



**Universidade de Brasília**

Instituto de Ciências Exatas  
Departamento de Ciência da Computação

## Verificação de Texto Manuscrito em Dispositivos Móveis

Nilson Donizete Guerin Júnior

Dissertação apresentada como requisito parcial para  
conclusão do Mestrado em Informática

Orientador

Prof. Dr. Bruno Luigi Macchiavello

Coorientador

Prof. Dr. Flávio de Barros Vidal

Brasília  
2015

Universidade de Brasília — UnB  
Instituto de Ciências Exatas  
Departamento de Ciência da Computação  
Pós-graduação em Informática

Coordenadora: Prof.<sup>a</sup> Dr.<sup>a</sup> Alba Cristina Magalhães Alves de Melo

Banca examinadora composta por:

Prof. Dr. Bruno Luigi Macchiavello (Orientador) — CIC/UnB  
Prof. Dr. Ricardo Pezzuol Jacobi — CIC/UnB  
Prof.<sup>a</sup> Dr.<sup>a</sup> Carla Liberal Pagliari — IME

### **CIP — Catalogação Internacional na Publicação**

Guerin Júnior, Nilson Donizete.

Verificação de Texto Manuscrito em Dispositivos Móveis / Nilson Donizete Guerin Júnior. Brasília : UnB, 2015.

109 p. : il. ; 29,5 cm.

Dissertação (Mestrado) — Universidade de Brasília, Brasília, 2015.

1. biometria, 2. sistemas biométricos, 3. palavras cursivas isoladas,  
4. assinaturas, 5. histogramas, 6. descritores de Fourier.

CDU 004

Endereço: Universidade de Brasília  
Campus Universitário Darcy Ribeiro — Asa Norte  
CEP 70910-900  
Brasília-DF — Brasil

# NILSON DONIZETE GUERIN JUNIOR

## VERIFICAÇÃO DE TEXTO MANUSCRITO EM DISPOSITIVOS MÓVEIS

Dissertação aprovada como requisito parcial para obtenção do grau de Mestre no Curso de Pós-graduação em Informática da Universidade de Brasília, pela Comissão formada pelos professores:

Orientador:

  
Prof. Dr. Bruno Luigi Macchiavello Espinoza – orientador  
(CIC/UnB)

  
Prof. Dr. Ricardo Pezzuol Jacobi  
(CIC/UnB)

  
Prof. Dr. Carla Pagliari  
(IME/RJ)

Vista e permitida a impressão.  
Brasília, 10 de março de 2015.

Prof.<sup>a</sup> Dr.<sup>a</sup> Alba Cristina Magalhães Alves de Melo  
Programa de Pós-Graduação em Informática  
Departamento de Ciência da Computação  
Universidade de Brasília



# Resumo

O reconhecimento biométrico pode ser definido como a ciência de estabelecer a identidade de um indivíduo baseado nas características físicas e/ou comportamentais da pessoa seja de maneira totalmente automática seja de forma semiautomática. Um sistema biométrico tem dois tipos de funcionalidades: a verificação e a identificação. Na verificação, o usuário afirma sua identidade e o sistema verifica se a afirmação é genuína. A identificação diz respeito a apontar uma identidade, dentre todas as cadastradas no sistema, para uma amostra de entrada.

A biometria tem vantagens sobre abordagens baseadas em *tokens* ou conhecimento de informações (por exemplo, uma senha). Isto porque *tokens* podem ser perdidos e o conhecimento pode ser esquecido. O reconhecimento de escritor baseado em texto é uma das áreas mais ativas na biometria, dentre as biometrias comportamentais. Tanto as abordagens online quanto offline tem sido estudadas em trabalhos anteriores. Mas quando se refere ao tipo de texto usado, o reconhecimento de assinatura tem atraído muito mais atenção do que abordagens baseadas em outros tipos de escrita, especialmente no caso online. Além disso, o cenário com dispositivos móveis foi pouco abordado nesse contexto.

O objetivo deste trabalho é o desenvolvimento de um sistema de verificação multi-biométrica aplicável a palavras cursivas isoladas e assinaturas, em um cenário com dispositivos móveis. Uma base de dados, denominada LISA-01, foi especialmente desenvolvida para avaliar o cenário desejado. O sistema considera descritores de Fourier, histogramas e velocidade para classificação, tendo em vista baixa complexidade computacional. Foram desenvolvidas e avaliadas duas propostas para otimização no uso de histogramas.

Taxas de erro de 6,47% e 4,39% são obtidas nas bases de dados LISA-01 e MCYT-100, respectivamente. Os resultados sugerem que a combinação das características utilizadas é complementar e que o uso de palavras cursivas isoladas pode ser adequado em sistemas biométricos. Além disso, evidenciam que há relevância entre comprimento da palavra e a performance de verificação.

**Palavras-chave:** biometria, sistemas biométricos, palavras cursivas isoladas, assinaturas, histogramas, descritores de Fourier.

# Abstract

The biometric recognition can be defined as the science of establishing the identity of an individual based on physical and / or behavioral characteristics of the person in a fully or semiautomatic way. A biometric system has two types of tasks: verification and identification. In verification, the user claim his/her identity and the system checks whether the claim is genuine. Identification concerns about relating a sample to an identity among all registered in the system.

Biometrics has advantages over approaches based on tokens or knowledge (e.g. passwords) as tokens can be lost and knowledge can be forgotten. The recognition of writer based on text is one of the most active areas in biometrics, specifically in the behavioral biometrics. Both online and offline approaches have been studied in previous work. But when it comes to the type of text used, the signature recognition has attracted much more attention than approaches based on other types of writing, especially in the case online. In addition, the scenario with mobile devices was not much discussed in this context.

The objective of this work is the development of a multi-biometric verification system applicable to isolated cursive words and signatures, in a scenario with mobile devices. A database, called LISA-01, was specially developed to evaluate the desired scenario. The system considers Fourier descriptors, velocity and histograms for classification, obtaining low computational complexity. Proposals for optimizing the use of histograms is also developed and evaluated.

Error rates of 6.47% and 4.39% are obtained in the database LISA-01 and MCYT-100, respectively. The results suggest that the combination of the used characteristics is complementary and the introduction of isolated cursive words can be appropriate to biometric systems. Furthermore, the relevance of word length on verification performance is also identified.

**Keywords:** biometrics, biometrics systems, cursive isolated words, signatures, histograms, Fourier descriptors.

# Sumário

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Introdução</b>  | <b>1</b>  |
| 1.1      | Descrição do problema . . . . .  | 3         |
| 1.2      | Estado da arte . . . . .   | 3         |
| 1.3      | Objetivos Geral e Específicos . . . . .                                | 4         |
| 1.4      | Contribuições . . . . .  | 5         |
| 1.5      | Organização . . . . .  | 5         |
| <b>2</b> | <b>Revisão Bibliográfica</b>   | <b>6</b>  |
| 2.1      | Sistemas biométricos . . . . .   | 6         |
| 2.1.1    | Cadastramento e reconhecimento . . . . .                               | 7         |
| 2.1.2    | Tarefas do reconhecimento . . . . .                                    | 8         |
| 2.1.3    | Os traços biométricos e suas características . . . . .                 | 10        |
| 2.1.4    | Medidas de Performance para Autenticação . . . . .                     | 10        |
| 2.2      | Sistemas multibiométricos . . . . .                                    | 12        |
| 2.2.1    | Tipos de sistemas de múltiplas fontes . . . . .                        | 13        |
| 2.2.2    | Níveis de fusão . . . . .  | 14        |
| 2.3      | Biometrias fisiológicas . . . . .                                      | 16        |
| 2.3.1    | Impressão digital . . . . .  | 16        |
| 2.3.2    | Impressão da palma da mão . . . . .                                    | 16        |
| 2.3.3    | Iris . . . . .   | 17        |
| 2.3.4    | Face . . . . .   | 17        |
| 2.3.5    | Voz . . . . .  | 17        |
| 2.4      | Biometrias comportamentais . . . . .                                   | 18        |
| 2.4.1    | Reconhecimento automático de escritor . . . . .                        | 19        |
| <b>3</b> | <b>Fundamentação Teórica</b>   | <b>26</b> |
| 3.1      | Análise de Fourier . . . . .   | 26        |
| 3.1.1    | A Transformada de Fourier de Função de uma Variável Contínua . . . . . | 26        |
| 3.1.2    | DFT - A Transformada Discreta de Fourier . . . . .                     | 27        |

|          |   |           |
|----------|---|-----------|
| 3.1.3    | FFT - A Transformada de Fourier Rápida . . . . .                      | 28        |
| 3.1.4    | Usos e aplicações da Transformada de Fourier . . . . .                | 29        |
| 3.2      | Histogramas . . . . .   | 30        |
| 3.2.1    | Definição . . . . .   | 30        |
| 3.2.2    | Densidades de kernels . . . . .                                       | 33        |
| 3.2.3    | Histogramas no reconhecimento de padrões . . . . .                    | 35        |
| 3.3      | Métricas de Comparação . . . . .                                      | 36        |
| 3.3.1    | Distância Euclidiana . . . . .  | 36        |
| 3.3.2    | Distância de Manhattan . . . . .                                      | 37        |
| 3.3.3    | DTW - Técnica da Deformação Dinâmica do Tempo . . . . .               | 38        |
| <b>4</b> | <b>Metodologia</b>  | <b>42</b> |
| 4.1      | A estrutura geral do sistema . . . . .                                | 42        |
| 4.2      | A etapa de aquisição de dados . . . . .                               | 44        |
| 4.3      | Pré-processamento . . . . .   | 45        |
| 4.4      | Extração de características . . . . .                                 | 46        |
| 4.4.1    | Histogramas de Ângulos . . . . .                                      | 47        |
| 4.4.2    | Descritores de Fourier . . . . .                                      | 50        |
| 4.4.3    | Característica de Velocidade . . . . .                                | 54        |
| 4.5      | Composição do Perfil . . . . .  | 56        |
| 4.5.1    | Análise das Amostras do Perfil . . . . .                              | 57        |
| 4.6      | Classificador . . . . .   | 61        |
| 4.6.1    | Combinação das características . . . . .                              | 64        |
| 4.7      | Decisão . . . . .   | 65        |
| 4.8      | Análise da Complexidade Assintótica para o Sistema Proposto . . . . . | 67        |
| <b>5</b> | <b>Resultados Experimentais</b>                                       | <b>69</b> |
| 5.1      | Bases de dados . . . . .  | 69        |
| 5.1.1    | A base de dados LISA-01 . . . . .                                     | 69        |
| 5.1.2    | A base de dados MCYT-100 . . . . .                                    | 73        |
| 5.2      | Parâmetros adotados nos experimentos . . . . .                        | 73        |
| 5.2.1    | Composição do perfil e simulações . . . . .                           | 73        |
| 5.2.2    | Utilização das amostras das bases de dados . . . . .                  | 74        |
| 5.2.3    | As características . . . . .  | 75        |
| 5.2.4    | Parâmetros do classificador . . . . .                                 | 76        |
| 5.2.5    | Limiares e Avaliação de desempenho . . . . .                          | 77        |
| 5.3      | Experimento 1 - Análise dos Histogramas . . . . .                     | 78        |
| 5.3.1    | Número ótimo de <i>bins</i> dos histogramas . . . . .                 | 78        |

|          |   |           |
|----------|---|-----------|
| 5.3.2    | Avaliação do sistema com as técnicas de histogramas . . . . .         | 79        |
| 5.4      | Experimento 2 - Performance do sistema completo . . . . .             | 81        |
| 5.4.1    | Relevância do <i>padding</i> para os descritores de Fourier . . . . . | 81        |
| 5.4.2    | Comparação com o estado da arte . . . . .                             | 82        |
| 5.4.3    | Análise da estimação de <i>drift</i> na performance . . . . .         | 84        |
| 5.5      | Experimento 3 - Comparação da escrita por dedo e caneta . . . . .     | 86        |
| <b>6</b> | <b>Conclusão</b>  | <b>88</b> |
| 6.1      | Trabalhos Futuros . . . . .   | 90        |
|          | <b>Referências</b>  | <b>92</b> |

# Lista de Figuras

|     |   |    |
|-----|---|----|
| 2.1 | Exemplos de traços que tem sido usados para reconhecimento biométrico. Traços físicos incluem face, impressão digital, iris, impressão da mão, geometria da mão, voz e formato da orelha. Postura / modo de andar, assinaturas/escrita são algumas características comportamentais. Fonte: [1]. . . . . | 7  |
| 2.2 | Exemplo de distribuição dos <i>scores</i> . O limiar determina da FAR e FRR do sistema. Note que dadas essas distribuições, a FAR e a FRR não podem ser reduzidas simultaneamente ao ajustar o limiar. Fonte: [1]. . . . .  | 12 |
| 2.3 | Exemplos de assinaturas. (a) Assinatura offline, onde constam somente informações de luminância dos pixels. (b) Assinatura dinâmica, representada pelos pontos amostrados (“*”: pontos amostrados no toque; “•”: pontos amostrados no ar). Fonte: [2] . . . . .   | 22 |
| 2.4 | Modelo de classificação de características. Fonte: [2]. . . . .   | 24 |
| 3.1 | Diagonal das notas de banco falsificadas. Histogramas com $x_0 = 137.8$ e $h = 0.1$ (superior esquerdo), $h = 0.2$ (inferior esquerdo), $h = 0.3$ (superior direito) e $h = 0.2$ (inferior direito). Fonte: [3]. . . . .  | 31 |
| 3.2 | Diagonal das notas de banco falsificadas. Histogramas com $h = 0.4$ e $x_0 = 137.65$ (superior esquerdo), $x_0 = 137.75$ (inferior esquerdo), $x_0 = 137.75$ (superior direito) e $x_0 = 137.85$ (inferior direito). Fonte: [3]. . . . .  | 32 |
| 3.3 | Histogramas ponderados pela média dos histogramas deslocados baseado em todos os tipos de notas (falsificadas e genuínas): 2 deslocamentos (superior esquerdo), 4 deslocamentos (inferior esquerdo), 8 deslocamentos (superior direito) e 16 deslocamentos (inferior direito). Fonte: [3]. . . . .      | 33 |
| 3.4 | Representação gráfica de um histograma bidimensional. Fonte: [3]. . . . .   | 34 |
| 3.5 | Densidades das diagonais de notas do banco genuínas e falsas. Fonte: [3]. . . . .   | 35 |
| 3.6 | Distância Euclidiana de 2 vetores bidimensionais. . . . .   | 36 |
| 3.7 | Comparação entre a distância Euclidiana (em verde) com a distância de Manhattan (outras cores) para dois pontos arbitrários. . . . .  | 37 |
| 3.8 | Alinhamento de duas sequências no tempo. Pontos alinhados são indicados pelas setas duplas. Fonte: [4]. . . . .   | 38 |

