. A Figura 25a mostra alguns frames cuja classificação é prejudicada pela descontinuidade de pontos das bordas e a Figura 25b mostra o efeito da sombra na imagem.

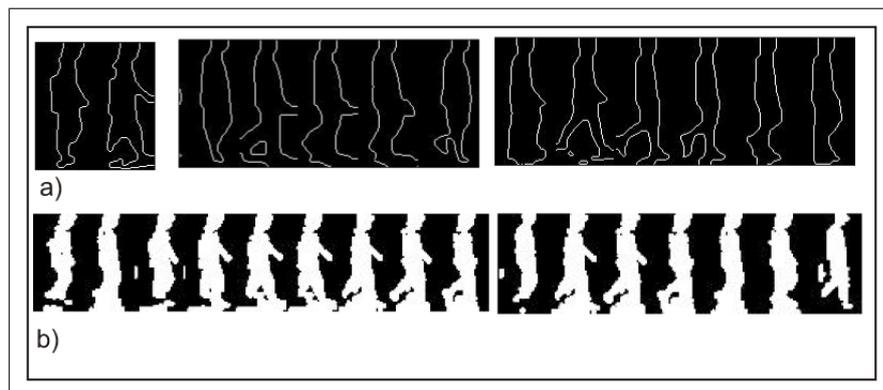


Figura 25: Frames defeituosos que prejudicam uma correta classificação (a) Frames do modelo SEW (b) Frames do modelo SBW.

As curvas $Recall \times Precision$ para os Teste-1 e Teste-2, para os quatro modelos independentes, são ilustradas na Figura 26. Nos dois testes, o melhor desempenho foi obtido utilizando-se a sequência SGW.

A curva obtida por meio da Pontuação Cumulativa de Acertos (CMS) para o Teste-2 é ilustrada na Figura 27, cujo Rank varia de 1 a 20 e na Tabela 12 são apresentados os percentuais de acertos, para os ranks de 1, 5 e 10.

Por meio de avaliação estatística para o modelo GBM, utilizando-se Pontuação Cumulativa de Acertos (CMS), constata-se uma precisão de 99,8% para o melhor caso.

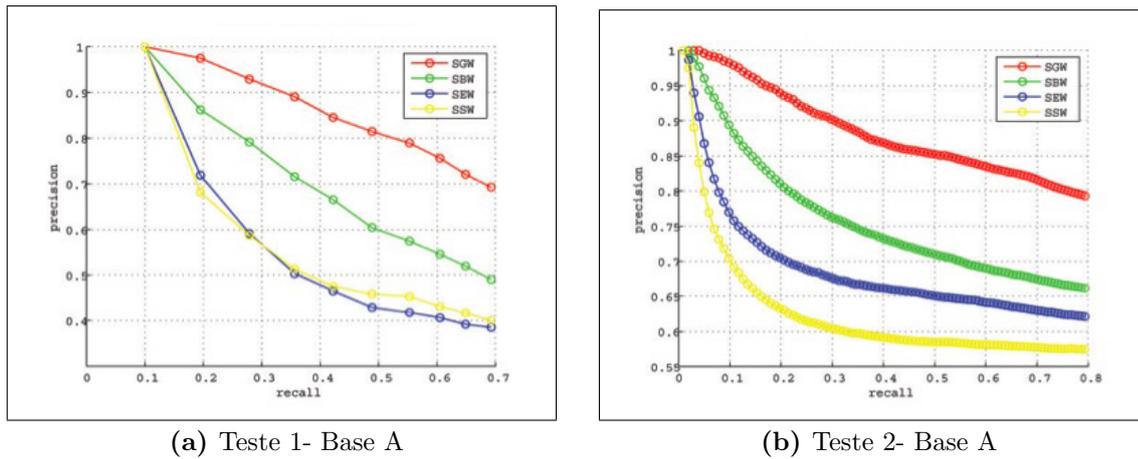


Figura 26: Curvas *Recall* \times *Precision*

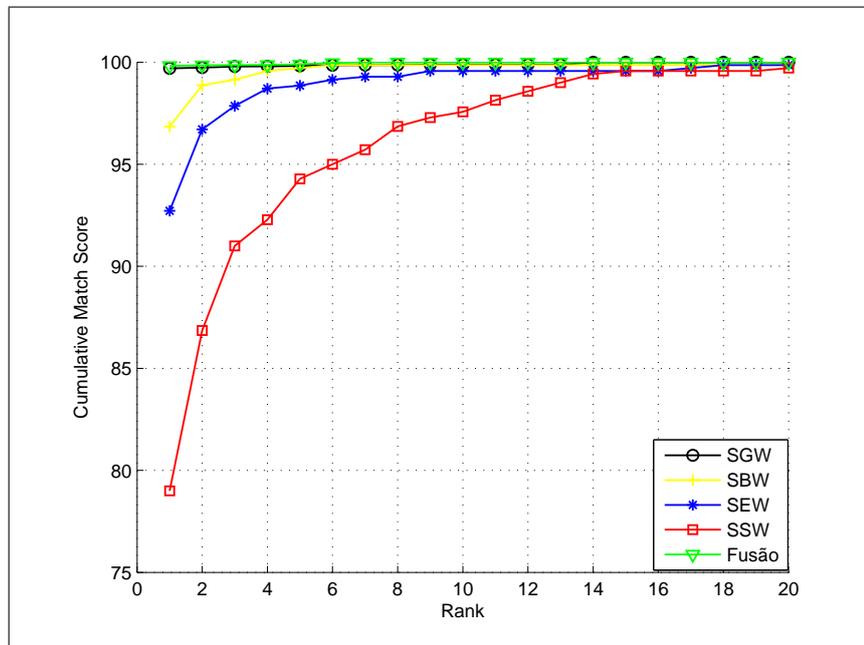


Figura 27: Curva CMS para o Teste-2 – Base A.

Tabela 12: Percentuais CMS obtidos para o Teste-2 – Base A.

Método	Rank %		
	1	5	10
GBM (SGW)	99,7	99,8	99,8
GBM (SBW)	96,8	99,7	99,8
GBM (SEW)	92,7	98,8	99,5
GBM (SSW)	79,0	94,2	97,5
GBM (Fusão)	99,8	100	100

5.3.2 Base B

As Matrizes de Confusão para os modelos SGW e SBW, com o indivíduo caminhando em ângulo de 90° , são apresentadas na Tabela 13 e na Tabela 14, respectivamente.

A Matriz de confusão disposta na Tabela 13 mostra os resultados da classificação com combinação de duas sequências da esquerda para a direita. Cada classe - combinação de duas sequências - possui 72 frames ou elementos da classe. Dentre as 20 classes da Base B, 9 apresentam 100% de classificações corretas (classes 1, 2, 5, 6, 10, 11, 14, 15, 16), ou seja, Falsa Rejeição igual a 0%. Onze classes (classes 2,3, 7, 9, 10, 11, 15, 16, 17, 18, 19) possuem Falsa Aceitação igual ou muito próximas de 0%, ou seja, 65% das classes não admitem ou admitem poucos impostores.

Tabela 13: Matriz de Confusão: Ângulo 90° - Modelo SGW - sequências 2 e 4.

Classe	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	FR%
1	72	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	72	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	59	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	6	26,3
4	0	0	0	56	0	0	0	0	0	0	2	2	0	0	0	0	0	0	0	12	22,2
5	0	0	0	0	72	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	72	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	1	3	0	51	3	1	0	0	0	13	0	0	0	0	0	0	0	29,1
8	29	0	0	0	0	0	0	36	0	0	0	0	0	7	0	0	0	0	0	0	50
9	0	0	0	0	0	17	0	0	51	0	0	0	0	0	2	0	0	0	0	2	29,1
10	0	0	0	0	0	0	0	0	0	72	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	72	0	0	0	0	0	0	0	0	0	0
12	0	0	0	4	0	0	0	0	0	0	0	68	0	0	0	0	0	0	0	0	5,6
13	0	0	0	0	0	0	0	0	0	0	0	4	64	0	0	0	0	0	0	4	11,1
14	0	0	0	0	0	0	0	0	0	0	0	0	0	72	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	72	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	72	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	67	0	0	5	6,94
18	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	70	0	0	2,8
19	0	0	0	0	0	0	0	2	0	1	0	0	0	0	0	0	0	0	69	0	4,2
20	0	0	0	0	23	0	0	0	0	0	0	0	0	0	0	0	1	0	0	48	33,3
FA%	1,5	0,0	0,0	0,6	1,4	0,9	0,1	0,3	0,1	0,1	0,0	0,3	1,2	0,4	0,1	0,0	0,1	0,0	0,0	1,5	89%

As classes que possuem maior índice de Falsa Rejeição são as 8 e 20. Frames pertencente à classe 8 são confundidos com frames da classe 1 e 14, e os da classe 20 com os da classe 5 e 17. No entanto, a classe 8 possui uma FA muito baixa, apenas 2 frames da classe 19 e 3 frames da classe 7 foram classificados erroneamente na classe 8. A classe 20 além da FR alta também possui a FA alta.

Outro dado interessante é em relação à classe 1, 100% de seus elementos são corretamente classificados (FR=0,0%), no entanto, possui Falsa Aceitação alta (FA=1,57%) devido exatamente ao problema detectado na classe 8, que confunde 29 de seus elementos como sendo da classe 1. As classes 2, 11, 15 e 16 além de terem FR iguais a zero, possuem FA iguais ou próximos a zero, ou seja, além de não aceitarem impostores, 100% dos elementos da classe são corretamente classificados.

Observando-se as silhuetas das sequências pertencentes às classes cujos índices de classificações corretas são baixos, notam-se resultados semelhantes aos citados para a

Base A, na seção 5.3.1.

A Matriz de Confusão disposta na Tabela 14 mostra os resultados da classificação com combinação das mesmas sequências (2 e 4) da Base B, considerando-se o modelo SBW.

Dentre as 20 classes da Base B, seis classes apresentam 100% de classificações corretas (classes 1, 5, 10, 14, 15, 18). É importante salientar que as 5 primeiras coincidem (precisão de 100%) com a classificação obtida no modelo SGW, mas a classe 18 tem maior precisão no modelo SBW que no SGW. Onze classes (classes 2, 3, 7, 9, 10, 11, 12, 16, 17, 18, 19) possuem Falsa Aceitação igual ou muito próximas de 0%. Observa-se a substituição da classe 15, no SGW, pela classe 12, no SBW, dentre as que não aceitam impostores. A precisão média manteve-se praticamente constante nos dois modelos.

Tabela 14: Matriz de Confusão: Ângulo 90° – Modelo SBW – sequências 2 e 4.

Class	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	FR%
1	72	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0
2	0	71	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1,4
3	0	0	62	2	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	13,9
4	0	0	0	56	0	0	0	0	0	0	0	0	4	1	0	0	0	0	0	11	22,2
5	0	0	0	0	72	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0,0
6	0	0	0	0	0	71	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1,4
7	0	0	0	0	4	0	54	3	1	0	0	0	10	0	0	0	0	0	0	0	25,0
8	30	0	0	0	0	0	0	36	0	0	0	0	0	6	0	0	0	0	0	0	50,0
9	0	0	0	0	0	4	0	0	62	1	0	0	0	0	4	0	0	0	0	1	13,9
10	0	0	0	0	0	0	0	0	0	72	0	0	0	0	0	0	0	0	0	0	0,0
11	0	0	0	0	0	0	0	0	0	0	71	0	1	0	0	0	0	0	0	0	1,4
12	0	0	1	2	0	0	0	0	0	0	0	69	0	0	0	0	0	0	0	0	4,2
13	0	0	1	1	0	0	0	0	0	0	0	0	60	0	6	0	0	0	0	4	16,7
14	0	0	0	0	0	0	0	0	0	0	0	0	0	72	0	0	0	0	0	0	0,0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	72	0	0	0	0	0	0,0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	69	0	0	0	0	4,2
17	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	66	0	0	4	8,3
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	72	0	0	0,0
19	0	0	0	0	0	0	0	2	0	0	0	0	0	1	0	0	0	0	69	0	4,2
20	0	0	0	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	49	31,9
FA%	1,5	0,0	0,1	1,4	0,2	0,2	0,0	0,26	0,05	0,05	0,0	0,0	1,32	0,58	0,53	0,0	0,0	0,05	0,0	1,05	90%

A Tabela 15 mostra o percentual de acertos global, para os quatro modelos e fusão, considerando-se as três vistas: lateral, oblíqua e frontal da Base B. Cada sequência (1, 2, 3, 4) possui 36 frames e cada combinação de sequências (1 – 3, 2 – 4) possui 72 frames ou elementos de classe e, as 4 sequências juntas possuem 144 frames cada.

Analisando o desempenho das sequências separadamente, o ângulo 0° tem o seu melhor desempenho para o modelo SEW com média de 80% de acertos; para o ângulo de 45° destaca-se a sequência 1 para o modelo SGW, no entanto, o melhor desempenho médio é para o modelo SBW; pode-se destacar a sequência 4 do modelo SEW para o ângulo

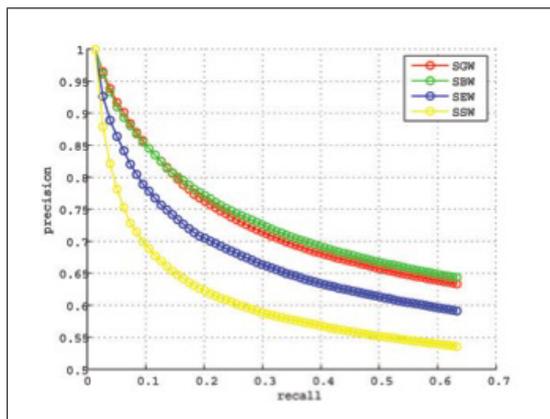
Tabela 15: Porcentagem de acertos para os modelos: SGW, SBW, SEW e SSW; considerando todas as variações de ângulo e sequências.

Seq	Ang.	Modelo				Fusão
		SGW	SBW	SEW	SSW	
1	0	74,3	75,6	83,6	66,3	88,2
2	0	79,1	80,4	82,6	68,3	83,5
3	0	79,1	82,0	85,4	68,1	84,2
4	0	76,6	79,0	82,3	71,1	78,4
1	45	87,2	85,9	81,3	62,0	86,3
2	45	75,6	80,1	82,6	64,7	79,9
3	45	78,8	84,7	77,5	58,1	84,8
4	45	76,8	82,0	80,8	70,0	81,7
1	90	90,5	91,2	88,7	72,9	92,5
2	90	89,8	90,2	87,9	80,5	91,7
3	90	88,6	88,4	87,9	75,4	89,4
4	90	85,4	87,0	92,3	78,1	94,6
1 e 3	0	81,1	82,1	83,0	67,1	89,9
1 e 3	45	78,3	80,8	75,2	58,4	88,7
1 e 3	90	81,1	83,3	86,2	67,8	93,4
2 e 4	0	80,0	80,0	80,9	67,0	82,6
2 e 4	45	78,4	81,0	81,2	66,0	83,7
2 e 4	90	89,0	90,0	89,3	73,4	96,2
1,2,3,4	0	70,5	69,7	70,4	49,6	78,3
1,2,3,4	45	55,0	48,4	59,6	42,7	79,1
1,2,3,4	90	60,2	53,8	65,8	52,5	81,0

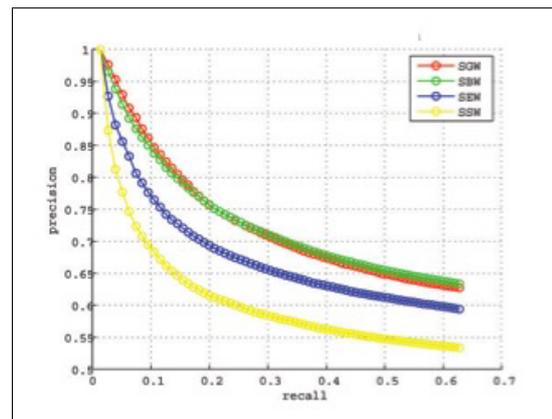
de 90°. No entanto, o desempenho médio dos modelos SGW, SBW e SEW é bastante similar. Para as combinações de sequências 1 – 3 e 2 – 4, o melhor desempenho é para o ângulo de 90°, com 90% de acertos para o para a combinação 2 – 4 do modelo SBW. Considerando-se a combinação das 4 sequências, o melhor desempenho também é para o ângulo de 0° para os modelos SGW, SBW e SEW. No processo de fusão, o melhor resultado é para o ângulo de 90° com 81% de classificações corretas.

Para uma avaliação global de desempenho do modelo GBM, considerando-se os quatro modelos de movimento do corpo humano separadamente(SGW, SBW, SEW, SSW), foram geradas as curvas *Recall* × *Precision* mostradas nas Figuras 28, 29 e 30 para os ângulos: 0°, 45° e 90°, respectivamente.

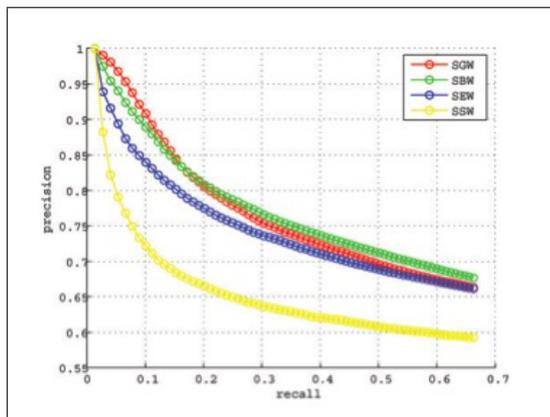
Como pode ser observado, o melhor desempenho geral é para o ângulo 90° com os



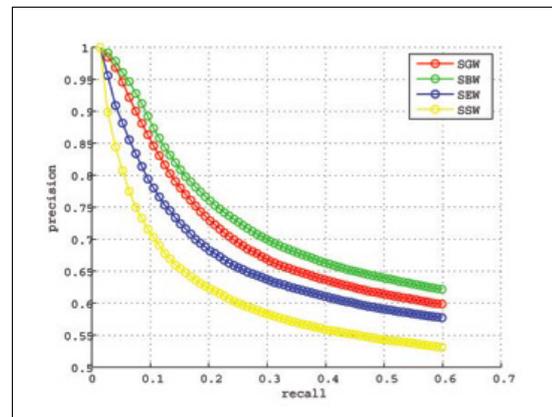
(a) seqüências 1 e 3



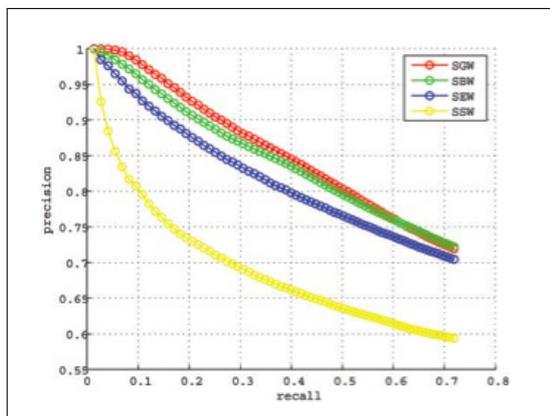
(b) seqüências 2 e 4

Figura 28: Curvas Recall x Precision - Ângulo 0° 

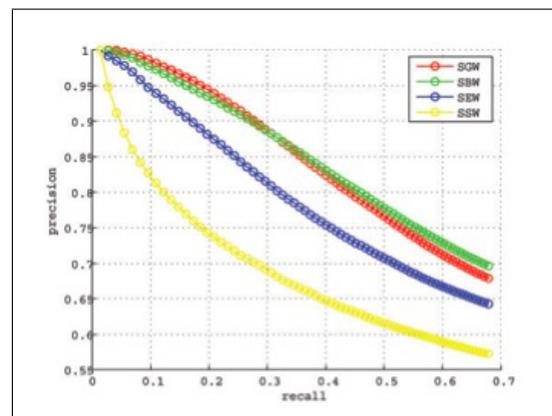
(a) seqüências 1 e 3



(b) seqüências 2 e 4

Figura 29: Curvas Recall x Precision - Ângulo 45° 

(a) seqüências 1 e 3



(b) seqüências 2 e 4

Figura 30: Curvas Recall x Precision - Ângulo 90°

modelos SGW e SBW, para a combinação de sequências Direita-Esquerda, utilizando-se 72 frames – conforme está disposto na Figura 30(b).