|      |  |    |
|------|--|----|
| 3.9  | Matriz de custos da comparação de duas sequências reais X e Y, usando a distância de Manhattan como função de custo c. Regiões de baixo custo são indicadas pelas cores escuras e regiões de alto custo são indicadas por cores claras. Fonte: [4]. . . . .  | 39 |
| 3.10 | (a) Matriz de custo C e (b) matriz de custo acumulado D com um caminho de deformação ótimo $p^*$ (caminho em branco). Fonte: [4]. . . . .  | 40 |
| 4.1  | Diagrama de fluxo contendo as principais tarefas realizadas pelo sistema proposto. . . . .   | 43 |
| 4.2  | Ângulos de azimute e altitude de uma caneta em relação ao plano da superfície de escrita. Fonte: [5]. . . . .  | 44 |
| 4.3  | Processo de derivar um histograma de ângulos: (a) a assinatura original, (b) a sequência de vetores dos 60 primeiros pontos da assinatura e (c) a distribuição dos ângulos derivados do vetor em (b). Fonte: [6]. . . . .  | 48 |
| 4.4  | Diagrama de fluxo para o cálculo de um histograma de ângulos. . . . .  | 49 |
| 4.5  | Um caso é mostrado para ilustração, usando somente os valores da coordenada $y$ . Da esquerda para a direita: (a) Assinatura genuína, seu perfil de $y$ e seus descritores de Fourier; (b) Assinatura falsificada, seu perfil de $y$ e seus descritores de Fourier. Os descritores de Fourier das assinaturas genuínas e falsas (apresentadas como pontos) são sobrepostos, mostrando os valores mínimos e máximos dos descritores de Fourier das assinaturas de referência, enquanto a linha no meio denota a média das assinaturas de referência. Fonte: [7] . . . . . | 55 |
| 4.6  | Diagrama de fluxo para o cálculo da característica de velocidade. . . . .  | 57 |
| 4.7  | Diagrama de fluxo representando os passos de análise de perfil. Ao receber o bloco de características das M amostras, são expandidos três fluxos em paralelo: um para calcular a característica de velocidade, outro para fazer amostras em cada um dos níveis dos histogramas e um último para realizar as mesmas análises nos descritores de Fourier. . . . .  | 61 |
| 4.8  | Diagrama de fluxo para os passos adotados no classificador. . . . .  | 66 |
| 5.1  | Interface desenvolvida para o recolhimento dos dados da base LISA-01. . .  | 70 |
| 5.2  | Conjunto de 3 amostras de cada uma das palavras da LISA-01: “december”, “intelligence”, “love” e “pattern”. Cada grupo de palavras pertence a um usuário distinto. . . . .   | 71 |

|      |  |    |
|------|--|----|
| 5.3  | Exemplos de amostras de 2 usuários na base de dados MCYT-100. Para cada usuário, as duas amostras mais à esquerda são assinaturas verdadeiras, enquanto que a amostra mais à direita é uma falsificação proficiente. Em cada caso, os <i>plots</i> gráficos apresentados correspondem às informações dinâmicas armazenadas na base de dados. Fonte: [7]. . . . .       | 74 |
| 5.4  | Exemplos de amostras de um usuário e possíveis tentativas de falsificação: (a) três amostras da palavra “dezembro” de um usuário, (b) conjunto de amostras da palavra “dezembro” de usuários diferentes (c) conjunto de amostras de palavras e usuários diferentes. . . . .  | 76 |
| 5.5  | A distribuição dos números de <i>bins</i> que melhor discriminam cada usuário do sistema, na base de dados MCYT-100: (a) análise feita com a distância euclidiana, (b) análise feita com a distância de Manhattan . . . . .  | 79 |
| 5.6  | EER para determinado número de <i>bins</i> para a execução do sistema na base de dados MCYT-100: (a) simulação com distância euclidiana, (b) simulação com distância de Manhattan. . . . .   | 80 |
| 5.7  | Gráficos da influência do tamanho para <i>padding</i> a EER do sistema, onde as linhas pontilhadas representam a abordagem cascata e as outras representam a abordagem completa: (a) resultados encontrados na MCYT-100, (b) resultados obtidos na LISA-100. . . . .   | 82 |
| 5.8  | Gráfico representando as variações na FAR e FRR em termos do limiar para a base de dados MCYT-100. O ponto de interseção representa a EER: (a) abordagem cascata, (b) abordagem completa. . . . .  | 83 |
| 5.9  | Gráfico representando as variações na FAR e FRR em termos do limiar para a base de dados LISA-01. O ponto de interseção representa a EER. (1) e (2) representam o caso onde as amostras são de palavras diferentes e da mesma palavra, respectivamente. (a), (b) e (c) representam a proposta de Fourier, a abordagem cascata e a abordagem completa, respectivamente. | 85 |
| 5.10 | Gráfico representando as variações na FAR e FRR em termos do limiar para a base de comparações. O ponto de interseção representa a EER: (a) entrada utilizando a caneta, (b) entrada utilizando o dedo . . . . .   | 87 |

# Lista de Tabelas

|     |   |    |
|-----|---|----|
| 3.1 | Funções de kernel mais comuns. . . . .  | 35 |
| 5.1 | Análise das três estratégias com o uso de histogramas como características, considerando a distância Euclidiana e de Manhattan como métricas de comparação, na MCYT-100, para o conjuntos de bins $\{6, 8, 10, 12, 14, 16\}$ , tanto para a abordagem completa quanto para a abordagem cascata. . . . | 80 |
| 5.2 | Análise das estratégias com o uso de histogramas como características, utilizando a distância de Manhattan, para ambas as bases de dados. $[b^i; p; b^f]$ representa a faixa de <i>bins</i> utilizada, iniciando em $b^i$ e finalizando em $b^f$ , com passo de $p$ nos valores. . . . .              | 81 |
| 5.3 | Abordagens de verificação automática de assinatura na base de dados MCYT-100, onde $x, y, t, p, az$ representam as coordenadas em $x$ , em $y$ , o tempo, a pressão e o azimute, respectivamente. Além destas características, estava disponível a altitude da caneta. . . . .                        | 83 |
| 5.4 | Resultados das simulações para a base de dados LISA-01: palavra 1 - “december”, palavra 2 - “intelligence”, palavra 3 - “love” e palavra 4 - “pattern”. . . . .   | 84 |
| 5.5 | Análise do impacto da estimação de <i>drift</i> para a eficiência do Sistema na MCYT-100 e LISA-01. Para a LISA-01, as falsificações são compostas de amostras da mesma palavra. . . . .  | 86 |
| 5.6 | Resultados da comparação de amostras escritas com o dedo e com a caneta, utilizando ambas as propostas descritas neste trabalho. . . . .  | 87 |

# Capítulo 1

## Introdução

A habilidade de identificar pessoas de forma única e associar atributos pessoais com um indivíduo tem sido crucial para a sociedade. Humanos tipicamente usam características do corpo e juntam com outras informações contextuais para se reconhecerem. A explosão do crescimento populacional, acompanhada da mobilidade crescente nas sociedades modernas impactou no desenvolvimento de sistemas de gerenciamento de identidade sofisticados que podem eficientemente armazenar, manter e obliterar identidades dos indivíduos. Esse gerenciamento é crítico em aplicações como regulamentações de transição de pessoas entre países, restrição de acesso físico, controle de acesso a recursos, etc [1].

A tarefa fundamental no gerenciamento de identidade é estabelecer a associação entre um indivíduo e sua identidade pessoal. Alguém deve ser capaz de determinar a identidade de uma pessoa ou verificar a afirmação de identidade de um indivíduo sempre que necessário. Esse processo é conhecido como reconhecimento de pessoas. Uma pessoa pode ser reconhecida baseado em três principais métodos:

1. O que sabe;
2. O que possui extrinsecamente;
3. Quem é intrinsecamente.

Enquanto o primeiro item recai sobre o fato de que o indivíduo tem conhecimento exclusivo de alguma informação secreta - por exemplo, senha, número de identificação pessoal ou chave criptográfica -, o segundo método assume que a pessoa possui um *token* - por exemplo, cartão de identificação, licença de motorista, passaporte, chave física ou um dispositivo pessoal como o celular. O terceiro método estabelece a identidade de uma pessoa baseado em seus traços físicos ou comportamentais inerentes e é conhecido como reconhecimento biométrico [1, 8].

Formalmente, o reconhecimento biométrico pode ser definido como a ciência de estabelecer a identidade de um indivíduo baseado nas características físicas e/ou comporta-

mentais da pessoa seja de maneira totalmente automática seja de forma semiautomática [9, 10]. O reconhecimento de pessoas baseado em conhecimento ou em *tokens* se suporta pela representação da identidade por senhas, cartões de identificação, os quais podem ser facilmente esquecidos/perdidos, adivinhados/roubados ou compartilhados. Além disso, eles não proveem funções vitais do gerenciamento de identidade como por exemplo a não-repudição, a detecção de múltiplos registros da mesma pessoa sob diferentes identidades e o reconhecimento de pessoas pós-evento, tal como identificação de suspeitos na cena de um crime [2, 11, 12, 13].

Um sistema biométrico tem dois tipos de funcionalidades: a verificação e a identificação. O termo reconhecimento, assim, faz referência a qualquer uma delas. Na verificação, o usuário afirma sua identidade e o sistema verifica se a afirmação é genuína. Nesse cenário, a amostra de entrada é comparada somente com o perfil correspondente à identidade afirmada - uma comparação de um para um [5, 14].

A identificação, por sua vez, pode ser classificada em positiva ou negativa. Na identificação positiva, o usuário tenta se identificar para o sistema sem explicitamente afirmar uma identidade. Nesse cenário, o sistema responde a questão “Você é alguém conhecido para o sistema?”. Em contraste, na identificação negativa considera-se que o usuário está mentindo sobre sua identidade, seja explicitamente ou implicitamente, mas que esteja incluído dentre os usuários do sistema. Neste último caso, o objetivo do sistema é descobrir “Você é quem você diz não ser?”. Dessa maneira, a identificação representa uma comparação de um para muitos [1].

Há duas grandes classes de biometrias: fisiológicas e comportamentais. As primeiras são baseadas na mensuração direta e partes do corpo humano, sendo a impressão digital e escâner da íris dois bons exemplos dessa classe. O segundo grupo refere-se à mensuração de dados derivados de ações desempenhadas pelo usuário e, portanto, medem as características do corpo humano de maneira indireta. Essa classificação, no entanto, é bastante artificial. Por exemplo, o sinal de voz depende de traços comportamentais tais quais a semântica, dicção, pronúncia, idiosincrasia, que, por sua vez, são relacionados ao status sócio-econômico, educação, local de nascimento, entre outros. Todavia, também depende da fisiologia do locutor, como por exemplo o formato do trato vocal [5, 14].

Cada traço biométrico tem seus prós e contras e, portanto, a escolha de um traço biométrico para um aplicação em particular depende de uma variedade de pontos a respeito de sua performance. Sete fatores são considerados nesse contexto: universalidade, unicidade, permanência, mensurabilidade, aceitabilidade e circumvenção. Nem uma única biometria eficientemente cumpre todos os requisitos impostos por todas as aplicações. Em outras palavras, nenhum traço biométrico é ideal, e sua relevância é estabelecida em relação à natureza e requerimentos da aplicação [11, 12, 1].

## 1.1 Descrição do problema

O reconhecimento de escritor baseado em texto é uma das áreas mais ativas na biometria, especificamente no que se refere às biometrias comportamentais. Quando dados espaço-temporais estão disponíveis (por exemplo, coordenadas espaciais, pressão, ângulos), o reconhecimento é dito *online*. Quando apenas a imagem escaneada da escrita está disponível, é dito *offline* [2, 15].

Tanto as abordagens *online* quanto *offline* tem sido estudadas em trabalhos anteriores. Mas quando se refere ao tipo de texto usado, o reconhecimento de assinatura [7, 6, 16, 17, 18] tem atraído muito mais atenção do que abordagens baseadas em outros tipos de escrita [14, 5], especialmente no caso *online*. Isso não deve causar surpresa visto que a assinatura tem longa tradição como um método para provar a identidade de uma pessoa. Para o campo *offline*, abordagens não baseadas em assinaturas [19, 20] tem recebido alguma atenção devido à sua relevância na investigação forense, principalmente. Mas no contexto *online*, abordagens não baseadas em assinaturas são limitadas e quando o olhar é específico para palavras avulsas ou pequenas sentenças, o número de trabalhos na literatura é realmente pequeno [5, 14].

Quando comparadas com assinaturas, palavras isoladas parecem carregar menos poder discriminativo. As assinaturas de duas pessoas tendem a ser bastante diferentes, enquanto que duas execuções de uma mesma palavra tendem a ser bem mais parecidas. Ou seja, assinaturas não colocam restrições para as diferenças de cada escritor enquanto que palavras são restringidas por dois fatores fundamentais: duas execuções de uma palavra que ser parecidas para serem consideradas a mesma palavras; elas tem que ser legíveis [14].

Os textos, todavia, tem uma propriedade singular, que as assinaturas por si só não tem: cancelabilidade. O usuário de um sistema biométrico não pode ser solicitado a escrever uma assinatura mais longa ou diferente, enquanto que tem disponibilidade de escrever um texto longo ou diferente. Dessa forma, um texto pode ser cancelado ou substituído por um novo, enquanto o mesmo não acontece com a assinatura. Portanto, sob esse ponto de vista, em alguns cenários de segurança o texto *online* manuscrito pode ser mais seguro do que a assinatura *online*. E mesmo que não fosse, a escrita cursiva que não seja assinatura é um traço biométrico que poderia complementar outros mecanismos de reconhecimento [5].

## 1.2 Estado da arte

Uma ampla variedade de técnicas foi aplicada para o reconhecimento de escritor. No cenário de assinaturas, é possível encontrar técnicas utilizado a Transformada de Fourier

[7, 21, 22], a Transformada de Cossenos [23, 24, 25], Redes Neurais [26], Cadeias de Markov [27] e Técnica da Deformação Dinâmica do Tempo (DTW, do inglês, *Dynamic Time Warping*) [28]. No que concerne a métricas de comparação, além da DTW, são vistos trabalhos utilizando distância euclidiana [7, 23] e distância de Manhattan [6].

Quando a visão é para o uso de palavras isoladas, é possível encontrar trabalhos recentes utilizando Mapas Auto-Organizáveis (SOM, do inglês, *Self Organizing Map*) e DTW para palavras em letra de forma [14, 5]. Além da pouca atenção a palavras isoladas, pouca ênfase foi dada a cenários em dispositivos móveis, cujo uso tem crescido cada vez mais, fazendo-os despontar como uma alternativa para reconhecimento biométrico [6]. O uso de histogramas também não foi muito estudado no reconhecimento *online* [29, 6].

### 1.3 Objetivos Geral e Específicos

Considerando todos os aspectos apresentados, a biometria é um campo com bastante relevância para o reconhecimento de identidade. Dessa forma, o objetivo geral do trabalho é o desenvolvimento de um sistema de verificação multibiométrica de palavras cursivas isoladas e assinaturas, que seja aplicável a dispositivos móveis.