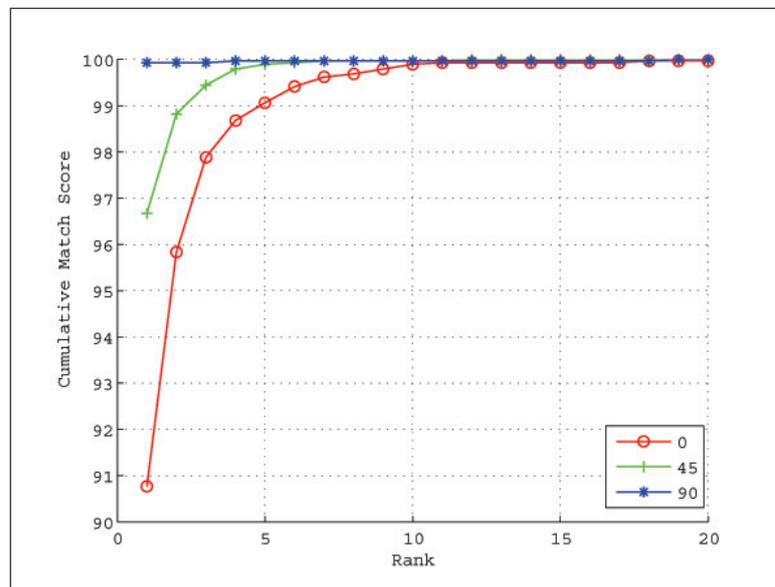


Figura 31: Modelo SGW, considerando a combinação das 4 sequências de caminhada.

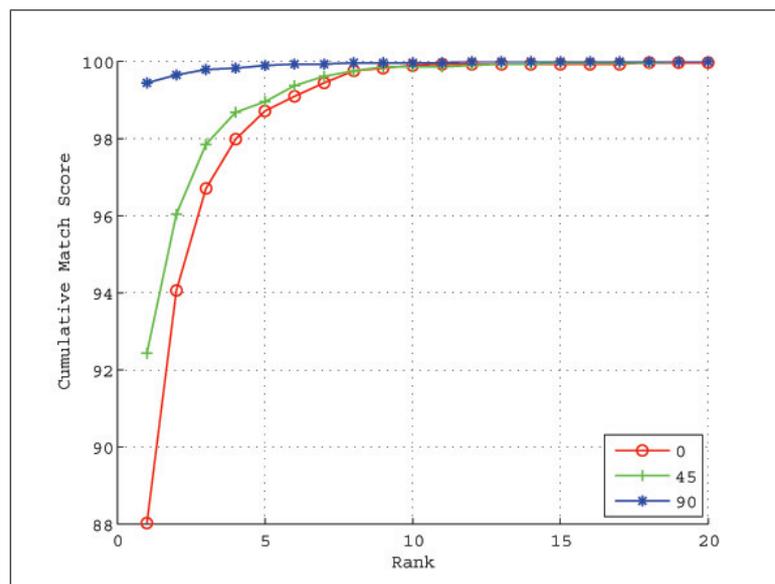


Figura 32: Modelo SBW, considerando a combinação das 4 sequências de caminhada.

Gráficos de desempenho do sistema foram gerados por meio das curvas CMS ilustradas na Figura 31, Figura 32, Figura 33 e Figura 34, cujo valor do rank varia de 1 a 20, para os 4 modelos do GBM para os ângulos 0° , 45° e 90° . Para cada um destes ângulos utilizam-se duas sequências de caminhada.

O gráfico de desempenho com o módulo de fusão está disposto na Figura 35, o valor

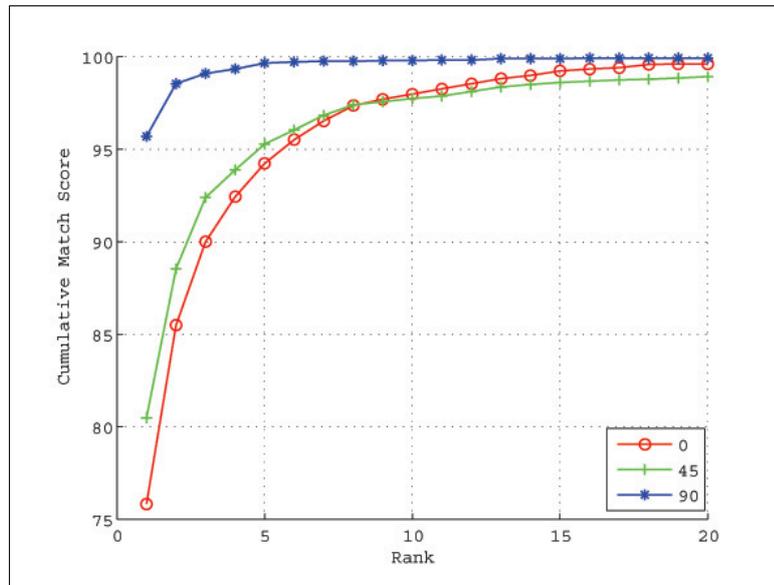


Figura 33: Modelo SEW, considerando a combinação das 4 seqüências de caminhada.

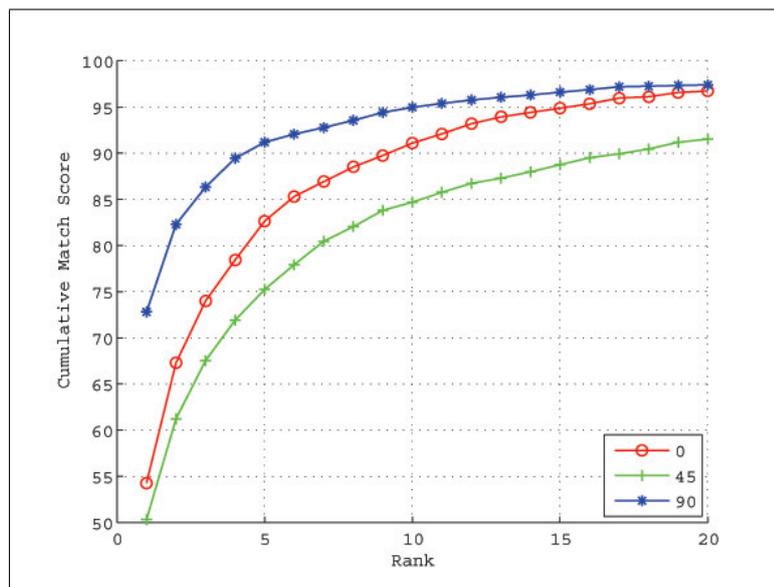


Figura 34: Modelo SSW, considerando a combinação das 4 seqüências de caminhada.

do rank varia de 1 a 20; com os ângulos 0° , 45° e 90° e utilizando-se a combinação das quatro seqüências de caminhada. Por meio da fusão dos quatro modelos, observa-se que o índice de classificações corretas aumenta em cada uma das seqüências avaliadas. Com isto, o desempenho global do sistema tem ganho significativo.

O desempenho do modelo GBM foi comparado com os trabalhos de BenAbdelKader (BENABDELKAER; CUTLER; DAVIS, 2002) Collins (COLLINS; SHI, 2002), Lee (LEE; GRIMSON, 2002), Philips (PHILLIPS et al., 2002), Wang (WANG et al., 2003), Kale (KALE et al.,

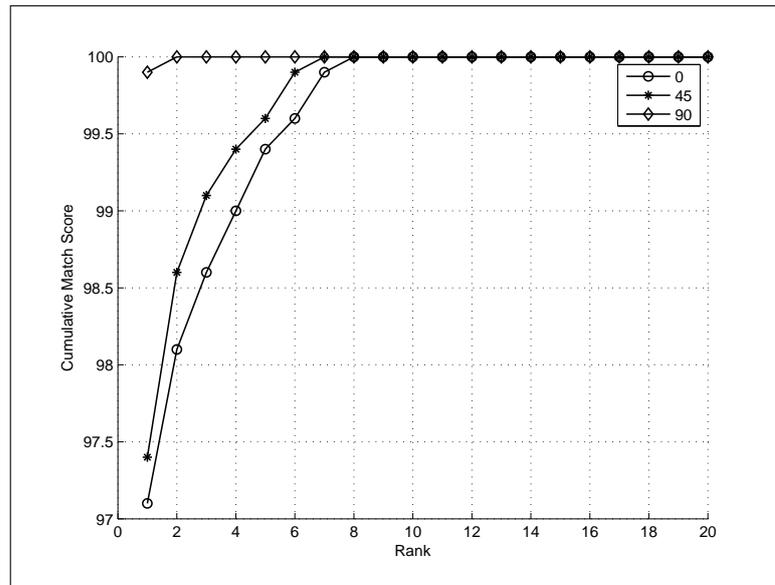


Figura 35: Fusão dos 4 modelos, considerando a combinação das 4 seqüências de caminhada.

2003) e Hong (HONG et al., 2007) - trabalhos recentes cuja base de vídeos para reconhecimento biométrico por meio da marcha é a mesma utilizada em nossos experimentos, aqui denominada de Base B. Na Tabela 16 são mostrados e comparados os valores percentuais de classificações corretas para os Ranks 1, 5 e 10, utilizando-se as 4 seqüências de caminhada, para as três vistas.

Os resultados para as seqüências esquerda-direita foram também comparados com o trabalho de Liu (LIU; ZHENG, 2007), conforme apresentado na Tabela 17.

Estas seqüências e ângulos foram escolhidos para comparação, devido ao fato das mesmas serem utilizadas em cada trabalho separadamente.

Os resultados obtidos para o modelo GBM aplicado na vista lateral, com combinação das quatro seqüências (direita-esquerda e esquerda-direita) (Tabela 16), mostraram o melhor desempenho estatístico global para o método SGW, com 90,9% de classificações corretas para o Rank 1. Ou seja, em 90,9% dos casos, a menor distância pontuada corresponde a um exemplar de teste verdadeiro. Com esta pontuação, supera-se os resultados das técnicas propostas por BenAbdelKader (BENABDELKAER; CUTLER; DAVIS, 2002) Collins (COLLINS; SHI, 2002), Lee (LEE; GRIMSON, 2002), Philips (PHILLIPS et al., 2002), Wang (WANG et al., 2003) e Kale (KALE et al., 2003).

Tabela 16: Tabela comparativa do modelo GBM para a vista lateral (0°) – Base B.

Ângulo	Método	Rank (%)		
		1	5	10
0°	BenAbdelkader	72.5	88.75	96.25
	Collins	71.25	78.75	87.50
	Lee	87.50	98.75	100
	Philips	78.75	91.25	98.75
	Wang	75.00/82.50	97.50/100	100/100
	Kale	82	92	96,2
	Hong	96	100	100
	GBM(SGW)	90,9	99,0	100
	GBM(SBW)	88,0	98,7	99,9
	GBM(SEW)	78,8	94,2	97,9
	GBM(SSW)	54,2	67,2	73,9
	GBM (Fusão)	97,1	99,4	100
45°	Kale	92,5	97,5	
	Hong	96,25	98,75	
	GBM (SGW)	96,7	99,9	
	GBM (SBW)	92,4	99,9	
	GBM (SEW)	80,4	95,2	
	GBM (SSW)	50,3	82,6	
	GBM (Fusão)	97,4	99,6	
90°	Kale	77,5	88,75	
	Hong	88,75	95	
	GBM (SGW)	99,9	99,9	
	GBM (SBW)	99,4	99,9	
	GBM (SEW)	95,6	99,6	
	GBM (SSW)	72,8	91,1	
	GBM (Fusão)	99,9	100	

Tabela 17: Tabela comparativa para a vista oblíqua, combinação de duas sequências sentido esquerda-direita.

Ângulo	Método	Rank (%)		
		1	2	3
45°	Liu	90	95	95
	GBM (SGW)	95,7	98	99
	GBM (SBW)	93,5	97	98,3
	GBM (SEW)	86,7	94	96
	GBM (SSW)	79	87	95,6
	GBM (Fusão)	96,8	98,9	99,8

Os resultados para o modelo SBW para a mesma vista também superam a maior parte dos outros algoritmos. Avaliando-se os resultados da Tabela 16, o melhor desempenho

estatístico foi obtido para o ângulo de 90° , com o modelo SGW, gerando 99,9% das classificações corretas no Rank 1. Este resultado é superior ao de Kale (KALE et al., 2003) e Hong (HONG et al., 2007). Os modelos SBW e SEW apresentaram respectivamente 99,4% e 95,6% dos acertos pertencentes ao Rank 1 superando também as técnicas dos demais autores.

Quando o sentido da direção de caminhada é avaliado separadamente, como, por exemplo, considerando-se as sequências esquerda-direita, os resultados são ainda melhores, considerando-se a vista lateral. Como pode ser visto nos resultados dispostos na Tabela 17, onde o modelo GBM(SGW) mostrou 95,7% de classificações corretas para o conjunto de imagens considerando-se Rank 1 e, mesmo para a imagem esqueletizada (SSW), que é a que tem menor desempenho global, foram obtidas 95,6% de classificações corretas com o Rank 3, tornando viável o reconhecimento apenas com o esqueleto da imagem, superando, assim, os resultados da técnica proposta por Liu (LIU; ZHENG, 2007).

5.3.3 Base C

As Figuras 36, 37, 38 e 39 mostram as curvas CMS para a combinação de seis sequências sem bolsa e sem casaco obtidas com os quatro modelos do GBM nos ângulos de 90° , 108° e 180° graus.

Em relação à variação de ângulos, para essa base os melhores resultados foram obtidos com os ângulos 0° , 90° , 108° e 180° . Um fato importante a se destacar é que as variações de vestimenta, utilizadas nos vídeos desta Base C, interferem muito pouco na classificação dos indivíduos em suas respectivas classes. Isto pode ser observado nos experimentos onde as sequências são combinadas, ou seja, sequências de pessoas carregando bolsa com sequências de pessoas vestindo casaco; sequências com bolsa com sequências sem bolsa e sem casaco e, sequências com casaco com sequências sem bolsa e sem casaco (normal).

As curvas CMS das Figuras 40, 41, 42 e 43 foram obtidas utilizando-se a combinação

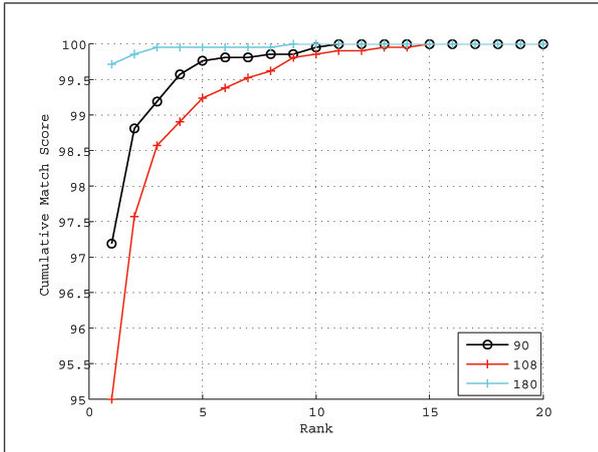


Figura 36: Curvas CMS SGW.

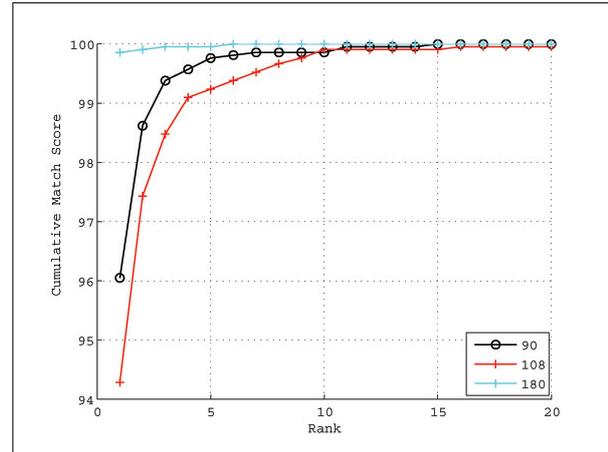


Figura 37: Curvas CMS SBW.

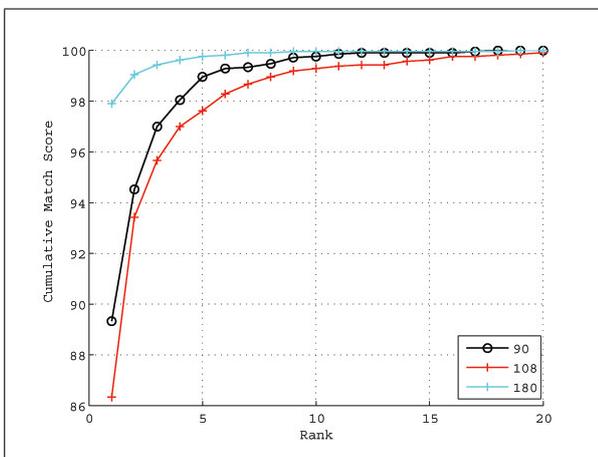


Figura 38: Curvas CMS SEW.

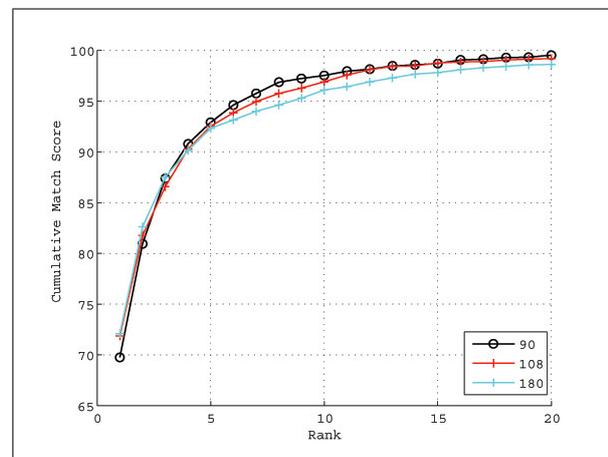


Figura 39: Curvas CMS SSW.

de duas sequências com casaco, variando-se os ângulos de 18° a 180° .

Nas Tabelas 18 e 19 estão dispostos os valores de classificações corretas obtidas por meio do método CMS para todas as variações de ângulos, considerando-se a combinação de duas sequências do tipo bg e de duas sequências do tipo cl. Observa-se que a sequência bg possui ótimo desempenho para uma grande variação de ângulos. Para os ângulos de 0° , 18° e 180° , o número de acertos corretos está próximo a 100% para o Rank 1. No entanto, para alguns ângulos (36° , 54° , 72°), a precisão de acertos cai para o Rank 1, e somente chega próximo a 100% para o Rank 10. Com relação às sequências do tipo cl, o número de acertos pertencentes ao topo é menor do que o número de acertos obtidos com as sequências do tipo bg, para todos os modelos GBM (SGW, SBW, SEW e SSW). Desta

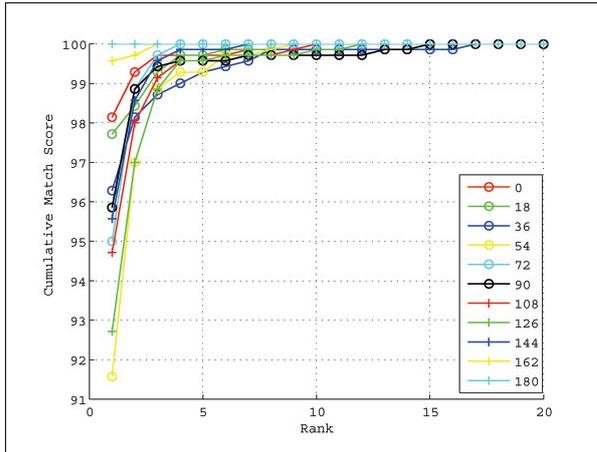


Figura 40: Curvas CMS SGW - cl.