Os objetivos específicos desse trabalho incluem:

- Criar uma base de dados para palavras cursivas isoladas usado a ponta do dedo como método de entrada e um *tablet* comum, visto que não há nenhuma desenvolvida;
- Desenvolver um sistema com baixa complexidade computacional, visto que este é um ponto fundamental para uso em dispositivos móveis;
- Avaliar metodologias no uso de histogramas como característica de sistemas de reconhecimento *online* de texto;
- Avaliar a combinação de histogramas com descritores de Fourier e velocidade como proposta para uma abordagem multibiométrica;
- Verificar a viabilidade do uso de palavras cursivas isoladas em termos de taxas de erro do sistema;
- Verificar o impacto do comprimento das palavras cursivas isoladas na performance de verificação;
- Criar uma pequena base de dados com usuários utilizando caneta e dedo para realizar análise do impacto do método de entrada na performance.

## 1.4 Contribuições

O sistema proposto neste trabalho é um tipo de sistema multibiométrico baseado em histogramas, descritores de *Fourier* e informação global de velocidade. Diferente de [5], que adota o uso de palavras isoladas em letra de forma, e de vários outros trabalhos encontrados na literatura, o objeto de estudo deste trabalho são as palavras cursivas isoladas (além das assinaturas), que, no melhor de nosso conhecimento, nunca foram avaliadas nos cenários considerados. Além disso, novas metodologias para o uso de histogramas como características de sistemas de reconhecimento de escritor *online* são apresentadas.

Os cenários comumente analisados e base de dados usualmente adotadas, consideram a utilização de mesas digitalizadoras com dispositivos especiais para captura de informações [6]. Assim, é possível obter informações como altitude e azimute da posição de escrita com a caneta. Diferentemente, aqui a avaliação das palavras cursivas foi feita com o uso de *tablets* comuns e sem o uso de dispositivos adicionais. Por fim, uma comparação do impacto do método de entrada para as palavras cursivas isoladas é apresentada.

A arquitetura, análises e resultados evidenciados neste trabalho serão publicados e apresentados na VISAPP 2015, 10<sup>a</sup> Conferência Internacional em Teoria de Visão Computacional e Aplicações (do inglês, *International Conference on Computer Vision Theory and Applications*) [30].

## 1.5 Organização

Este trabalho está organizado como segue:

- Capítulo 2 - aborda os detalhes e conceitos a respeito da área de biometria, assim como referências a trabalhos da área;
- Capítulo 3 - apresenta definições de algumas ferramentas envolvidas na proposta, métricas de comparação, análise na frequência e histogramas;
- Capítulo 4 - mostra os detalhes metodológicos pertinentes ao trabalho apresentado, a coleta de dados, pré-processamento, extração de características, caracterização dos perfis, classificação das amostras e a decisão do sistema;
- Capítulo 5 - apresenta os principais experimentos realizados, além dos resultados obtidos. Também são mostrados detalhes da composição da base de palavras cursivas isoladas.
- Capítulo 6 - apresenta as conclusões do trabalho apresentado.

# Capítulo 2

## Revisão Bibliográfica

Este capítulo abordará os detalhes e conceitos a respeito da área de biometria em geral. Serão explicitadas características gerais de biometrias, sendo apresentado, em seguida, informações a respeito das biometrias fisiológicas. Por último, serão abordadas as biometrias comportamentais, o foco do trabalho aqui apresentado. Nesta seção será focado o reconhecimento automático de escritor, com ênfase nas abordagens online e offline.

### 2.1 Sistemas biométricos

A extração automática de identidade a partir de traços pessoais deu nascimento a um ramo particular do reconhecimento de padrões, que tem sido alvo de várias pesquisas, a biometria [29]. O termo biometria refere-se ao reconhecimento individual baseado em características discriminantes de uma pessoa [9, 10]. Enquanto outras técnicas usam um *token* (como um crachá, cartão de identificação, etc.) ou conhecimento *a priori* (uma senha, palavra chave, etc.) para realizar o reconhecimento pessoal [8], técnicas biométricas oferecem o potencial de usar características inerentes da pessoa a ser reconhecida para desempenhar essa tarefa. Dessa forma, os atributos biométricos não sofrem das desvantagens de abordagens baseadas em *tokens*, cujos atributos podem ser perdidos ou roubados e de abordagens com conhecido, cujos atributos podem ser esquecidos [2, 11, 12, 13].

Assim, um sistema biométrico é aquele que mede uma ou mais características físicas e/ou comportamentais, incluindo impressão digital, impressão da mão, face, iris, retina, ouvido, voz, assinatura, DNA, etc. de um indivíduo para determinar ou verificar sua identidade. Essas características são relacionadas com diferentes termos tais quais traços, indicadores, identificadores ou modalidades [1]. A Figura 2.1 mostra alguns tipos de traços comumente usados para reconhecimento.

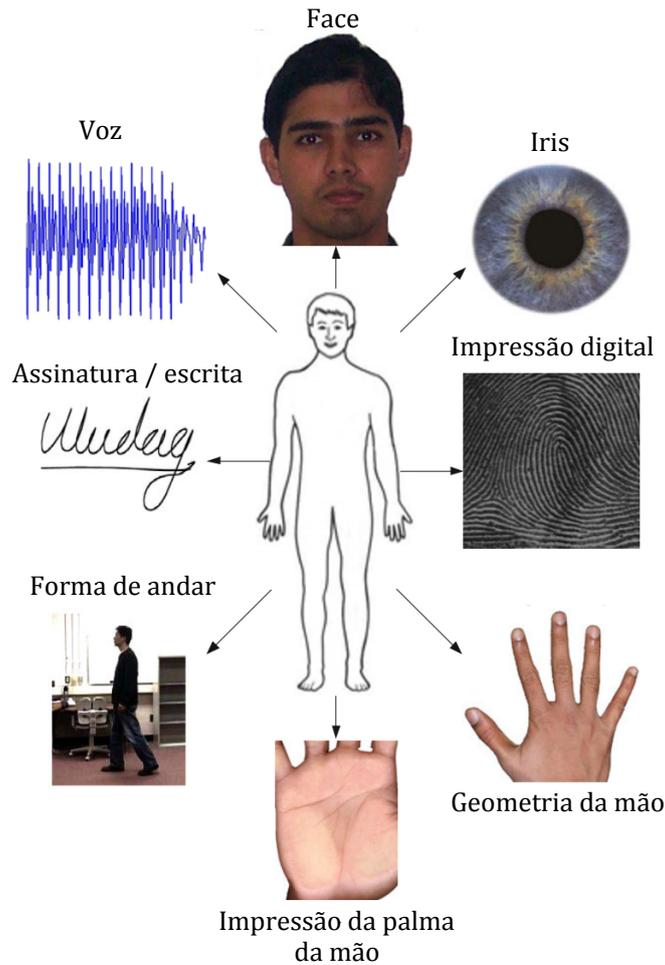


Figura 2.1: Exemplos de traços que tem sido usados para reconhecimento biométrico. Traços físicos incluem face, impressão digital, iris, impressão da mão, geometria da mão, voz e formato da orelha. Postura / modo de andar, assinaturas/escrita são algumas características comportamentais. Fonte: [1].

### 2.1.1 Cadastramento e reconhecimento

Um sistema biométrico identifica/verifica um usuário baseado em seus traços físicos/-comportamentais a partir de duas tarefas, nomeadamente, cadastramento e reconhecimento [12]. Durante a fase de cadastramento, os dados biométricos são adquiridos do indivíduo e armazenados em uma base de dados juntamente à identidade da pessoa. Tipicamente, os dados biométricos adquiridos são processados para extrair características (ou no inglês, *features*) distintivas e salientes. Em vários casos, somente as características extraídas são armazenadas, enquanto os dados biométricos brutos são descartados.

Na fase de reconhecimento, os dados biométricos são readquiridos do indivíduo e comparados com os armazenados para determinar a identidade do usuário. Assim, um sistema biométrico é essencialmente um sistema de reconhecimento de padrões, constituído

de quatro blocos básicos de construção, apresentados abaixo [1, 31, 12, 10]:

1. **Módulo do sensor:** uma interface que incorpore um sensor ou leitor biométrico é necessária para medir ou armazenar os dados biométricos brutos do usuário. Por exemplo, um sensor ótico de impressões digitais pode ser usado para construir a imagem a partir do padrão de fricção da ponta do dedo.
2. **Módulo de extração de características:** usualmente, os dados biométricos brutos são submetidos a um pré-processamento antes que as características sejam extraídas. Alguns exemplos desses processamentos são melhoria da qualidade, segmentação e realce. A extração de características propriamente dita refere-se ao processo de geração de uma representação compacta mas expressiva do traço biométrico em questão, chamado de *template* / perfil. Espera-se que o perfil tenha apenas a informação discriminatória essencial para o reconhecimento.
3. **Módulo da base de dados:** o sistema biométrico age como um repositório de informação biométrica. Durante o processo de cadastramento, o conjunto de características extraídas é armazenado na base de dados junto com outras informações pessoais.
4. **Módulo de comparação:** o propósito de um comparador é realizar a comparação das características da amostra de entrada em relação aos perfis armazenados para gerar uma pontuação/medida de comparação. Essa medida de comparação é um referencial da similaridade entre o perfil e a amostra de entrada.

### 2.1.2 Tarefas do reconhecimento

A etapa de reconhecimento em um sistema biométrico pode ter como objetivo a identificação ou a verificação de usuários. Na verificação, também conhecida como autenticação, o usuário afirma sua identidade e o sistema verifica se a afirmação é genuína. Nesse cenário, a amostra de entrada é comparadamente somente com o perfil correspondente à identidade afirmada, representando uma comparação de um para um [14, 5]. A afirmação de identidade é usualmente feita através de um número de identificação pessoal, um nome de usuário, ou um *token*. Se a entrada do usuário e o perfil da identidade provida tem um alto grau de similaridade, então a afirmação é tida como verdadeira. Caso contrário, a afirmação é rejeitada e o usuário é considerado um “impostor”. Na literatura da biometria, os termos “cliente” ou “autêntico” são algumas vezes usados no lugar do termo “genuíno”. A verificação é tipicamente utilizada em aplicações onde a meta é prevenir pessoas não autorizadas de usar os serviços [1].

Formalmente, a verificação pode ser colocada como o seguinte problema de classificação de duas categorias: dada uma identidade  $I$  e um conjunto de características de entrada  $x^A$ , é necessário decidir se  $(I, x^A)$  pertence à classe “genuína” ou “impostora”. Seja  $x_I^E$  o perfil armazenado correspondente à identidade  $I$ . Tipicamente,  $x^A$  é comparado com  $x_I^E$  e um *score*  $s$ , que mede a similaridade entre  $x^A$  e  $x_I^E$ , é calculado. A regra de decisão é dada por:

$$(I, x^A) \in \begin{cases} \text{genuína, se } s \geq \eta, \\ \text{falsa, se } s < \eta, \end{cases} \quad (2.1)$$

onde  $\eta$  é um limiar pré-definido. Se um *score* de dissimilaridade é usado ao invés do de similaridade, então as desigualdades da equação são invertidas.

Identificação, a outra tarefa de reconhecimento, pode ser classificada em identificação positiva e negativa. Na identificação positiva, o usuário tenta se identificar para o sistema sem explicitamente afirmar uma identidade. Nesse cenário, o sistema responde a questão “Você é alguém conhecido para o sistema?”. Em contraste, na identificação negativa considera-se que o usuário está mentindo sobre sua identidade, seja explicitamente ou implicitamente. Neste último caso, o objetivo do sistema é descobrir “Você é quem você diz não ser?” [1].

O propósito da identificação negativa é prevenir de uma única pessoa usar múltiplas identidades. Em ambas as identificações, a entrada biométrica do usuário é comparada com os perfis de todas as pessoas cadastradas na base de dados e o sistema retorna ou a identidade da pessoa cujo perfil tem o maior grau de similaridade ou uma decisão indicando se o usuário apresentado a entrada não está cadastrado no sistema.

Formalmente, o problema de identificação pode ser descrito da seguinte maneira: dado um conjunto de características de entrada  $x^A$ , é necessário decidir a identidade  $I$  do usuário, com  $I \in \{I_1, I_S, \dots, I_N, I_{N+1}\}$ . Aqui,  $I_1, I_2, \dots, I_N$  corresponde às identidades dos  $N$  usuários cadastrados no sistema e  $I_{N+1}$  indica o caso onde nenhuma identidade viável pode ser determinada, dada a entrada. Se  $x_{I_n}^E$  é o perfil armazenado correspondente à identidade  $I_n$  e  $s_n$  é o *score* entre  $x^A$  e  $x_{I_n}^E$ , para  $n = 1, 2, \dots, N$ , a regra de decisão é:

$$\begin{cases} I_{n_0}, \text{ se } n_0 = \operatorname{argmax} s_n \text{ e } s_{n_0} \geq \eta, \\ I_{N+1}, \text{ caso contrário,} \end{cases} \quad (2.2)$$

onde  $\eta$  é um limiar pré-definido. Também é possível forçar o sistema a retornar um dentre as  $N$  identidades cadastradas, indiferente ao valor de  $s_{n_0}$ .

Em alguns sistemas a identificação é semiautomática. Um sistema biométrico semiautomático retorna as identidades das  $t$  melhores comparações ( $1 < t \ll N$ ) e um especialista humano determina a identidade dentro desse grupo.

### 2.1.3 Os traços biométricos e suas características

Um grande número de traços biométricos estão sendo usados em várias aplicações. Cada traço biométrico tem seus prós e contras e, portanto, a escolha de um traço biométrico para uma aplicação particular depende de uma variedade de assuntos a respeito de sua performance. Em geral, sete fatores devem ser considerados para determinar a conveniência do traço físico/comportamental para uma aplicação biométrica [1, 11, 12]:

1. **Universalidade:** cada indivíduo acessando a aplicação deve possuir o traço. Esse fator determina a taxa de erro para cadastramento do sistema.
2. **Unicidade:** o traço dado deve ser suficientemente diferente entre os indivíduos da população de usuários. Caso contrário, a taxa de aceitação falsa do sistema seria inaceitavelmente alta.
3. **Permanência:** o traço biométrico de um indivíduo deve ser suficientemente invariante sob um período de tempo com respeito ao algoritmo de comparação. Um traço que muda significativamente sob o tempo não é útil como um biometria porque irá levar a uma taxa de rejeição falsa alta.
4. **Mensurabilidade:** deve ser possível adquirir e digitalizar o traço biométrico usando dispositivos viáveis que não causem inconveniência ao indivíduo. Além disso, os dados brutos devem ser passíveis de processamento a fim de extrair conjuntos de características discriminativas.
5. **Aceitabilidade:** os indivíduos na população que irá utilizar a aplicação devem concernir a respeito de apresentar seus traços biométricos ao sistema.
6. **Circumvenção:** se refere à facilidade com a qual o traço de um indivíduo pode ser imitado usando artefatos, no caso de traços físicos, e repetição, no caso de traços comportamentais.