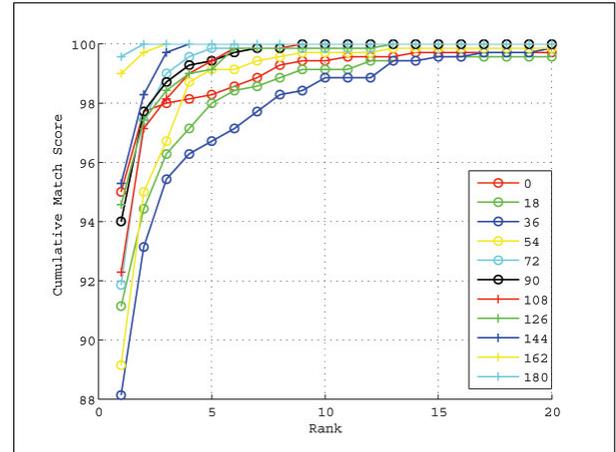


Figura 41: Curvas CMS SBW - cl.

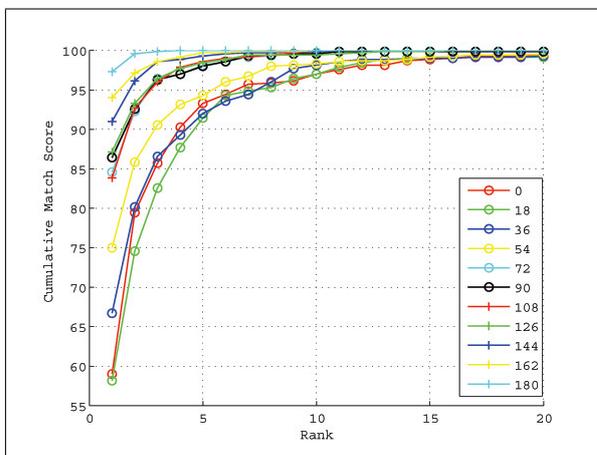


Figura 42: Curvas CMS SEW - cl.

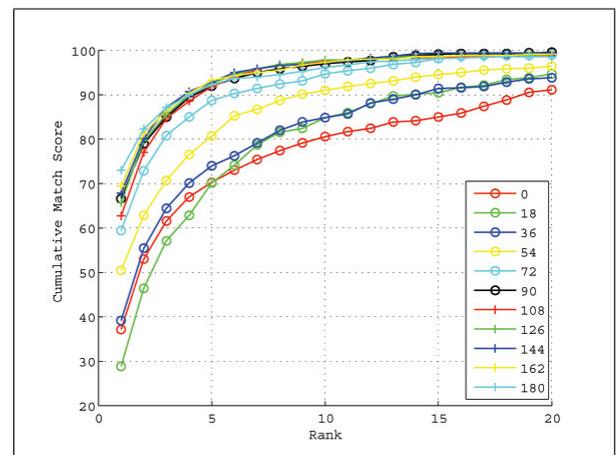


Figura 43: Curvas CMS SSW - cl.

forma, conclui-se que o uso de vestimenta, como o casaco, interfere mais negativamente em comparação ao uso de bolsa para a correta classificação dos indivíduos. O fator que mais influencia na incorreta classificação dos mesmos em suas respectivas classes são os ângulos de caminhada.

Na tabela 20 os valores das classificações foram obtidos com o método CMS, no qual são reproduzidos os melhores resultados em relação à variação de ângulo, após a aplicação do módulo de fusão. Nesta tabela consta a combinação de duas sequências com pessoas carregando bolsa e duas sequências de pessoas com casaco. A primeira sequência mostra excelente desempenho sobre a maioria dos ângulos do experimento. Para os ângulos 0° e 180° , o número de respostas corretas é menor em relação aos resultados obtidos com a

Tabela 18: Tabela Comparativa para variação de ângulos (0° a 90°), considerando-se duas combinações das sequências do tipo bg e duas sequências do tipo cl.

Ângulo	Sequência	Modelo	Rank (%)		
			1	5	10
0°	bg	BGM (SGW)	99,5	100	100
		BGM (SBW)	97,1	99,8	99,8
		BGM (SEW)	59,8	93,2	98,1
		BGM (SSW)	46,2	84,2	92,4
	cl	BGM (SGW)	98,1	99,7	99,8
		BGM (SBW)	95	98,2	99,4
		BGM (SEW)	59	93,2	97
		BGM (SSW)	37,1	70,2	80,5
18°	bg	BGM (SGW)	99	100	100
		BGM (SBW)	93,8	99,1	99,7
		BGM (SEW)	70,8	95,4	97,7
		BGM (SSW)	46,1	79,4	91,5
	cl	BGM (SGW)	97,7	99,7	100
		BGM (SBW)	91,1	98	99,1
		BGM (SEW)	58,1	91,4	97
		BGM (SSW)	28,8	70,1	84,8
36°	bg	BGM (SGW)	99,1	99,8	100
		BGM (SBW)	93,4	98,4	99,5
		BGM (SEW)	76,5	96,7	99,1
		BGM (SSW)	49,5	82,2	91,2
	cl	BGM (SGW)	96,2	99,2	99,8
		BGM (SBW)	88,1	96,7	98,8
		BGM (SEW)	66,7	92	98,1
		BGM (SSW)	39,1	74	84,8
54°	bg	BGM (SGW)	95,4	99,2	100
		BGM (SBW)	94,2	98,7	100
		BGM (SEW)	80,7	96,7	99
		BGM (SSW)	61,4	88,4	93,4
	cl	BGM (SGW)	91,5	99,2	100
		BGM (SBW)	89,1	99,1	99,7
		BGM (SEW)	75	94,2	98,2
		BGM (SSW)	50,4	80,7	91
72°	bg	BGM (SGW)	91,7	99,2	99,8
		BGM (SBW)	91,4	98,8	99,8
		BGM (SEW)	81,5	96,4	98,5
		BGM (SSW)	64,1	91,2	96,5
	cl	BGM (SGW)	95	100	100
		BGM (SBW)	91,8	99,8	99,8
		BGM (SEW)	84,5	98,1	99,7
		BGM (SSW)	59,4	88,7	94,7
90°	bg	BGM (SGW)	93,8	99,7	100
		BGM (SBW)	91,5	99,5	100
		BGM (SEW)	83	97,8	99,1
		BGM (SSW)	70	89	97,4
	cl	BGM (SGW)	95,8	99,5	99,7
		BGM (SBW)	94	99,4	100
		BGM (SEW)	86,4	98	99,5
		BGM (SSW)	66,5	92	97

primeira sequência. Considerando-se estes dados obtidos, pode-se concluir que a roupa influi mais negativamente no processo de classificação que o uso de bolsas.

As curvas *Recall* \times *Precision* também são geradas para este modelo, ilustradas nas Figuras 44, 45 e 46 para o modelo SGW, considerando-se as sequências com pessoas, carregando uma bolsa, vestindo casaco e sem bolsa e sem casaco. Foram utilizadas as combinações de duas sequências carregando bolsa, duas sequências vestindo casaco e seis

Tabela 19: Tabela Comparativa para variação de ângulos (108° a 180°), considerando-se duas combinações das sequências do tipo bg e duas sequências do tipo cl.

Ângulo	Sequência	Modelo	Rank (%)		
			1	5	10
108°	bg	BGM (SGW)	94,5	99,2	99,8
		BGM (SBW)	92,1	99,7	100
		BGM (SEW)	86,1	98,1	99,5
		BGM (SSW)	71	93,8	97,5
	cl	BGM (SGW)	94,7	99,5	100
		BGM (SBW)	92,2	99,4	100
		BGM (SEW)	83,8	98,5	99,8
		BGM (SSW)	62,7	91,8	97,5
126°	bg	BGM (SGW)	96,8	99,7	99,8
		BGM (SBW)	96,1	99,7	99,8
		BGM (SEW)	92,7	98,8	99,5
		BGM (SSW)	75,7	94,1	97,7
	cl	BGM (SGW)	92,7	99,5	99,8
		BGM (SBW)	94,5	99,1	99,8
		BGM (SEW)	87,1	98,4	99,4
		BGM (SSW)	65,7	92,5	97,8
144°	bg	BGM (SGW)	97,7	99,8	99,8
		BGM (SBW)	98,1	100	100
		BGM (SEW)	93,8	99	99,5
		BGM (SSW)	79	94,2	97,5
	cl	BGM (SGW)	95,5	99,8	100
		BGM (SBW)	95,2	100	100
		BGM (SEW)	91	99,2	99,8
		BGM (SSW)	67,8	92,8	97,2
162°	bg	BGM (SGW)	99,5	99,8	100
		BGM (SBW)	100	100	100
		BGM (SEW)	98,8	100	100
		BGM (SSW)	80,8	95,7	98
	cl	BGM (SGW)	99,5	100	100
		BGM (SBW)	99	100	100
		BGM (SEW)	94	99,7	100
		BGM (SSW)	69,4	93,2	97,2
180°	bg	BGM (SGW)	99,8	100	100
		BGM (SBW)	100	100	100
		BGM (SEW)	99,5	100	100
		BGM (SSW)	84,8	97,1	98,7
	cl	BGM (SGW)	95	98,2	99,4
		BGM (SBW)	99,5	100	100
		BGM (SEW)	97,2	100	100
		BGM (SSW)	73	92,2	96,1

sequências sem bolsa e sem casaco, estas figuras encontram-se no apêndice A.

As Figuras 47, 48 e 49 ilustram as curvas $Recall \times Precision$ para o modelo SBW, considerando os três tipos de sequências (bolsa, casaco e normal), estas figuras encontram-se no apêndice A.

As Figuras 50, 51 e 52 ilustram as curvas $Recall \times Precision$ para o modelo SEW e, as Figuras 53, 54 e 55 ilustram as curvas $Recall \times Precision$ para o modelo SSW. Considerando também, os três tipos de sequências, estas figuras encontram-se no apêndice A.

Tabela 20: Variação de ângulos, considerando duas sequências de caminhada com pessoas carregando bolsa e duas sequências com pessoas vestindo casaco, após a fusão.

Ângulo	Sequência	Rank		
		1	5	10
0°	Bolsa	99,6	99,8	100
	Casaco	98,4	99,8	99,9
90°	Bolsa	95,7	100	100
	Casaco	97,6	100	100
108°	Bolsa	98,7	100	100
	Casaco	97,7	99,8	99,9
180°	Bolsa	99,9	100	100
	Casaco	97,8	99,7	100

Na Figura 56(a) são representadas as curvas $Recall \times Precision$ com a combinação de duas sequências carregando bolsa com duas sequências vestindo casaco para os ângulos de 0°, 18°, 36° e 54°. Na Figura 56(b) são mostradas as curvas $Recall \times Precision$, com a combinação de duas sequências com bolsa com duas sequências com casaco e na Figura 56(c), a combinação de duas sequências com casaco com duas sequências sem bolsa e sem casaco, para os ângulos de 0°, 72°, 90° e 108°. Todas as curvas referem-se ao modelo SGW e encontram-se no apêndice A.

5.4 Considerações Finais

Neste capítulo foram descritos e analisados os resultados dos experimentos realizados para a avaliação da metodologia proposta, considerando os quatro modelos de representação da forma humana.

Os resultados obtidos com a metodologia desenvolvida para o modelo GBM demonstram a validade da proposta.

Capítulo 6

Conclusão

O modelo GBM foi designado levando em consideração os diferentes tipos de informações contidas no movimento do corpo humano, capturadas utilizando-se diferentes segmentações de imagens de sequência de vídeo.

Como as pessoas têm suas próprias características de caminhar, e geralmente se vestem diferentemente umas das outras, o modelo proposto visa capturar as informações globais contidas no movimento humano, para reconhecê-las. De acordo com este trabalho, estas informações globais estão contidas nas representações em escala de cinza ($3D$), juntamente com as imagens binarizadas ($2D$), no contorno das silhuetas do corpo em movimento e no movimento das juntas (esqueleto). Para validar o modelo proposto, foram analisadas sequências de vídeo capturadas em diversos ângulos.

Quatro novos modelos de representação humana foram gerados por meio da segmentação de *frames* com as características globais do movimento humano. O modelo SGW (Silhouette-Gray-Wavelet) oferece informações globais sobre o movimento, em escala de cinza. O modelo SBW (Silhouette-Binary-Wavelet) fornece as informações globais do movimento humano, contidas nas imagens binárias. O modelo SEW (Silhouette-Edge-Wavelet) contém as informações globais do movimento humano em relação ao contorno da silhueta. Finalmente, o modelo SSW (Silhouette-Skeleton-Wavelet) fornece as informações globais do movimento humano contidas no esqueleto do corpo.

As seqüências de vídeo, pertencentes a diferentes pessoas, ou classes, são consideradas como um elemento da classe de cada indivíduo. O resultado é um conjunto de bases de dados, no qual cada frame é considerado como um elemento da classe. O objetivo do método é classificar cada indivíduo em sua respectiva classe.

Para avaliação do método, três bases de vídeo foram utilizadas e os resultados obtidos por uma das bases foram comparados com trabalhos recentes encontrados na literatura.

Os resultados obtidos utilizando o método GBM sugerem a viabilidade do método para melhorar o desempenho. Por exemplo, considerando-se o modelo GBM(SGW), foram obtidas 100% de classificações corretas para as três bases de dados testadas, para um indivíduo de uma das classes com movimento perpendicular ao observador em ângulo de 0° .

Os experimentos realizados com os modelos SEW e SSW também alcançaram taxas de classificação corretas satisfatórias, quando comparadas com outros métodos encontrados na literatura, com a vantagem da redução do conjunto de dados ou informações.

Quando as pessoas caminham naturalmente ao longo de um trajeto perpendicular ao observador (90°), dependendo do ângulo de visão, seu tamanho não é alterado. No entanto, em outros casos, o tamanho da silhueta altera-se de acordo com a projeção geométrica. Estas mudanças afetam o desempenho de classificação. O desempenho do método é então reduzido para o ângulo 0° da base B e para os ângulos 18° , 36° , 54° , 126° , 144° e 162° da base C.

A validação de desempenho do método GBM foi realizada por meio de matrizes de confusão, curvas *Recall* \times *Precision* e curvas CMS. Por meio das curvas CMS, pode-se concluir que as respostas corretas pertencem ao topo de k buscas na maioria dos experimentos realizados.

Quando o processo de fusão é aplicado sobre as taxas de classificação corretas obtidas separadamente em cada modelo (SGW, SBW, SEW e SSW), é possível o refinamento dos resultados. A combinação das pontuações de similaridades obtidas em cada modelo proporciona melhor desempenho do sistema. Esta melhora no desempenho pode ser com-

provado pela redução das taxas de Falsa Aceitação (FA) e Falsa Rejeição (FR) e, na geração das curvas CMS. Estes resultados provam a hipótese inicial de que cada modelo carrega diferentes informações do corpo humano durante a caminhada.

A metodologia proposta não é adequada para aplicações em tempo-real, mas é apropriada para busca em bases de vídeo.

Tabela 21: Desempenho do sistema para a base B.

GMM	Haar	NN	Fusão
10 – 12 <i>frames</i> por segundo	\cong 3500 segundos para processar 2880 <i>frames</i>	\cong 501,6715 segundos para classificar 2880 <i>frames</i>	\cong 0,1696 segundos de processamento

O custo computacional do sistema é da ordem de $O(n^2)$. Para chegar a esta conclusão, foram somados os custos computacionais dos vários módulos que integram o sistema, no qual influenciam o desempenho do sistema: a redução de escala pela TW de Haar, o classificador do vizinho mais próximo e o processo de fusão. O custo computacional da TW de Haar é $O(n)$, do classificador do vizinho mais próximo é $O(n^2)$ e da fusão é $O(n^2)$. Portanto, o custo computacional total do sistema é $O(n^2)$ - complexidade quadrática, que é o maior custo de um dos módulos que compõem o sistema, sendo este um custo aceitável para aplicações em vídeo.

O desempenho do sistema, utilizando-se um *MackBook Apple* com processador *Intel Core 2 DUO*, 2GB de memória RAM e com o sistema operacional MAC OSX v.10.5.6, para o processamento da base B, pode ser avaliado na tabela 21.

6.1 Publicações

- *Método de Reconhecimento da Marcha Humana por meio da Fusão das Características do Movimento Global.* WVC 2009 - V Workshop de Visão Computacional. São Paulo-SP.

- *Human Gait Recognition using Extraction and Fusion of Global Motion Features.*
Multimedia Tools and Applications, Springer,ISSN: 1380-7501 (sob revisão)

Referências

- ANDRIACCHI, T.; OGLE, J.; GALANE, J. Walking speed as a basis for normal and abnormal gait measurements. *Biomech Journal*, v. 10, n. 4, p. 261–268, 1977.
- BENABDELKAER, C.; CUTLER, R.; DAVIS, L. Motion-based recognition of people in eingengait space. *Proceedings of fifth IEEE International Conference on Automatic Face and Gesture Recognition*, p. 267–272, 2002.
- BOBICK, A.; DAVIS, J. The recognition of human movement using temporal templates. *IEEE Trans. PAMI*, v. 23, n. 3, p. 257–267, March 2001.
- BOULGOURIS, N.; ZHIWEI, X. Gait recognition using radon transform and linear discriminant analysis. *IEEE Transactions on Image Processing*, v. 16, 2007.
- CASIA. *NLPR*. January 2005. Disponível em: <www.cbsr.ia.ac.cn>.
- CLARK, J.; YUILLE, A. *Data Fusion for Sensory Information Processing Systems*. [S.l.]: Kluwer, 1990. (The Springer International Series in Engineering and Computer Science, v. 105).
- COLLINS, R.; SHI, J. Silhouette-based human identification from body shape and gait. *Proc. International Conference Automatic face and gesture recognition*, 2002.
- CUTTING, J.; KOZLOWSKI, L. Recognizing friends by their walk:gait perception without familiarity cues. *Bulletn of the Psychonomic Society*, p. 9:353–356, 1977.
- CUTTING, J.; PROFFI, D. *Gait perception as an example of how we perceive events*. [S.l.]: Plenum Press, London., 1981.
- DITTRICH, W. H. Action categories and the perception of biological motion. *Perception*, v. 22, n. 1, p. 15–22, 1993.
- GROSS, R.; SHI, J. *The CMU Motion of Body MOBO database*. [S.l.], 2001.
- HAN, J.; BHANU, B. Individual recognition using gait energy image. *IEEE Transactions On Patterns Analysis and machine Intelligence*, v. 28, n. 2, p. 316–322, February 2006.

- HAVASI, L.; ZOLTÁN, S.; SZIRÁNYI, T. Detection of gait characteristics for scene registration in video surveillance system. *IEEE Transactions on Image Processing*, v. 16, n. 2, 2007.
- HE, Q.; DEBRUNNER, C. Individual recognition from periodic activity using hidden markov models. *Proc. IEEE Workshop Human Motion*, 2000.
- HONG, L.; JAIN, A.; PANKANT, S. Can multibiometrics improve performance? *Proc. AutoID'99*, v. 12, p. 59–64, 1999.
- HONG, S. et al. Fusion of multiple gait features for human identification. *International Conference on Control, Automation and Systems*, p. 2121–2125, 2008.
- HONG, S. et al. A new gait representation for human identification: Mass vector. *IEEE - Second Conference on Industrial Eletronics and Applications.*, p. 669–673, 2007.
- HUANG, P.; HARRIS, C.; NIXON, M. Canonical space representation for recognition humans by gait and face. *IEEE - Southwest Symposium on Image Analysis and Interpretation*, p. 180–185, April 1998.
- INTEL. *Opencv Library*. 2006. Disponível em: <<http://opencvlibrary.sourceforge.net>>.
- JAIN, A. K.; ROSS, A.; PRABHAKAR, S. An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, v. 14, n. 1, p. 4–20, 2004.
- J.LITTLE; J.BOYD. Recognizing people by their gait: the shape of motion. *Videre: Journal of Computer Vision Research*, p. 2–32, 1998.
- JOHANSSON, G. Visual motion perception. *Scientific American*, p. 232:76–88, 1975.
- KALE, A. et al. Gait analysis for human identification. *AVPA*, p. 706–714, 2003.
- KALE, A. et al. Gait-based recognition of humans using conituous hmms. *International Conf. Automatic Face and Gesture Recognition.*, 2002.
- KAWTRAKULLPONG, P.; BOWDEN, R. An improved adaptive background mixture model for real-time tracking with shadow detection. *Video Based Surveillance Systems: Computer Vision and Distributed Processing*, p. 135–144, September 2001.
- KENT, J. *New directions in Shape Analysis. The art of statistical science: a tribute to g.s. Watson*. [S.l.]: Wiley, Chchester, 1992.