Nem uma única biometria eficientemente cumpre todos os requisitos impostos por todas as aplicações - por exemplo forense, controle de acesso, programas de benefício do governo, etc. Em outras palavras, nenhum traço biométrico é ideal, mas um número deles é admissível. A relevância de uma biometria específica para uma aplicação é estabelecida dependendo da natureza e requerimentos da aplicação [11, 12].

### 2.1.4 Medidas de Performance para Autenticação

As medidas básicas de eficácia de uma sistema biométrico são a Taxa de Não-alinhamento Falso (FNRM, do inglês, *False Non-Match Rate*) e a Taxa de Alinhamento Falso (FMR,

do inglês, *False Match Rate*). FNMR refere-se à probabilidade de que duas amostras irmãs (obtidas do mesmo traço biométrico do mesmo usuário) serão falsamente declaradas desalinhadas. FMR é a probabilidade de que duas amostras não-irmãs serão incorretamente alinhadas [1, 31, 32].

No contexto de verificação biométrica, a FNMR e FMR são geralmente referidas como Taxa de Rejeição Falsa (FRR, do inglês, *False Rejection Rate*) e Taxa de Aceitação Falsa (FAR, do inglês, *False Acceptance Rate*), respectivamente. Formalmente:

$$FAR(\eta) = p(s \geq \eta | \omega_0) = \int_{\eta}^{\infty} p(s | \omega_0) ds, \quad (2.3)$$

$$FRR(\eta) = p(s < \eta | \omega_1) = \int_{-\infty}^{\eta} p(s | \omega_1) ds, \quad (2.4)$$

onde  $p(s | \omega_1)$  e  $p(s | \omega_0)$  são as funções de probabilidades de scores genuínos e impostores, respectivamente. Tanto a FAR, quanto a FRR são funções do limiar  $\eta$ . Se o limiar cresce, FAR irá decrescer mas FRR irá crescer e vice versa. Dessa forma, dado um sistema biométrico, é impossível decrescer ambas as taxas de erro simultaneamente [15]. Decorrente da FAR e FRR há, ainda, uma terceira medição, a Taxa de Erro Igual (EER, do inglês, *Equal Error Rate*). A EER é definida como a taxa de erro do sistema quando  $FAR = FRR$  [2, 10]. É um valor amplamente considerado para medir a eficácia de um sistema biométrico, sendo visto em vários trabalhos [33, 28, 34, 25, 14, 5, 7, 35]. A Figura 2.2 mostra um exemplo de gráfico da distribuição de probabilidade das classes genuínos e impostores em função dos *scores*.

Os cálculos discretos para a FAR e FRR decorrem naturalmente da definição contínua. Considere um conjunto de *scores*  $\{s_i\}_{i=1}^L$ , onde os primeiros  $L_1$  *scores* correspondem à classe genuína, os  $L_0$  *scores* subsequentes correspondem à classe impostora e o número total de *scores* é  $L = (L_1 + L_0)$ . A FAR e a FRR, considerando um limiar  $\eta$ , são dadas por:

$$FRR(\eta) = \frac{1}{L_1} \sum_{i=1}^{L_1} I(s_i < \eta), \quad (2.5)$$

$$FAR(\eta) = \frac{1}{L_0} \sum_{i=L_1+1}^L I(s_i \geq \eta), \quad \text{onde} \quad (2.6)$$

$$I(X) = \begin{cases} 1, & \text{se } x \text{ é verdadeiro,} \\ 0, & \text{caso contrário.} \end{cases} \quad (2.7)$$

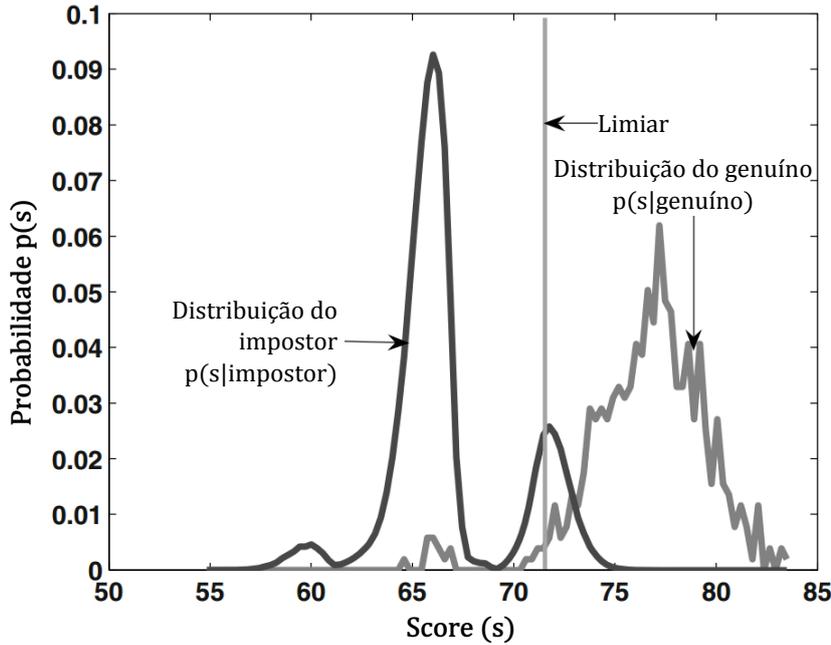


Figura 2.2: Exemplo de distribuição dos *scores*. O limiar determina da FAR e FRR do sistema. Note que dadas essas distribuições, a FAR e a FRR não podem ser reduzidas simultaneamente ao ajustar o limiar. Fonte: [1].

## 2.2 Sistemas multibiométricos

Sistemas de reconhecimento baseados em apenas um traço biométrico são denominados unibiométricos. Esses sistemas têm algumas limitações. Por exemplo, quando a fonte biométrica se torna não viável devido ao sensor ou ao software, a baixa qualidade de um traço biométrico específico ou mesmo manipulação deliberada. Além disso, aplicações de alta segurança e larga escala impõe requerimentos de eficácia rigorosos, que não podem ser alcançados por sistemas unibiométricos [1, 36].

Seres humanos reconhecem-se baseado na evidência de múltiplas características biométricas, além dos detalhes contextuais. O processo de reconhecimento pode ser visto como uma reconciliação de evidências pertencentes à múltiplas fontes de informação. Assim, sistemas biométricos também podem ser designados para reconhecer pessoas baseado na informação de múltiplas fontes biométricas. Tais sistemas são denominados sistemas multibiométricos e é esperado que sejam mais eficazes que os sistemas unibiométricos [9, 37].

A melhora da eficácia, que é a motivação primária para usar sistemas multibiométricos, acontece devido à duas razões. Primeiro, a fusão de múltiplas fontes biométricas efetivamente aumenta a dimensionalidade do espaço de características e reduz a sobreposição entre as distribuições de características de indivíduos diferentes. Em outras palavras, a

combinação de múltiplas fontes biométricas é mais única para um indivíduo [1].

Em segundo lugar, o ruído, a imprecisão ou o *drift* em um subconjunto de fontes biométricas pode ser compensado pela informação discriminatória das fontes remanescentes. Assim, a disponibilidade de múltiplas fontes biométricas provê redundância e tolerância a falhas. Outras vantagens incluem a amenização do problema de não-universalidade, um maior grau de flexibilidade para a autenticação, etc.

Embora esses sistemas ofereçam várias vantagens, eles são usualmente mais caros que sistemas unibiométricos devido à necessidade de maior processamento e maior tempo de cadastramento e reconhecimento. Assim, é essencial analisar o custo-benefício do custo adicionado e os benefícios alcançados. Exemplos de trabalhos que utilizam várias fontes biométricas podem ser vistos em [37, 36].

### 2.2.1 Tipos de sistemas de múltiplas fontes

Há cinco possíveis cenários que podem prover diferentes informações biométricas. Baseado nas fontes, sistemas multibiométricos podem ser classificados em multisensores, multialgoritmos, multi-instâncias, multiamostras e multimodais [1, 9, 13, 32]:

1. **Sistemas multisensores:** um único traço biométrico é imageado ou capturado usando múltiplos sensores a fim de extrair diversas informações. Por exemplo, um sistema pode armazenar informações bidimensionais da face de uma pessoa usando uma câmera e a modelagem tridimensional da face usando um sensor de profundidade.
2. **Sistemas multialgoritmos:** o mesmo traço biométrico é processado usando vários algoritmos. É o caso do sistema proposto neste trabalho. Por exemplo, um algoritmo baseado em textura de impressão digital e outro baseado em detalhes de impressão digital podem operar na mesma impressão digital a fim de extrair diversos conjuntos de características que podem melhorar a performance do sistema. Ou mesmo utilizar um conjunto de características comum para diferentes tipos de comparações. Esse tipo de sistema não requer o uso de novos sensores e, portanto, tem bom custo-benefício. Uma abordagem que examina diferentes abordagens multibiométricas em textos manuscritos pode ser vista em [13]. O trabalho avalia a abordagem multialgorítmica, multi-instâncias e multiamostras.
3. **Sistemas multi-instâncias:** usam múltiplas instâncias da mesma característica. Por exemplo, os dedos indicadores direito e esquerdo podem ser usados para verificar a identidade de um indivíduo. Esses sistemas geralmente não necessitam de novos sensores ou novas estratégias de extração de características e algoritmos de comparação sendo, assim, fáceis de implementar. Todavia, em alguns casos, um

novo arranjo dos sensores pode ser necessário para facilitar a captura simultânea das várias instâncias.

4. **Sistemas multiamostras:** um único sensor pode ser usado para adquirir múltiplas amostras do mesmo traço biométrico para contabilizar pelas variações que podem ocorrer no traço, ou para obter uma representação mais completa do traço. Um sistema de faces, por exemplo, pode capturar e armazenar os perfis direito e esquerdo em conjunto com a imagem frontal da face de uma pessoa para contabilizar pelas variações na postura facial. O sistema aqui desenvolvido também se enquadra nessa categoria.
5. **Sistemas multimodais:** combinam a evidência apresentada por diferentes traços corporais para estabelecer a identidade. Alguns dos primeiros sistemas multimodais usaram as características de face e voz para estabelecer a identidade de uma pessoa. Visto que diferentes traços biométricos de um indivíduo são descorrelacionados, o uso de sistemas multimodais geralmente leva a uma melhora maior comparados com outros sistemas multibiométricos. Todavia, algumas combinações de traços podem exibir correlação significativa e aplicar tais combinações pode não prover melhora na performance. O custo de sistemas multimodais também é substancialmente maior devido ao uso de múltiplos sensores.

### 2.2.2 Níveis de fusão

Um ponto fundamental na estrutura de sistemas multibiométricos é determinar o tipo de informação que deveria ser consolidada por um módulo de fusão. Essa fusão pode ser obtida utilizando a informação disponível em qualquer um dos módulos biométricos - módulos do sensor, do extrator de características, do comparador e da decisão. A fusão biométrica pode ser classificada em fusão anterior à comparação e fusão posterior à comparação.

Anterior à comparação, a integração de informação de múltiplas fontes pode tomar lugar tanto no nível de sensor, quanto no nível de características. Esquemas de integração de informações depois da classificação/comparação podem ser divididos em fusão em nível de decisão, fusão em nível de ranqueamento e fusão em nível de *score* [1, 31, 9, 32]:

1. **Nível do sensor:** consolida múltiplas fontes de dados brutos antes que eles sejam sujeitos à extração de características. Na literatura do processamento de imagens, é referido como fusão em nível de imagem. Em geral, uma função de fusão de nível do sensor  $f_S(\cdot)$  converte uma coleção de  $M$  amostras biométricas  $B = \{B_1, B_2, \dots, B_M\}$

em uma única amostra fundida  $B_f$ :

$$B_f = f_S(B_1, B_2, \dots, B_M). \quad (2.8)$$

Por exemplo, um pequeno sensor de impressão digital pode capturar duas ou mais impressões do dedo de uma pessoa para criar uma imagem composta que revela uma estrutura mais rígida do dedo.

2. **Nível de características:** envolve a consolidação de dois conjuntos de características de um mesmo indivíduo. Matematicamente, uma função de fusão de nível de características  $f_R(\cdot)$  converte uma coleção de  $M$  conjuntos de características  $X = \{x_1, x_2, \dots, x_M\}$  em um único conjunto de características  $x_f$ :

$$x_f = f_R(x_1, x_2, \dots, x_M) \quad (2.9)$$

Como pode haver variabilidade nos tipos de características vários processamentos podem ser adotados para normalização, equalização, etc.

3. **Nível do *score*:** É o tipo de fusão adotado no sistema em questão. Ocorre quando a saída de diferentes comparadores/classificadores são consolidados de forma a chegar em uma decisão final. Também é conhecida como fusão ao nível de mensuração ou nível de confiança. É relativamente fácil acessar e combinar os *scores* gerados por diferentes classificadores. Consequentemente, esse tipo de fusão é o mais comum nos sistemas multibiométricos. Seja  $s = [s_1, s_2, \dots, s_M]$  o vetor de *scores* obtido comparando um conjunto de características de entrada  $X^A$  com o conjunto de um perfil cadastrado  $X^E$  usando  $M$  classificadores diferentes. Baseado em treinamento ou algum outro critério/parâmetro específico, o classificador utiliza uma região de referência para classificar a amostra como genuína ou impostora. [36] apresenta um sistema multibiométrico com fusão em nível de *score*.
4. **Nível de ranqueamento:** é utilizada para a tarefa de identificação. Nesse caso, a saída do sistema pode ser vista como um ranqueamento das identidades cadastradas. O objetivo nesse caso é consolidar todos os ranques dos subsistemas biométricos individuais a fim de derivar um consenso para cada identidade.
5. **Nível de decisão:** é realizada quando somente as decisões feitas por cada classificador individual estão disponíveis. Métodos propostos pela literatura envolvem regras de “E” e “OU”, votação majoritária, votação majoritária ponderada, abordagem de função Bayesiana, entre outras. Um exemplo de trabalho que utiliza esse tipo de fusão pode ser visto em [37].

## 2.3 Biometrias fisiológicas

A relevância de uma biometria específica para uma aplicação é estabelecida em relação à natureza e requerimentos da aplicação, além de algumas propriedades da própria biometria. Dependendo dos traços considerados, as biometrias podem ser divididas em comportamentais e fisiológicas. As biometrias fisiológicas são baseadas em mensurações diretas de uma parte do corpo humano. Assim, o reconhecimento é baseado em “algo que você é”. Impressão digital, face, iris, escâner de mão são exemplos que pertencem a essa classe. Esta seção trará uma breve introdução de algumas biometrias fisiológicas [5, 10].