- LAM, T. H.; LEE, R. S. Advantaces in biometrics. In: _____. [S.l.]: Spinger Berlin / Heidelberg, 2005. (Lecture notes in Computer Science, v. 3832), cap. A new representation for Human Gait Recognition - Motion Silhouettes Image (MSI), p. 612–618.
- LAPTEV, I. On space-time interest point. *International Journal of Computer Vision*, v. 64, n. 2-3, p. 107–123, September 2005.
- LEE, L.; GRIMSON, W. Gait analysis for recognition and classification. *IEEE Conference on Face and Gesture Recognition*, p. 155–161, 2002.
- LIU, J.; ZHENG, N. Gait history image: A novel temporal template for gait recognition. *IEEE - International Conference on Multimedia and Expo 2007*, p. 663–666, 2007.
- LIU, Z.; SARKAR, S. Improved gait recognition by gait dynamics normalization. *IEEE Transactions on Pattern Analysis and machine Intelligence*, v. 28, n. 6, June 2006.
- MURASE, H.; SAKAI, R. Moving object recognition in eigenspace recognition: Gait analysis and lip reading. *Pattern Recognition*, v. 17, p. 155–162, 1996.
- MURRAY, M.; DROUGHT, A.; KORY, R. Walking patterns of normal men. *Bone Joint Surg.*, v. 46A, n. 2, p. 335–360, 1964.
- NIYOGI, S.; ADELSON, E. Analyzing and recognizing walking figures in xyt. *Proceedings CVPR*, p. 469–474, 1994.
- PERRY, J. *Análise da Marcha*. [S.l.]: Manole, 2005.
- PHILIPS, P. et al. The feret evaluation methodology for face-recognition. *IEEE - Transactions on Pattern Analysis and Machine Intelligence.*, v. 22, p. 1090–1104, 2000.
- PHILLIPS, P. et al. Baseline results for challenge problem of human id using gait analysis. *Proc. IEEE International Conference on Automatic Face and Gesture recognition.*, p. 137–142, 2002.
- RIBEIRO, H.; GONZAGA, A. Hand image segmentation in video sequence by gmm: A comparision analysis. *XIX Symposium on Computer Graphics and Image Processing*, p. 357–364, 2006.
- SARKAR, S. et al. The humanid gait challenge problem: data sets, performance, and analysis. *IEEE Trans. on Pattern Analysis and Machine Inteligence*, v. 27, n. 2, p. 162–177, Feb 2005.

- SHUTLER, J. et al. On a large sequence-based human gait database. *Proc. 4th International Conference on Recent Advances in Soft Computing*, p. 66–71, 2002.
- SHUTLER, J.; NIXON, M.; HARRIS, C. Statistical gait recognition via temporal moments. *IEEE - Southwest Symposium on Image Analysis and Interpretation*, p. 291–295, 2000.
- STAUFFER, C.; GRIMSON, W. Adaptive background mixture models for real-time tracking. *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, p. 252–254, 1999.
- SUNDARESAN, A.; ROYCHOWDHURY, A.; CHELLAPPA, R. A hidden markov model based framework for recognition of human from gait sequences. *IEEE - International Conference Image Processing*, v. 3, p. II–93–96, Sep. 2003.
- WANG, L. et al. Gait recognition based on procrustes shape analysis. *IEEE International Conference on Image processing*, September 2002.
- WANG, L. et al. Silhouette analysis-based gait recognition for human identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 25, n. 12, p. 1505–1518, 2003.
- WINTER, D. *The Biomechanis and Motor Control of Human Gait*. 2nd.. ed. [S.l.]: Walterloo Press, 1991.
- YANG, J.; WU, X.; PENG, Z. Gait recognition based on difference motion slice. *Proc. of 8th International Conference on Signal Processing.*, v. 4, 2006.
- ZUEV, Y.; IVANOV, S. The voting as a way to increse the decision reliability. *Journal of Franklin Institute*, p. 361–378, 1999.

APÊNDICE A

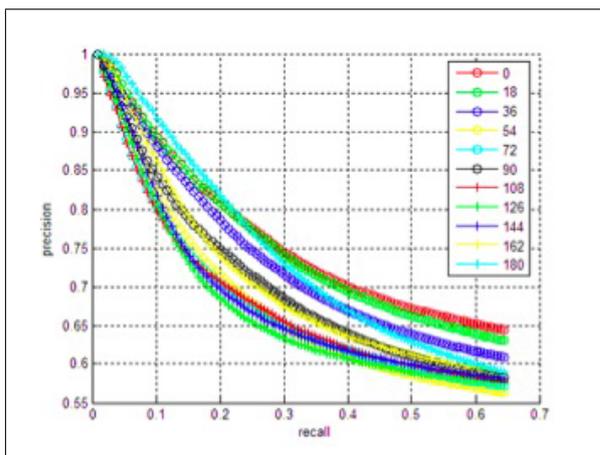


Figura 44: Curvas Recall x Precision SGW - Combinação de duas seqüências do tipo bg.

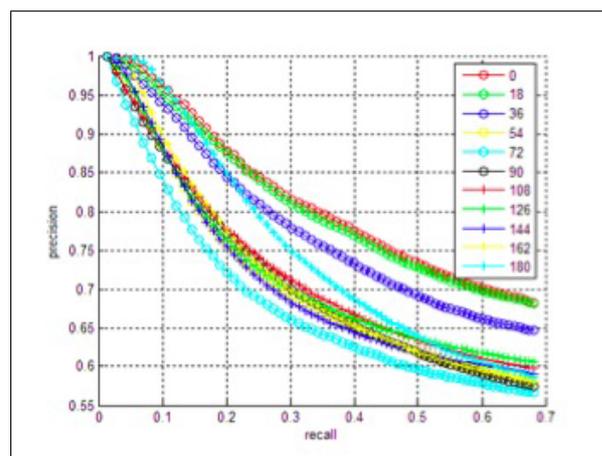


Figura 45: Curvas Recall x Precision SGW - Combinação de duas seqüências do tipo cl.

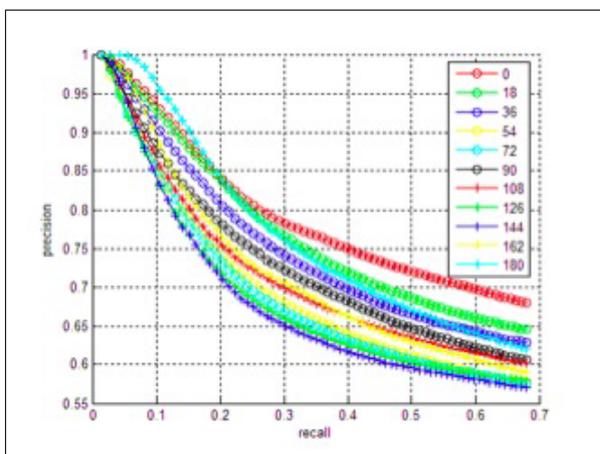


Figura 46: Curvas Recall x Precision SGW- Combinação de seis seqüências do tipo nm.

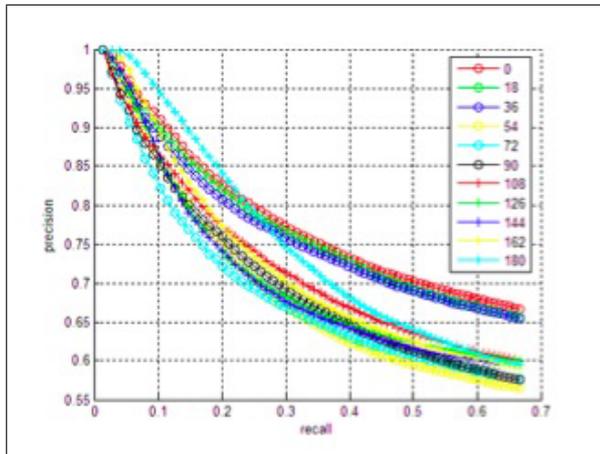


Figura 47: Curvas Recall x Precision SBW - Combinação de duas seqüências do tipo bg.

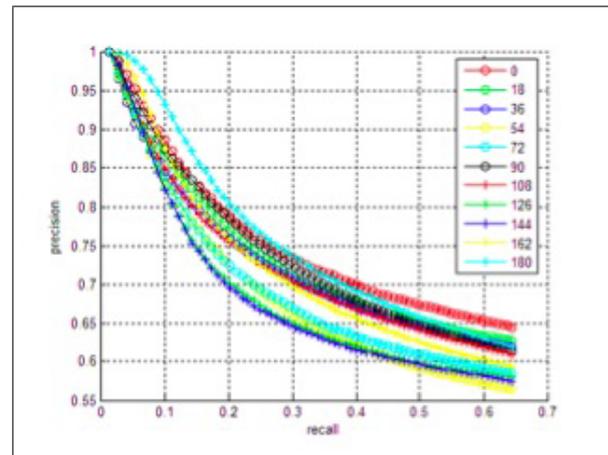


Figura 48: Curvas Recall x Precision SBW - Combinação de duas seqüências do tipo cl.

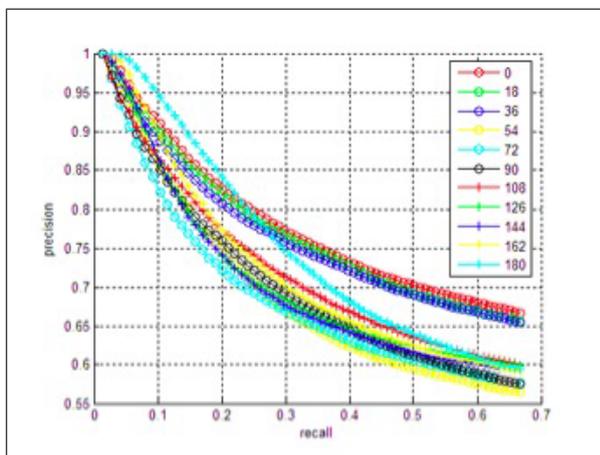


Figura 49: Curvas Recall x Precision SBW - Combinação de seis seqüências do tipo nm.