### 2.3.1 Impressão digital

Seres humanos usam impressões digitais para identificação pessoal desde muitas décadas. A digital é o padrão de cumes e vales na superfície da ponta do dedo, cuja informação é determinada durante os primeiros sete meses de desenvolvimento fetal. Enquanto impressões digitais vem sendo usadas para aplicações forenses por volta de cem anos, o advento de escâneres compactos e de baixo custo impactou na criação de um grande número de aplicações somente nos últimos dez anos [1, 10].

Em aplicações requerendo identificação em larga escala, envolvendo milhões de identidades, as impressões digitais múltiplas de uma pessoa podem ser usadas para melhorar a performance das comparações, embora seja feito ao custo de mais recursos computacionais. As impressões digitais de uma pequena fração da população pode ser inviável para identificação automática devido a fatores genéticos, idade, ambiente, razões ocupacionais, etc. Um trabalho envolvendo o uso de impressão digital e escâner de iris utilizando fusão em nível de decisão foi feito anteriormente [37].

### 2.3.2 Impressão da palma da mão

As palmas das mãos contém padrões de cumes e vales parecidos com os das impressões digitais. A área da palma é muito maior do que a área de um dedo e, como resultado, é esperado que as impressões das palmas da mão sejam ainda mais distintivas do que impressões digitais.

Como escâneres da palma precisam capturar uma área maior, eles são volumosos e mais caros do que sensores do dedo. Finalmente, usando um escâner de alta resolução, todas as características da mão tais como geometria, cumes, vales, linhas principais e rugas podem ser combinadas para construir um sistema biométrico com bastante eficácia [1]. [38] apresenta um sistema biométrico para controle de acesso baseado no escâner da palma da mão.

### 2.3.3 Iris

A iris é a região anelar do olho delimitada pela pupila e esclera (parte branca do olho) em ambos os lados. A textura visual da iris é formada durante o desenvolvimento fetal e estabiliza durante os primeiros dois anos de idade. A textura complexa da iris carrega informação bastante distintiva para reconhecimento pessoal. A eficácia e velocidade do desenvolvimento de sistemas de reconhecimento baseados na iris é promissora e pode suportar reconhecimento em larga escala [1, 10]. Pesquisas envolvendo a utilização de iris para fins biométricos podem ser vistas em [37, 36]

### 2.3.4 Face

O reconhecimento de face é um método não intrusivo e atributos faciais são, provavelmente, as características mais comuns utilizadas por humanos para se reconhecerem. Embora a performance de autenticação de sistemas de reconhecimento de face seja aceitável para uso em certas aplicações, eles impõe um número de restrições em como as imagens da face são obtidas, comumente requerendo um fundo fixo e simples com iluminação controlada. Esses sistemas também tem dificuldade em comparar imagens de faces capturadas de diferentes vistas, sob condições de iluminação distintas [1, 10]. A utilização da face como característica em um sistema biométrico pode ser vista em [36].

### 2.3.5 Voz

A voz é uma combinação de características biométricas físicas e comportamentais. As características físicas da voz de um indivíduo são baseadas na forma e tamanho dos apêndices, que são usados para a síntese do som. As características físicas da fala humana são invariantes para um indivíduo, mas os aspectos comportamentais mudam no decorrer do tempo devido à idade, condições médicas, etc [1, 31].

Uma desvantagem de sistemas de reconhecimento baseados em voz é que as características da fala são muito sensíveis a fatores como ruído do fundo e características do microfone. O reconhecimento do interlocutor é mais apropriado em aplicações baseadas em telefonia mas o sinal de voz é tipicamente degradado em qualidade pelo canal de comunicação [1, 10]. [39] apresenta um sistema biométrico para usuários pré-definidos em interações dependentes de texto e independentes de texto. A distância de Mahalanobis e a quantização vetorial são utilizadas na classificação.

## 2.4 Biometrias comportamentais

Há uma variedade de biometrias comportamentais, e o objeto desse estudo recai nessa classe de biometrias. Pode-se denominar biometrias comportamentais como aquelas baseadas na maneira como os indivíduos proveem informação ao sistema de autenticação. Nessa classe entra o reconhecimento de voz (que também se encontra nessa classe, conforme citado na seção anterior), a verificação de assinatura e texto, a dinâmica das batidas no teclado, a dinâmica do uso do *mouse*, entre outros [8, 10].

A maioria da pesquisa científica focou o estudo das biometrias físicas bem estabelecidas tais como impressão digital e escâner de íris. Sistemas de biometrias comportamentais são usualmente menos estabelecidos. Entretanto, biometrias comportamentais proveem um número de vantagens sobre as tecnologias de biometria tradicionais. Elas podem ser coletadas de maneira não invasiva/intrusiva [8] e, em alguns casos, até mesmo sem o conhecimento do usuário. A coleta de dados comportamentais comumente não requer dispositivos especiais e tem um bom custo-benefício [40].

As biometrias comportamentais podem ser classificadas em cinco categorias baseadas no tipo de informação sobre o usuário [40]:

1. **De autoria:** são obtidas com a análise de um pedaço de texto ou desenho produzido pela pessoa. A verificação pode ser feita observando peculiaridades do autor do trabalho. Esse é o tipo de biometria adotada no sistema em questão, que é baseado tanto em assinaturas quanto em palavras cursivas isoladas.
2. **Interação direta homem-computador:** na interação diária com computadores, seres humanos empregam diferentes estratégias, usam diferentes estilos e aplicam habilidades e conhecimento único. O foco é quantificar tais traços.
3. **Interação indireta homem-computador:** bastante relacionada com a anterior, sendo o conjunto de biometrias indiretas da interação homem-computador. São eventos que podem ser obtidos ao monitorar o comportamento do usuário por ações como traços de chamadas, traços de execução de programas, registro de acesso, atividade de armazenamento, entre outras.
4. **Motoras:** a categoria mais pesquisada dentre as biometrias comportamentais. Recai nas habilidades motoras dos usuários. Habilidade motora é a capacidade de usar músculos e são aprendidas, não herdadas.
5. **Puramente comportamental:** medem o comportamento humano não se concentrando diretamente em mensurações de partes do corpo ou ações musculares. Seres humanos utilizam diferentes estratégias, habilidades e conhecimento durante a reali-

zação de tarefas mentais. Essas biometrias focam-se na quantificação de tais traços comportamentais.

### 2.4.1 Reconhecimento automático de escritor

O reconhecimento automático de escritor é o tipo de reconhecimento biométrico baseado em amostras de texto do usuário [10]. No trabalho aqui apresentado, a palavra “texto” representará tanto assinaturas cursivas quanto palavras isoladas cursivas e o termo reconhecimento se refere especificamente à verificação.

As assinaturas manuscritas ocupam um lugar bastante especial nesse conjunto de traços biométricos. Isto deve-se principalmente ao fato de que assinaturas manuscritas têm se estabelecido como o meio mais amplo de verificação pessoal. Assinaturas são geralmente reconhecidas como um meio legal de verificar a identidade de um indivíduo por instituições financeiras e administrativas. Além disso, a verificação por assinatura não requer nenhum tipo de mensuração invasiva e as pessoas estão familiares com o uso delas no seu dia a dia. Uma assinatura manuscrita é o resultado de um processo complexo que depende das condições sob as quais o processo de assinatura ocorre. Dessa maneira, embora teorias complexas tenham sido propostas para modelar os mecanismos psicológicos sob a escrita, a verificação de assinatura ainda permanece um desafio em aberto [2].

O reconhecimento automático de assinatura envolve disciplinas que vão da anatomia humana para a engenharia, da neurociência à ciência da computação. Nos anos recentes, estudos sobre verificação de assinatura têm atraído pesquisadores de diferentes campos, os quais estão interessados não só nos desafios científicos mas também nas valiosas aplicações que este campo oferece.

Em conjunto com os avanços recentes e o crescimento extraordinário da internet, a verificação automática de assinatura está sendo considerada com novos interesses. A criação de leis específicas e regulamentações, que tem sido aprovadas em vários países e a atenção que várias associações nacionais e institutos internacionais tem dado à padronização dos dados de assinaturas são evidências do interesse nesse campo. O objetivo desses esforços é facilitar a integração de tecnologias de verificação de assinatura para uma ampla gama de aplicações comerciais, tais quais bancos, segurança médica, gerenciamento de documentos, *e-commerce*, entre outros [2].

Dentro de todo este contexto, há aspectos muito semelhantes no uso de palavras cursivas. Por exemplo, a familiaridade que as pessoas tem com o ato de escrever. Embora não haja vínculo aparente com escrever e verificar, essa familiaridade em conjunto com o uso de métodos não invasivos para a coleta de palavras cursivas as fazem uma boa proposta também. Quando comparadas com assinaturas, palavras isoladas parecem carregar menos poder discriminativo. As assinaturas de duas pessoas tendem a ser bastante diferentes,

enquanto que duas execuções de uma mesma palavra tendem a ser bem mais parecidas [14]. Em outras palavras, a assinatura não coloca restrições para as diferenças de cada escritor enquanto que palavras são restringidas por dois fatores fundamentais [5]:

- Duas execuções de uma palavra tem que ser parecidas para serem consideradas a mesma palavra;
- Elas tem que ser legíveis.

Os textos, todavia, tem uma propriedade singular, que as assinaturas por si só não tem: cancelabilidade. O usuário de um sistema biométrico não pode ser solicitado a escrever uma assinatura mais longa ou diferente, enquanto tem a disponibilidade de escrever um texto mais longo ou diferente. Dessa maneira, um texto pode ser cancelado ou substituído por um novo, enquanto o mesmo não acontece com a assinatura. Portanto, sob esse ponto de vista, em alguns cenários de segurança, o texto online manuscrito pode ser mais seguro que a assinatura online. E mesmo que não fosse, a escrita cursiva que não é assinatura é um traço biométrico que poderia complementar outros mecanismos de reconhecimento [14].

O reconhecimento automático de escritor pode seguir duas abordagens: online e offline. Tanto as abordagens online quanto offline tem sido estudadas em trabalhos anteriores. mas quando se refere ao tipo de texto usado, o reconhecimento de assinatura tem atraído muito mais atenção do que abordagens baseadas em outros tipos de escrita, especialmente no caso online. Para o campo offline, abordagens não baseadas em assinatura tem recebido alguma atenção devido à sua relevância na investigação forense principalmente. Mas no contexto online, constata-se que abordagens não baseadas em assinaturas são limitadas e quando considera-se apenas palavras avulsas ou pequenas sentenças, poucos trabalhos abordaram este problema [14, 5]. As próximas seções irão abordar a respeito das duas abordagens para o reconhecimento, tanto de assinaturas quanto de textos no geral.

## Reconhecimento offline

O reconhecimento offline de textos utiliza um escâner para obter uma imagem de uma amostra [15]. Tal qual, elas contém apenas as características estáticas da amostra como a inclinação, a caixa de limitação e a largura da caneta. São basicamente informações de espaço-luminância da imagem adquirida, não havendo referência temporal, e são denominadas características offline. Sistemas estáticos realizam a aquisição de dados após o processo de escrita ter sido completado. Nesse caso, a amostra é representada como uma imagem em tons de cinza,  $\{S(x, y)\}_{0 \leq x < X, 0 \leq y < Y}$ , onde  $S(x, y)$  denota o nível de cinza na posição  $(x, y)$  da imagem e  $X \times Y$  representa o tamanho total da imagem [2].

Em adição ao conjunto mínimo de características para trabalhar, a qualidade dos dados pode variar dependendo da qualidade da imagem contendo a assinatura e aspectos relacionados ao processo de aquisição da assinatura. Tipicamente, a imagem pode necessitar ser pré-processada a fim de remover artefatos que possam ter ficado integrados com a amostra de assinatura - por exemplo, imagem levemente borrada, fundo manchado, etc. Depois desse pré-processamento, pode-se começar o trabalho de verificar a amostra. A fim de desenvolver um esquema de verificação, uma amostra de referência deve estar disponível. Usuários serão tipicamente solicitados a produzir suas assinaturas/textos um número de vezes, e esses dados são então usados para formar uma referência que será utilizada no processo subsequente de reconhecimento [31].

Tipicamente, o estágio de pré-processamento envolve a limpeza do fundo e a aplicação de algum tipo de filtragem bidimensional que preserve as bordas da imagem. O próximo estágio é a binarização da imagem com respeito a algum tipo de limiar. Outro tipo de pré-processamento envolve a normalização da imagem com respeito às dimensões horizontais e verticais. Isso pode ser necessário ao comparar imagens de tamanhos diferentes, que pode ocorrer quando as amostras são produzidas em ambientes diferentes. Todavia, esse processo de escala pode distorcer algumas informações contidas na amostra, assim essa opção só deve ser empregada quando necessário [31].

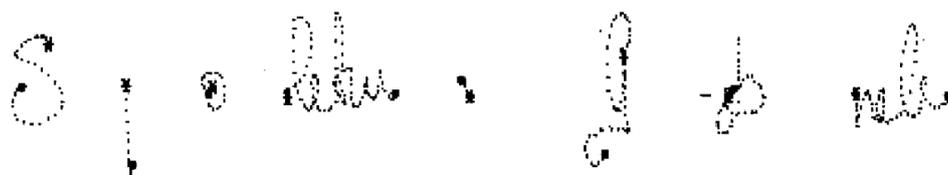
Outros pontos relacionados envolvem a largura da caneta, a qual pode alterar as dimensões gerais da amostra, assim como a possibilidade de distorção de algumas informações locais tais como curvatura. Uma vez que os dados tenham sido pré-processados e normalizados, os estágios subsequentes envolvem a extração de características, a construção do modelo de referência e as estratégias de classificação, como é usual. A literatura nesse campo é bastante vasta, sendo possível encontrar vários trabalhos envolvendo uma grande variedade de técnicas.

## Reconhecimento online

O reconhecimento online de texto refere-se a um método para automatização do reconhecimento que ocorre essencialmente em tempo real. Os sistemas dinâmicos usam dispositivos de aquisição online que geram sinais eletrônicos para representar a imagem durante o processo de escrita [15]. Nesse caso, uma amostra é representada como uma sequência  $\{S(n)\}_{n=0,1,\dots,N}$ , onde  $S(n)$  é o valor das características extraídas do sinal amostrado no tempo  $n\Delta t$  do processo de escrita/assinatura,  $0 \leq n \leq N$ , com  $\Delta t$  sendo o período de amostragem. Dessa forma, o caso offline envolve o tratamento da espaço-luminância da imagem enquanto o caso online se concentra no tratamento da representação espaço-temporal da amostra [2].



(a)



(b)

Figura 2.3: Exemplos de assinaturas. (a) Assinatura offline, onde constam somente informações de luminância dos pixels. (b) Assinatura dinâmica, representada pelos pontos amostrados (“\*”: pontos amostrados no toque; “•”: pontos amostrados no ar). Fonte: [2]

Os desafios de sistemas online são, portanto, diferentes de uma abordagem offline, tanto em termos de requerimentos de processamento (em outras palavras, o tempo para decisão) e a natureza a informação disponível para gerar o processo de reconhecimento. Muitos autores fazem a distinção entre as características estáticas e dinâmicas dos dados, havendo alegações de vantagens ao incorporar ambas as formas nos modelos. Autores também reportam que os dados dinâmicos sozinhos proveem mais capacidade discriminativa do que características estáticas [31, 15].