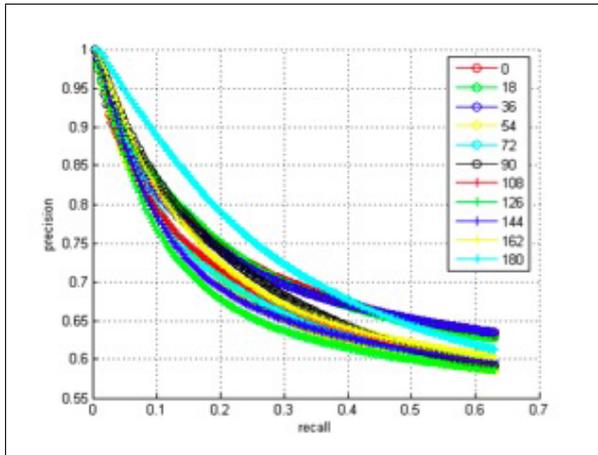


Figura 50: Curvas Recall x Precision SEW - Combinação de duas seqüências do tipo bg.

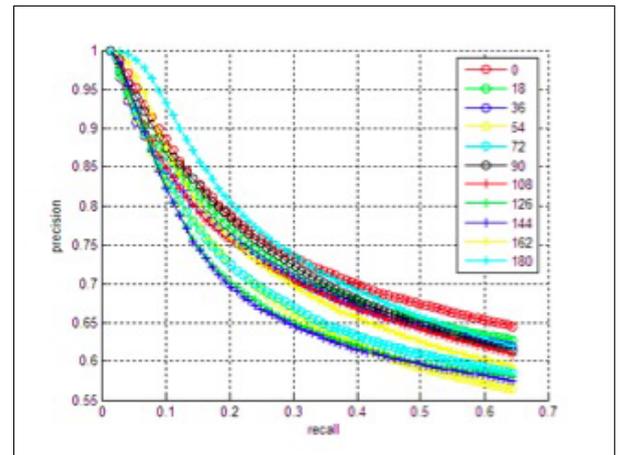


Figura 51: Curvas Recall x Precision SEW - Combinação de duas seqüências do tipo cl.

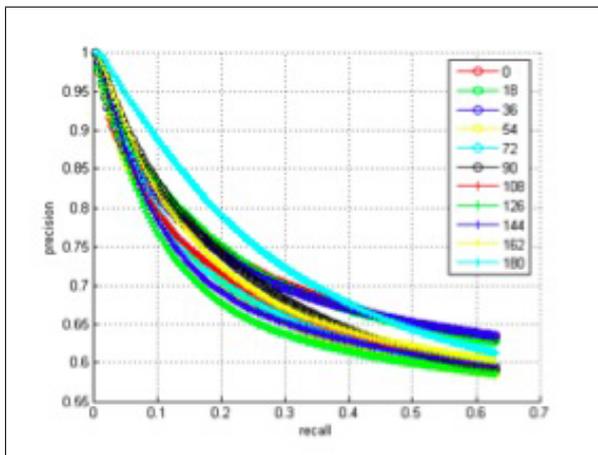


Figura 52: Curvas Recall x Precision SEW- Combinação de seis seqüências do tipo nm.

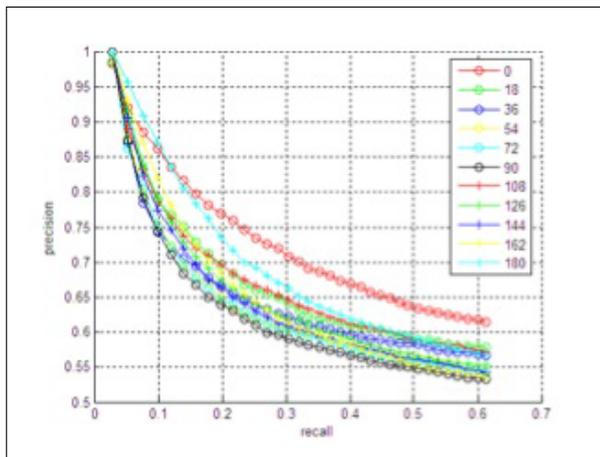


Figura 53: Curvas Recall x Precision SSW - Combinação de duas seqüências do tipo bg.

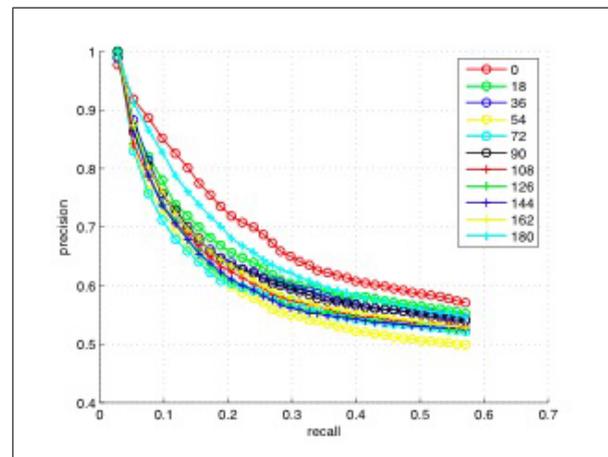


Figura 54: Curvas Recall x Precision SSW - Combinação de duas seqüências do tipo cl.

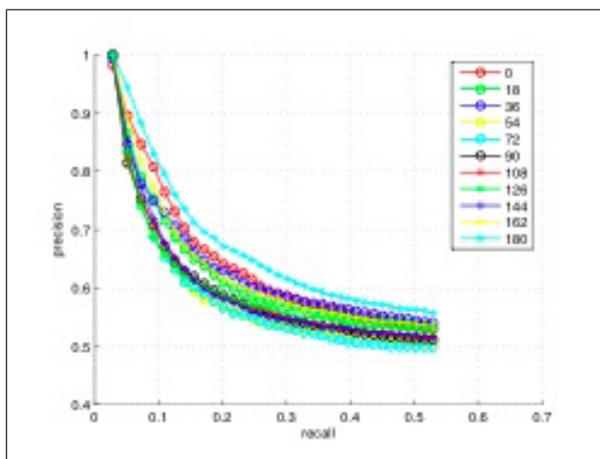


Figura 55: Curvas Recall x Precision SSW- Combinação de seis seqüências do tipo nm.

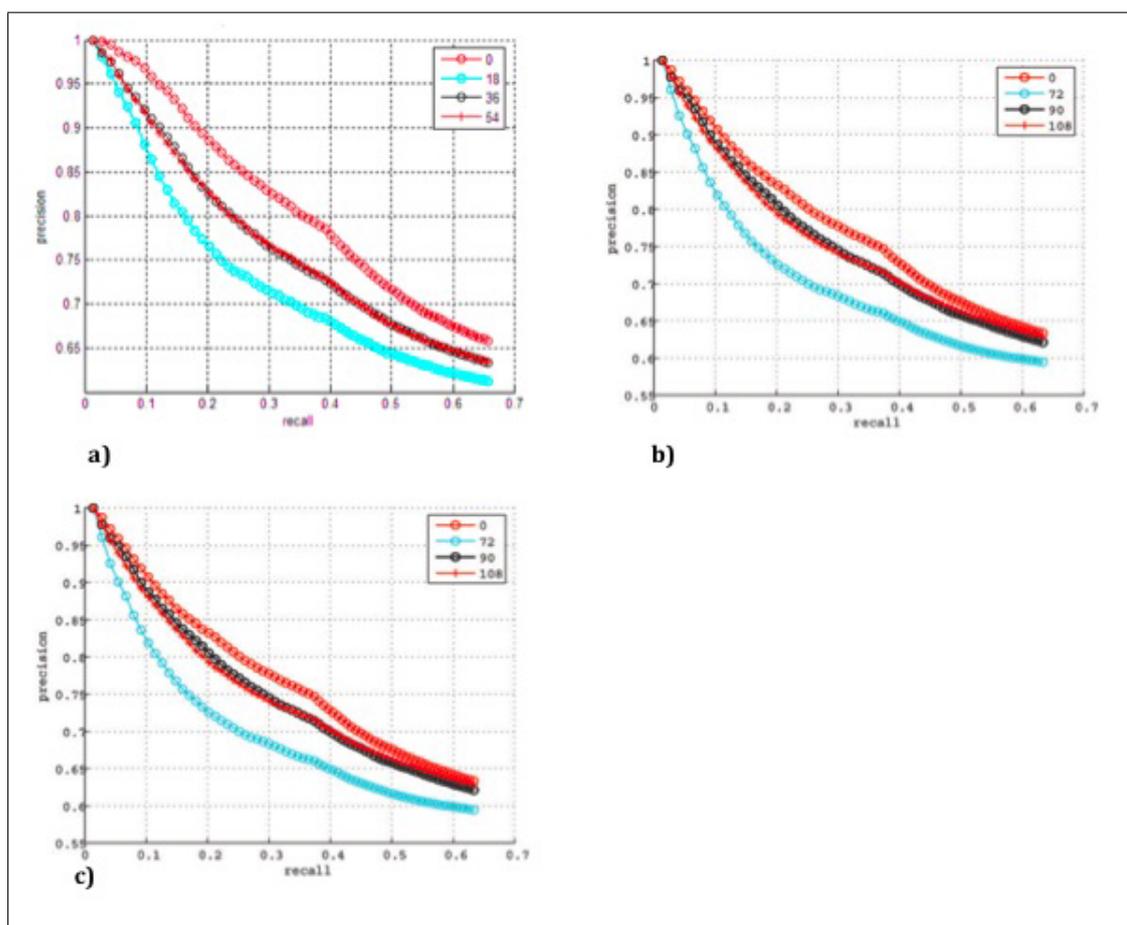


Figura 56: Curvas Recall x Precision SGW: (a) Combinação de seqüências bg com seqüências cl; (b) combinação de seqüências bg com seqüências nm; (c) combinação de seqüências cl com seqüências nm.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)