### Classificação das características

Conforme visto nas seções anteriores, as características podem ser online (decorrem de informação temporal) e offline (não decorrem de informação temporal), o que, por sua vez, diferencia os sistemas em online e offline, respectivamente [41]. Uma forma de definir as características é classificá-las simplesmente como locais, globais e regionais [41]. Locais são as características referentes à cada ponto amostrado, como por exemplo as coordenadas espaciais, a pressão, etc. Sistemas baseados nesse tipo de características são computacionalmente mais complexos visto que há necessidade de técnicas elásticas de comparação, como por exemplo a técnica da deformação dinâmica do tempo (DTW, do inglês, *Dynamic Time Warping*), detalhada no próximo capítulo. Este tipo de características são sempre funções dependentes do tempo. As características globais são aquelas que representam a amostra como um todo tais quais tempo total de assinatura, altura, largura, velocidade média, etc. São bem mais eficientes em questão de tempo de proces-

samento, todavia têm resultados piores do que as características locais. Por último, as características regionais são aquelas que se referem à determinadas regiões da assinatura.

Uma outra classificação, ainda, pode ser adotada. As características podem ser classificadas em funções e parâmetros [41, 2]. Funções são uma outra forma de denominar as características locais. Quando parâmetros são usados, a amostra é caracterizada como um vetor de elementos, cada elemento representando o valor de uma característica. Parâmetros são classificados em duas grandes categorias: globais e locais. Parâmetros globais abordam a amostra como um todo. Parâmetros globais típicos são o tempo total de duração, o número de toques na tela, número de componentes da amostra, orientação global da amostra, coeficientes obtidos por transformações matemáticas, etc. Esses são os tipos de características utilizadas no trabalho aqui apresentado.

Parâmetros locais referem-se a características extraídas de partes específicas da amostra. Dependendo do nível de detalhes considerado, parâmetros locais podem ser divididos em parâmetros orientados a componentes, que são extraídos no nível de cada componente (por exemplo, taxa de altura por largura de um traço, posições relativas a um traço, orientação de um traço, etc.), e parâmetros orientados a pixel, os quais são extraídos em nível dos pixels (por exemplo, informação baseada em *grid*, densidade dos pixels, intensidade dos níveis de cinza, textura, etc.). É interessante notar que alguns parâmetros, os quais são geralmente considerados características globais, também podem ser aplicados localmente e vice versa. Por exemplo, características baseadas em contorno podem ser extraídas em nível global (por exemplo, envelopes extraídos no nível global da amostra) ou nível local (por exemplo, ao nível de cada componente conectado).

Os parâmetros, ao contrário das funções, podem representar tanto características online como offline. Sempre que os parâmetros utilizarem de informações dependentes do tempo, eles serão considerados características online. Caso contrário, se foram extraídos a partir de análises da imagem da amostra, serão offline [2]. A Figura 2.4 mostra um diagrama desse modelo de classificação de características.

## Principais características utilizadas

As características mais utilizadas no reconhecimento de texto/assinatura manuscrito online são a posição, velocidade, aceleração [33, 35, 28]. A função de posição é convertida diretamente pelo dispositivo de aquisição enquanto a velocidade e aceleração podem ser providas diretamente pelo dispositivo de entrada ou derivadas diretamente. Nos anos recentes, a pressão tem sido explorada frequentemente a fim de tirar vantagem da não dependência dessa característica com as outras. Além disso, direção do movimento da caneta e a inclinação da caneta [5, 14] têm sido consideradas para melhorar a performance no reconhecimento online. Em geral, as funções de posição, velocidade e inclinação da

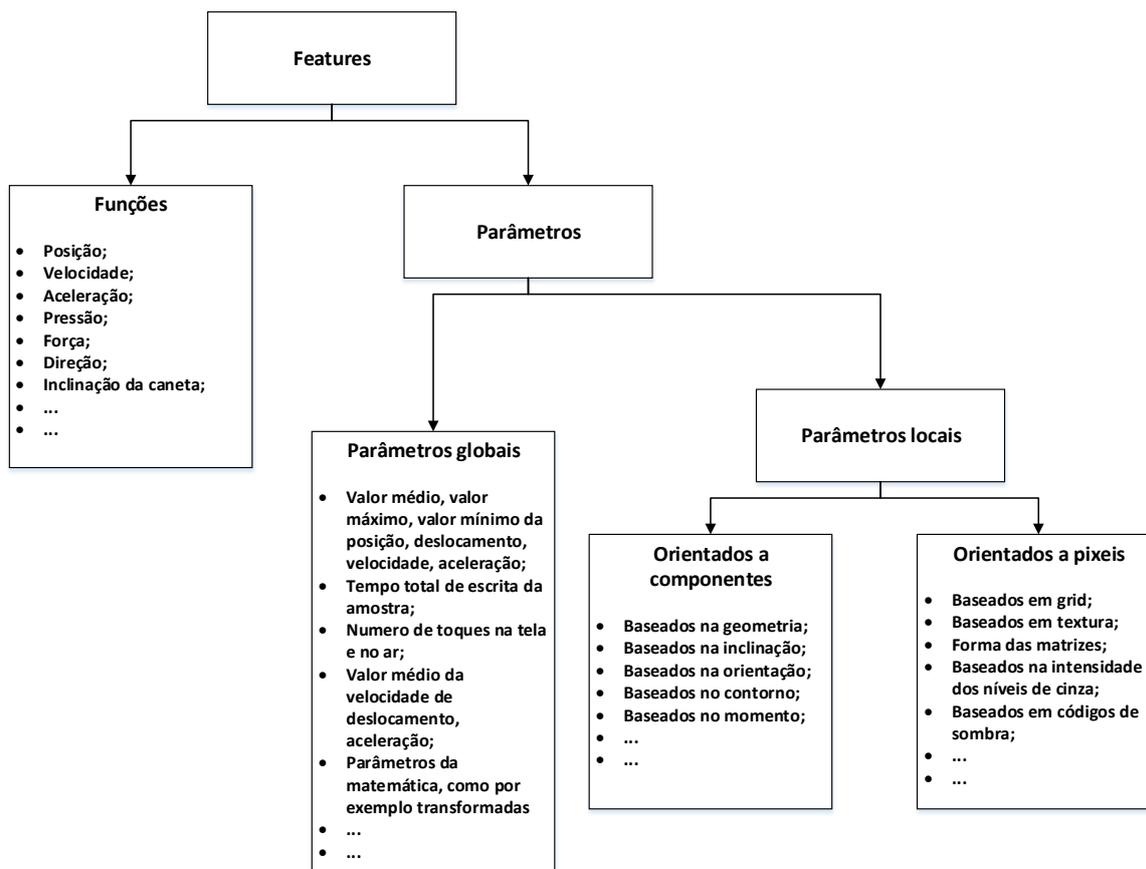


Figura 2.4: Modelo de classificação de características. Fonte: [2].

caneta são consideradas dentre as mais consistentes características no reconhecimento online de texto, quando modelos de consistência baseados em distância são aplicados. Esse modelo considera que características também podem ser estimadas usando uma medida de distância associada com as características por si só [2].

Dentre os parâmetros utilizados em sistemas de reconhecimento de texto online estão a Transformada de Fourier [7, 21, 22], a Transformada de Cossenos [23, 24, 25], a Transformada de *Wavelets*, Transformada de Radon [42], o número de traços [5, 14], valores extremos e médios de velocidade, aceleração, posição [29], entre outras. Histogramas não foram muito explorados no reconhecimento online de texto, sendo alguns trabalhos apresentados em [6, 29].

### Técnicas de classificação comumente adotadas

No processo de verificação, a autenticidade de uma assinatura/texto é avaliada pela comparação de suas características com aquelas armazenadas na base de conhecimento

durante o estágio de cadastramento. Esse processo produz uma resposta booleana que certifica a autenticidade da amostra. Várias técnicas foram adotadas para a classificação das amostras.

Quando técnicas de comparação de perfil são utilizadas, a amostra questionada é comparada com um perfil de referência. Nesse caso, as abordagens mais comuns usam a DTW para a comparação [5, 28, 14]. O uso de DTW, ou de outras técnicas próprias para comparar vetores de tamanhos diferentes se torna obrigatório no uso de funções como características, que naturalmente terão variabilidade nos comprimentos. Para evitar evitar o uso de comparações com este tipo de técnicas, a DTW também é utilizada para o alinhamento dos tamanhos de características de funções [33, 35]. É importante notar que há um custo computacional alto, qualquer seja o caso, ao utilizar este grupo de técnicas para comprimentos variáveis.

Quando abordagens estatísticas são adotadas, classificadores baseados em distância são utilizados. Essas abordagens estatísticas são comumente utilizadas quando parâmetros são adotados características. As abordagens mais comuns envolvem o uso de distância Euclidiana [7, 23], distância de Mahalanobis e distância de Manhattan [6].

Redes neurais são amplamente utilizadas para a verificação de assinatura devido às suas capacidades de aprender e generalizar [26]. Além disso, especial atenção tem sido dada ao uso de modelos de cadeias ocultas de Markov tanto para o reconhecimento online quanto offline [2, 27, 43, 44]. É possível encontrar uma ampla gama técnicas utilizadas para classificação, por exemplo as baseadas em lógica nebulosa [33, 35], máquinas de suporte vetorial [45] e abordagens Bayesianas [46].

Dessa forma, o sistema proposto neste trabalho é um tipo de sistema multibiométrico, com fusão em nível de *score*. Além disso, por se basear em histogramas, descritores de *Fourier* e informação global de velocidade, pode ser caracterizado como um sistema online baseado em métodos/características globais. Diferente de [5], que adota o uso de palavras isoladas em letra de forma, e de vários outros trabalhos encontrados na literatura, o objeto de estudo deste trabalho são as palavras cursivas isoladas, além das assinaturas. Por fim, diferente dos cenários comumente analisados e base de dados usualmente adotadas, que consideram a utilização de mesas digitalizadoras com dispositivos especiais para captura [6], a avaliação das palavras cursivas foi feita com o uso de *tablets* comuns e sem o uso de dispositivos adicionais. O próximo capítulo abordará alguns dos conceitos matemáticos relacionados neste trabalho.

# Capítulo 3

## Fundamentação Teórica

Este capítulo abordará definições e aplicações de algumas das ferramentas envolvidas na proposta. Primeiramente as ferramentas das análises no domínio da frequência serão introduzidas, nomeadamente, a Série e a Transformada de Fourier, assim como alguns de seus possíveis usos. Em seguida, serão apresentados conceitos relacionados aos histogramas, assim como o seu uso na área de reconhecimento de padrões. Por último, serão discutidas algumas métricas de comparação, a distância Euclidiana, a distância de Manhattan e a DTW.

### 3.1 Análise de Fourier

Qualquer função periódica pode ser expressa como soma de senos e cossenos de frequências diferentes, cada um multiplicado por um coeficiente diferente. Não importa o quão complicada é uma função, se é periódica e contínua, ela pode ser representada por tal soma, chamada de Série de Fourier [47]. Mesmo funções não periódicas, mas cuja área sobre a curva é finita, podem ser expressas como a integral de senos e cossenos multiplicados por um peso. A formulação, nesse caso, é a Transformada de Fourier, e sua utilidade é ainda maior que a das Séries de Fourier em muitos campos teóricos e aplicados [47].

#### 3.1.1 A Transformada de Fourier de Função de uma Variável Contínua

A transformada de Fourier de uma função contínua  $f(t)$  de uma variável contínua  $t$ , é definida pela equação

$$F(\mu) = \mathcal{F}\{f(t)\} = \int_{-\infty}^{\infty} f(t)e^{-i2\pi\mu t} dt \quad (3.1)$$

onde  $\mu$  também é uma variável contínua. Inversamente, dada uma função  $F(\mu)$ , é possível obter  $f(t)$  de volta usando a transformada inversa, dada por

$$f(t) = \mathcal{F}^{-1}\{F(\mu)\} = \int_{-\infty}^{\infty} F(\mu)e^{i2\pi\mu t}d\mu \quad (3.2)$$

Na Transformada de Fourier, pelo fato da única variável a restar depois da integração é  $\mu$ , que representa a frequência, diz-se que o domínio de Fourier é o domínio da frequência. Essas duas equações formam o chamado par de transformação de Fourier. Isso indica o importante fato de que uma função pode ser recuperada de sua transformada [47].

### 3.1.2 DFT - A Transformada Discreta de Fourier

Seja uma função amostrada

$$\tilde{f}(t) = f(t)s_{\Delta T}(t) = \sum_{n=-\infty}^{\infty} f(t)\delta(t - n\Delta T) = f(k\Delta T) \quad (3.3)$$

onde  $\delta(t) = 1$  se  $t = 0$  e 0 caso contrário e  $s_{\Delta T}(t) = \sum_{n=-\infty}^{\infty} \delta(t - n\Delta T)$  é o trem de impulsos, tal que  $n, k \in \mathbb{Z}$  (conjunto dos números inteiros). O teorema da convolução diz que a transformada de Fourier do produto de duas funções no domínio espacial/temporal é a convolução das transformadas das duas funções e vice-versa. Assim, a transformada de Fourier da função discreta é dada por

$$\tilde{F}(\mu) = \mathcal{F}\{\tilde{f}(t)\} = \mathcal{F}\{f(t)s_{\Delta T}(t)\} = F(\mu) * S(\mu) \quad (3.4)$$

onde

$$\mathcal{F}\{s_{\Delta T}(t)\} = S(\mu) = \frac{1}{\Delta T} \sum_{n=-\infty}^{\infty} \delta(\mu - \frac{n}{\Delta T}) \quad (3.5)$$

tal que  $n \in \mathbb{Z}$  [47].

A Transformada de Fourier de uma função discreta de banda limitada é uma função contínua e periódica. Dada a função  $\tilde{F}(\mu) = \int_{-\infty}^{\infty} \tilde{f}e^{-i2\pi\mu t}dt$ , e usando o fato que  $\tilde{f}(t) = f(t)s_{\Delta T}(t) = \sum_{n=-\infty}^{\infty} f(t)\delta(t - n\Delta T)$ , é possível derivar a transformada discreta de Fourier:

$$\begin{aligned} \tilde{F}(\mu) &= \int_{-\infty}^{\infty} \sum_{n=-\infty}^{\infty} f(t)\delta(t - n\Delta T)e^{-i2\pi\mu t}dt \\ &= \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} f(t)\delta(t - n\Delta T)e^{-i2\pi\mu t}dt \\ &= \sum_{n=-\infty}^{\infty} f_n e^{-i2\pi\mu n\Delta T} \end{aligned} \quad (3.6)$$

Embora  $f_n$  seja uma função discreta, sua transformada de Fourier  $\tilde{F}(\mu)$  é uma função contínua e periódica, com período  $\frac{1}{\Delta T}$ . Dessa maneira, só é necessário caracterizar um único período de  $\tilde{F}(\mu)$ . Suponha que deseja-se obter  $M$  amostras igualmente espaçadas de  $\tilde{F}(\mu)$  dado o período  $\mu = [0, \frac{1}{\Delta T}]$ . Isso pode ser obtido usando amostras das seguintes frequências [47]:

$$\mu = \frac{m}{M\Delta T} \quad m = 0, 1, 2, \dots, M - 1 \quad (3.7)$$

Substituindo esse valor for  $\mu$ , é possível obter a expressão para a DFT:

$$F_m = \sum_{n=0}^{M-1} f_n e^{-\frac{i2\pi mn}{M}} \quad m = 0, 1, 2, \dots, M - 1 \quad (3.8)$$

A transformada inversa (IDFT) é definida por:

$$f_n = \frac{1}{M} \sum_{m=0}^{M-1} F_m e^{\frac{i2\pi mn}{M}} \quad n = 0, 1, 2, \dots, M - 1 \quad (3.9)$$

### 3.1.3 FFT - A Transformada de Fourier Rápida

Trabalhar com o domínio da frequência não seria prático se fosse necessário implementar as equações apresentadas para a DFT diretamente. Uma implementação de força bruta para as equações requer somas e adições na ordem de  $\mathcal{O}(MN)^2$ . Para isso foi desenvolvido o algoritmo FFT, que reduz os cálculos para ordem de  $\mathcal{O}(MN \times \log_2 MN)$  adições e somas [47].

Para explicar a ideia pela qual reduzem-se os cálculos, será utilizada a versão 1D da transformada. É comum expressar a equação na forma

$$F(u) = \sum_{x=0}^{M-1} f(x) W_M^{ux} \quad (3.10)$$

$u = 0, 1, \dots, M - 1$ , onde

$$W_M = e^{-\frac{i2\pi}{M}} \quad (3.11)$$

e assume-se que  $M$  estará na forma  $M = 2^n$ , com  $n$  sendo um inteiro positivo. Assim,  $M$  pode ser expressa na forma  $M = 2K$ , com  $K$  sendo um inteiro positivo também. Assim:

$$F(u) = \sum_{x=0}^{2K-1} f(x) W_{2K}^{ux} + \sum_{x=0}^{K-1} f(2x) W_{2K}^{u(2x)} + \sum_{x=0}^{K-1} f(2x+1) W_{2K}^{u(2x+1)} \quad (3.12)$$

Todavia, tem-se que  $W_{2K}^{2ux} = W_K^{ux}$ , dessa forma, a equação pode ser representada por:

$$F(u) = \sum_{x=0}^{K-1} f(2x)W_K^{ux} + \sum_{x=0}^{K-1} f(2x+1)W_K^{ux}W_{2K}^u \quad (3.13)$$

Definindo  $F_{\text{par}}(u) = \sum_{x=0}^{K-1} f(2x)W_K^{ux}$  e  $F_{\text{impar}}(u) = \sum_{x=0}^{K-1} f(2x+1)W_K^{ux}$ , para  $u = 0, 1, 2, \dots, K-1$  e, considerando que  $W_M^{u+M} = W_M^u$  e  $W_{2M}^{u+M} = -W_{2M}^u$ , a equação se torna:

$$F(u+K) = F_{\text{par}}(u) - F_{\text{impar}}(u)W_{2K}^u \quad (3.14)$$

Analisando essa equação, uma propriedade interessante pode ser observada. Uma transformada de  $M$  pontos pode ser calculada dividindo a expressão original em duas partes. Avaliar a primeira parte de  $F(u)$  requer a computação de duas transformadas de  $\frac{M}{2}$  pontos. A outra metade segue diretamente. Essa é a propriedade utilizada para diminuir a quantidade de cálculos da transformada de Fourier, sendo a base do algoritmo da FFT [47].

### 3.1.4 Usos e aplicações da Transformada de Fourier

A publicação do algoritmo da transformada rápida de Fourier, em 1965, abriu uma nova área no processamento digital de sinais por reduzir a ordem de complexidade do cálculo da transformada [48]. Desde então, análises no domínio da frequência tem sido utilizadas em várias áreas, como por exemplo no processamento de imagens, no reconhecimento de padrões e em visão computacional.

No campo de processamento de imagens, é possível ver aplicações para o processo de filtragem. Suavização e realce com filtros ideais, Butterworth e Gaussianos são comumente utilizados. Além disso, filtragem seletivas, como as passa-bandas e os filtros de Notch também são adotados. Uma outra utilização para a transformada de Fourier é na compressão de informações [47].

No campo de reconhecimento de padrões e visão computacional, a transformada tem sido aplicada de variadas maneiras para extrair características de objetos e ações, como pode ser visto em [49], onde são extraídas características espectrais para reconhecimento de faces. Ou, por exemplo, em [50] onde a transformada é utilizada para o reconhecimento de formas não rígidas. Em [51, 52] a transformada é usada como característica para a recuperação de imagens. [53] utiliza descritores de Fourier para o reconhecimento de caracteres em chinês e [54] faz seu uso em modulação de frequência de ondas de radares para reconhecimento de veículos.

Quando o assunto é o reconhecimento de escritores, é possível encontrar vários trabalhos que utilizam os descritores de Fourier. [7] traz uma nova forma de utilizar os descritores para o reconhecimento automático de assinaturas *online* manuscritas. Em [55] é possível ver a aplicação dos descritores no reconhecimento automático de assinaturas *offline*, enquanto [56] utiliza características espectrais numa abordagem híbrida, considerando características *online* e *offline*.

## 3.2 Histogramas

Um histograma é uma representação da distribuição dos dados. É um estimador da distribuição de probabilidade de uma variável contínua (variável quantitativa) e foi introduzido por Karl Pearson [3, 57]. Para construir um histograma, o primeiro passo é agrupar o intervalo dos valores, isto é, dividir o intervalo completo em uma série de intervalos menores. Após isso, conta-se quantos valores estão em cada um dos intervalos.

Um histograma pode também ser normalizado, mostrando, assim, as frequências relativas. Dessa forma, mostra a proporção de casos que estão em cada uma das várias categorias, com a soma dos pesos totalizando 1. Esta seção tratará a respeito dessa forma de representar os dados.

### 3.2.1 Definição

Histogramas são estimações de densidade [3, 57]. Uma estimação da densidade representa a distribuição dos dados, além de mostrar a sua multimodalidade. Dada uma origem  $x_0$ , a ideia é representar a densidade dos dados contando o número de observações em uma sequência de intervalos consecutivos (os chamados *bins*). Seja  $B_j(x_0, h)$  um *bin* de comprimento  $h$ , que é um elemento em uma grade de *bins* começando em  $x_0$ :

$$B_j(x_0, h) = [x_0 + (j - 1) \times h, x_0 + j \times h), j \in \mathbb{Z} \quad (3.15)$$

onde  $[\bullet, \bullet)$  denota um intervalo fechado à esquerda e aberto à direita [3, 57]. Se  $\{x_i\}_{i=1}^n$  é uma amostra i.i.d (independente e identicamente distribuída) com densidade  $f$ , o histograma é definido como segue:

$$\hat{f}_h(x) = n^{-1}h^{-1} \sum_{j \in \mathbb{Z}} \sum_{i=1}^n I\{x_i \in B_j(x_0, h)\} \times I_{x \in B_j(x_0, h)}. \quad (3.16)$$

Na soma explicitada pela Equação 3.16, a primeira função indicadora,  $I\{x_i \in B_j(x_0, h)\}$  conta o número de observações que estão no *bin*  $B_j(x_0, h)$ . A segunda função indicadora é responsável por “localizar” as contagens em torno de  $x$ . O parâmetro  $h$  é um parâmetro

de suavização ou localização e controla a largura dos *bins* do histograma. Um  $h$  que seja muito grande levará a blocos muito grandes e, conseqüentemente, a um histograma desestruturado. Por outro lado, um  $h$  muito pequeno dá uma estimação muito variável com muitos picos não importantes.

O efeito de mudar  $h$  pode ser visto na Figura 3.1, que contém histogramas sobre análise da diagonal de notas falsificadas do banco (diagonais cuja análise ajuda a diferenciar notas verdadeiras de falsas) para  $x_0 = 137.8$  (o mínimo dessas observações) e  $h = 0.1$ . Aumentar  $h$  para  $h = 0.2$  e usar a mesma origem,  $x_0 = 137.8$ , resulta no histograma mostrado no canto inferior esquerdo da figura. Esse histograma está, de certa forma, mais suave devido ao  $h$  maior. A largura do *bin* é fixada em  $h = 0.3$  (canto superior direito).

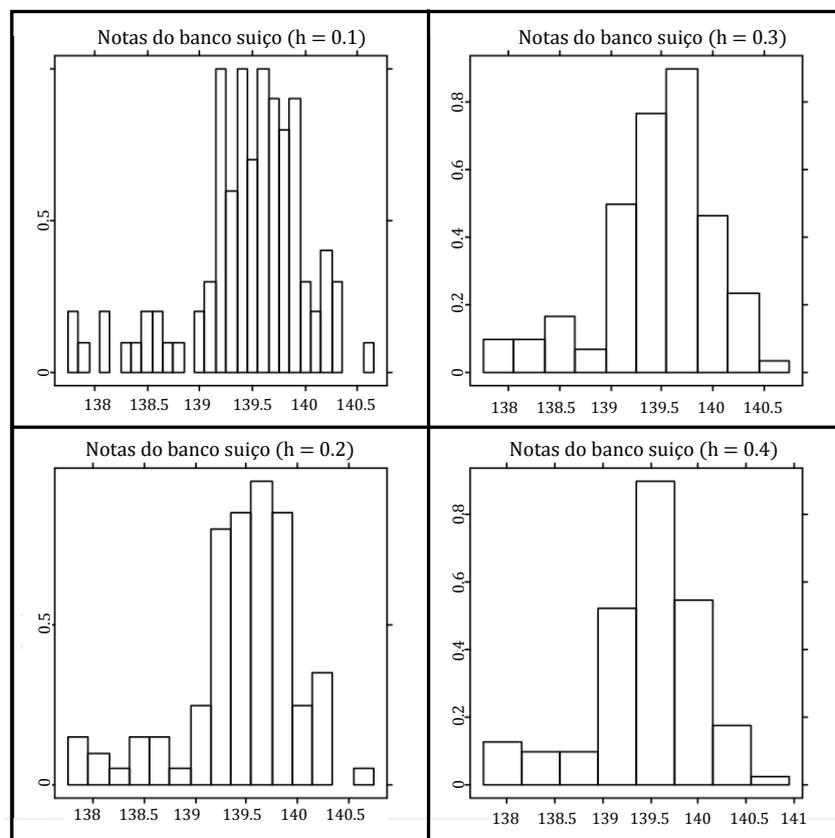


Figura 3.1: Diagonal das notas de banco falsificadas. Histogramas com  $x_0 = 137.8$  e  $h = 0.1$  (superior esquerdo),  $h = 0.2$  (inferior esquerdo),  $h = 0.3$  (superior direito) e  $h = 0.4$  (inferior direito). Fonte: [3].

Desse histograma, tem-se a impressão que a distribuição da diagonal é bimodal com picos em torno de 138.5 e 139.9. Para decidir a largura dos bins, pode-se ser necessário alguma estratégia. Uma proposta pode ser vista em [58], baseada em métodos de suavi-

zação. A partir dessa ideia pode-se estimar uma largura  $h$  “ótima” para  $n$  observações:

$$h_{otm} = \left( \frac{24 \times \sqrt{\pi}}{n} \right)^{\frac{1}{3}} \quad (3.17)$$

A largura  $h$  não é o único parâmetro a determinar o formato de  $\hat{f}$ . Na Figura 3.2 é possível ver um histogramas com  $x_0 = 137.65$  (canto superior esquerdo),  $x_0 = 137.75$  (canto inferior esquerdo),  $x_0 = 137.85$  (canto superior direito) e  $x_0 = 137.95$  (canto inferior direito).

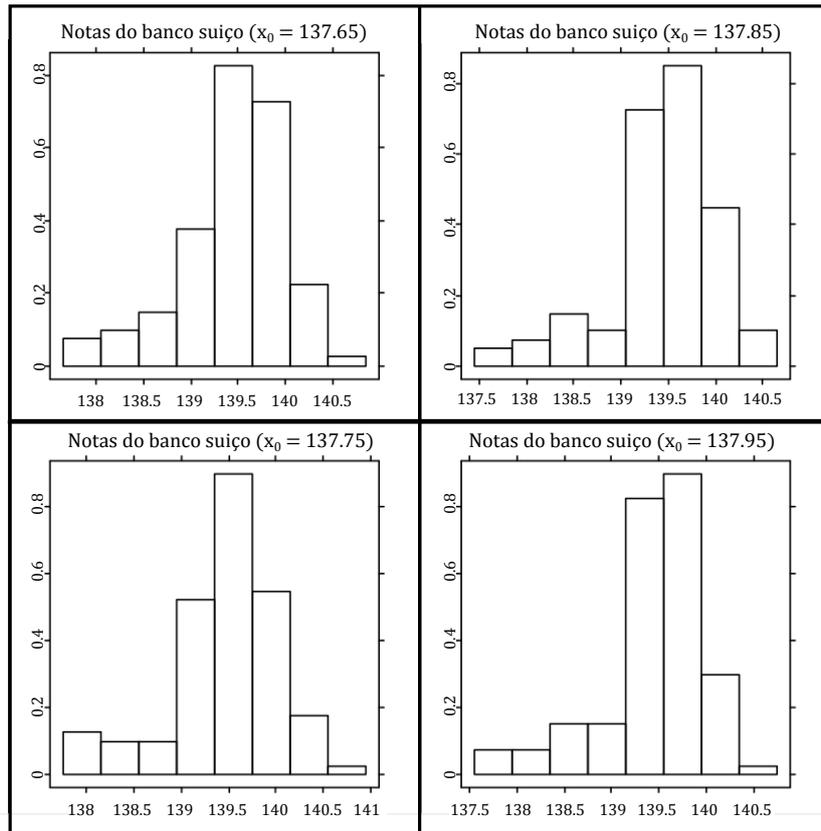


Figura 3.2: Diagonal das notas de banco falsificadas. Histogramas com  $h = 0.4$  e  $x_0 = 137.65$  (superior esquerdo),  $x_0 = 137.75$  (inferior esquerdo),  $x_0 = 137.75$  (superior direito) e  $x_0 = 137.85$  (inferior direito). Fonte: [3].

Todos os gráficos foram escalados em relação ao eixo  $y$  igualmente para permitir comparações. Apesar a largura fixa para  $h$ , a interpretação não é facilitada.

O deslocamento na origem  $x_0$  para 4 diferentes localizações criou 4 histogramas diferentes. Essa propriedade de histogramas contradiz severamente o objetivo de apresentar características de dados. Os mesmos dados são representados de maneira bastante diferente pelos 4 histogramas. Uma das correções para isso foi proposta por [59]: fazer a média dos histogramas deslocados. Os resultados são apresentados na Figura 3.3.

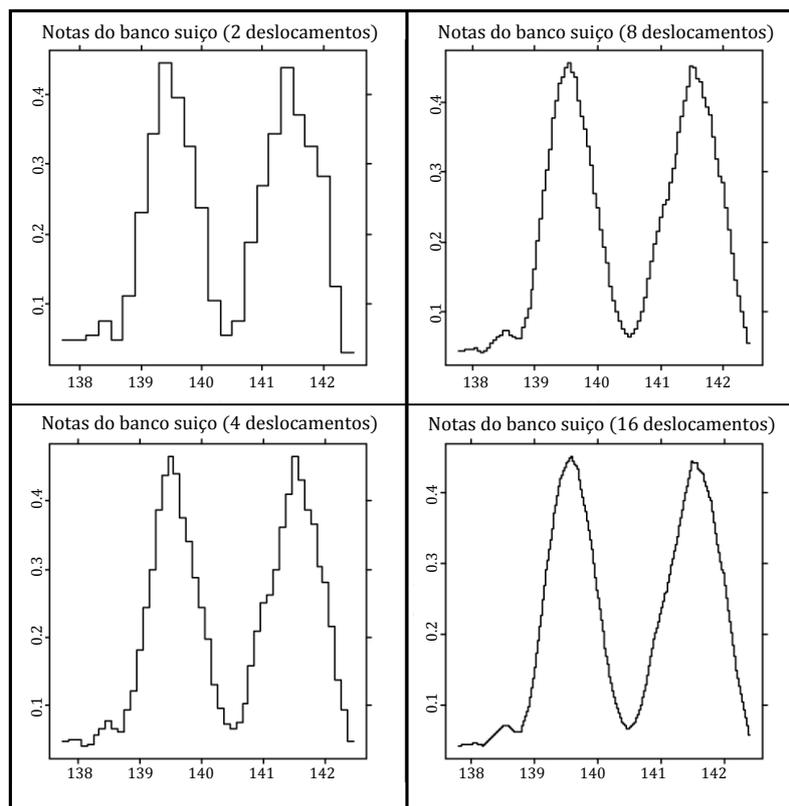


Figura 3.3: Histogramas ponderados pela média dos histogramas deslocados baseado em todos os tipos de notas (falsificadas e genuínas): 2 deslocamentos (superior esquerdo), 4 deslocamentos (inferior esquerdo), 8 deslocamentos (superior direito) e 16 deslocamentos (inferior direito). Fonte: [3].

Nessa figura todas as observações de notas do banco foram usadas. O histograma ponderado pelos deslocados não é mais dependente da origem e mostra claramente uma bimodalidade das diagonais das notas do banco da Suíça.

É interessante notar que as definições acerca de histogramas podem ser facilmente estendidas para várias dimensões. Em algumas situações os dados podem ser multidimensionais, por exemplo várias características de um objeto, e podem ser especificados intervalos de distribuição para cada uma dessas características. A Figura 3.4 mostra um exemplo gráfico de histograma bidimensional.

### 3.2.2 Densidades de kernels

As maiores dificuldades na estimação de histogramas são [3]:

- A determinação da largura  $h$ , que controla o formato do histograma;
- A escolha da origem dos *bins*  $x_0$ , que também influencia a forma;

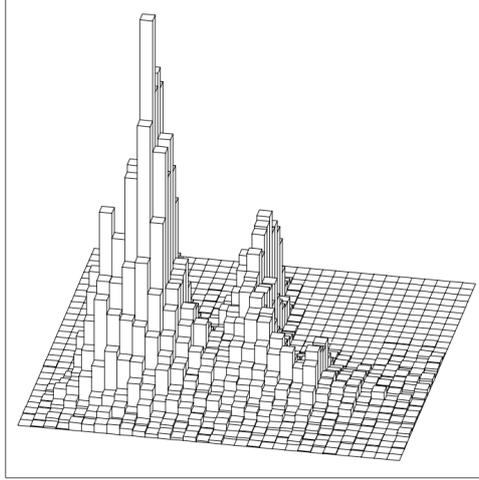


Figura 3.4: Representação gráfica de um histograma bidimensional. Fonte: [3].

- A perda de informações, já que as observações são substituídas pelo ponto central do intervalo que elas estão contidas;
- A função de densidade é comumente tida como suave, mas o histograma não é suave.

Para esse fim foi desenvolvida uma abordagem que evita as três últimas dificuldades. Primeiro, uma função *kernel* suave é usada ao invés de um bloco. Segundo, a função suave é centralizada diretamente sobre cada observação. Supondo  $x$  o valor central de um *bin*, o histograma pode ser reescrito como

$$\hat{f}_h(x) = n^{-1}h^{-1} \sum_{i=1}^n I(|x - x_i| \leq \frac{h}{2}). \quad (3.18)$$

Definindo  $K(u) = I(|u| \leq \frac{1}{2})$ , então a Equação 3.18 muda para

$$\hat{f}_h(x) = n^{-1}h^{-1} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right). \quad (3.19)$$

Essa é a forma geral de um estimador de kernel [3]. Permitir funções de kernel mais suaves, como o kernel quártico,

$$K(u) = \frac{15}{16}(1 - u^2)^2 I(|u| \leq 1), \quad (3.20)$$

e calcular  $x$  não só no centro dos *bins* gera um estimador de densidade de kernel.

Estimadores de kernel também podem ser derivados via média ponderada de pontos arredondados ou pela média de histogramas com diferentes origens. Alguns kernels comuns podem ser vistos na Tabela 3.1.

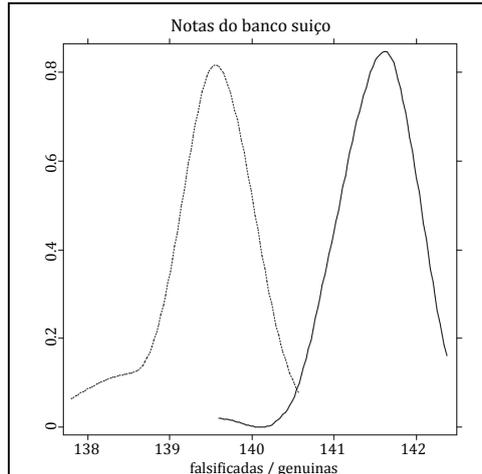


Figura 3.5: Densidades das diagonais de notas do banco genuínas e falsas. Fonte: [3].

Tabela 3.1: Funções de kernel mais comuns.

| Kernel $K(\bullet)$  | Nome         |
|--|--------------|
| $K(u) = \frac{1}{2} \times I( u  \leq 1)$                      | Uniforme     |
| $K(u) = (1 -  u ) \times I( u  \leq 1)$                        | Triangular   |
| $K(u) = \frac{3}{4} \times (1 - u^2) \times I( u  \leq 1)$     | Epanechnikov |
| $K(u) = \frac{15}{16} \times (1 - u^2)^2 \times I( u  \leq 1)$ | Quártico     |
| $K(u) = \frac{1}{\sqrt{2\pi}} e^{(-\frac{u^2}{2})} = \phi(u)$  | Gaussiano    |

É interessante destacar que a largura de banda  $h$  determina o grau de suavidade da estimação  $\hat{f}$ . A estimação de densidade por kernels é uma boa ferramenta descritiva para ver as modas, localização, de assimetria, caudas, assimetria, etc [3].

### 3.2.3 Histogramas no reconhecimento de padrões

Histogramas são utilizados amplamente como conjuntos de características para capturar os atributos estatísticos em várias tarefas de reconhecimento [6]. Por exemplo, [60] apresentou uma técnica para determinar objetos em uma cena usando histogramas das respostas de um vetor de operações lineares de vizinhança. Já [61] aborda o uso de vetores de histogramas para o reconhecimento *offline* de assinaturas.

O uso de histogramas como características para o reconhecimento *online* de assinaturas foi primeiramente proposto por [62], que utilizou métricas Euclidianas para realizar comparação. É possível encontrar outros três trabalhos que utilizam histogramas de ângulos, apresentados em [16, 29]. A mais recente proposta de uso de histogramas como

características no reconhecimento *online* de assinaturas pode ser vista em [6] onde várias características são extraídas com base em histogramas bidimensionais.

### 3.3 Métricas de Comparação

No campo de reconhecimento de padrões, quando as abordagens são baseadas em *template*, o processo de verificação/identificação de objetos é feito pelo *matching* de suas características em relação àquelas armazenadas na base de conhecimento. Esse processo também é chamado de *template matching*, onde o *template* tem a referência para comparação [2]. Várias métricas podem ser utilizadas para esse fim, como, por exemplo, distância euclidiana, riemanniana, DTW (*Dynamic Time Warping*), etc. Essa seção retratará algumas dessas métricas.

#### 3.3.1 Distância Euclidiana

O comprimento quadrado de um vetor  $x = [x_1, x_2]$  é a soma dos quadrados de suas coordenadas. E a distância do vetor  $x$  para um vetor  $y = [y_1, y_2]$  é obtida aplicando o teorema de Pitágoras à diferença de suas coordenadas,  $d_{x,y} = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$ . A Figura 3.6 retrata geometricamente esse fato.

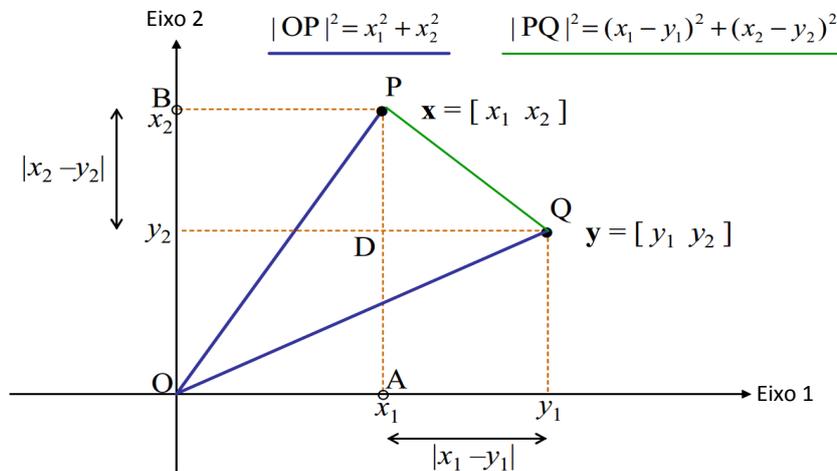


Figura 3.6: Distância Euclidiana de 2 vetores bidimensionais.

Essa noção pode ser estendida para quantas dimensões forem necessárias. Dado um vetor, em um espaço euclidiano, representado por  $J$  dimensões, tem sua distância dada por [63, 64]

$$d_{x,y} = \sqrt{\sum_{j=1}^J (x_j - y_j)^2} \quad (3.21)$$

Assim, essa equação generaliza a noção de distância para espaços multidimensionais e é chamada de Distância Euclidiana.

Para o caso da comparação de perfis relatado anteriormente, um conjunto de características pode ser visualizado como um vetor em um espaço  $J$ -dimensional. Portanto, para comparar dois vetores desse tipo, basta calcular o valor dado pela Equação 3.21. Sob essa perspectiva, o valor  $d_{x,y}$  é a medida de similaridade/dissimilaridade desses dois conjuntos de características.

### 3.3.2 Distância de Manhattan

A distância de Manhattan, considerada por Hermann Minkowski no século XIX, é uma forma de geometria em que a usual métrica da geometria euclidiana é substituída por uma nova métrica em que a distância entre dois pontos é a soma das diferenças absolutas de suas coordenadas. A Figura 3.7 faz um comparativo entre a distância de Manhattan, também conhecida como distância  $L_1$ , e a distância Euclidiana.

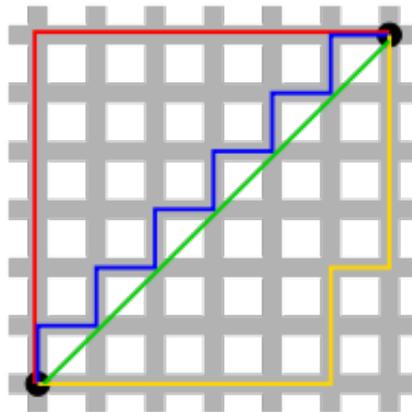


Figura 3.7: Comparação entre a distância Euclidiana (em verde) com a distância de Manhattan (outras cores) para dois pontos arbitrários.

Formalmente, a distância de Manhattan entre dois pontos em um espaço euclidiano com sistema de coordenadas cartesianas fixo é a soma dos comprimentos das projeções do segmento de reta que liga os pontos sobre os eixos coordenados. Assim, a distância  $L_1$  entre um vetor  $x = [x_1, x_2, \dots, x_n]$  e um vetor  $y = [y_1, y_2, \dots, y_n]$  num espaço  $n$ -dimensional com sistema de coordenadas cartesianas fixo é dada por [63]:

$$L_1(x, y) = \|x - y\|_1 = \sum_{i=1}^n |x_i - y_i| \quad (3.22)$$

A distância de Manhattan é um caso específico da distância de Minkowski. Pode-se generalizar essa noção de distância, também conhecida como distância  $L_p$  [64]:

$$L_p(x, y) = \left( \sum_{i=1}^n |x_i - y_i|^p \right)^{\frac{1}{p}} \quad (3.23)$$

Além de métrica para comparação de características de amostras, a distância de Manhattan é usada para medir distâncias no xadrez. Por fim, ao resolver sistemas de equações lineares indeterminados, o termo de regularização do vetor de parâmetros é expressado em termos dessa distância. Essa abordagem aparece em sistemas de recuperação de sinais, na área de compressão sensível.

### 3.3.3 DTW - Técnica da Deformação Dinâmica do Tempo

A DTW (do inglês *Dynamic Time Warping*) é uma técnica bastante conhecida, utilizada para encontrar o alinhamento ótimo entre duas sequências sob certas restrições. As duas sequências são deformadas de uma forma não linear para ser feito o *matching* das duas. Originalmente, a DTW era usada para comparar diferentes padrões de fala no reconhecimento automático de fala. Em campos como mineração de dados e recuperação de informações, a DTW tem sido aplicada com sucesso para lidar com as deformações temporais e velocidades diferentes de dados dependentes do tempo [4].

#### DTW Clássica

O objetivo da DTW é comparar duas sequências,  $X = (x_1, x_2, \dots, x_N)$  de comprimento  $N \in \mathbb{N}$  e  $Y = (y_1, y_2, \dots, y_M)$  de comprimento  $M \in \mathbb{N}$  [4, 65].

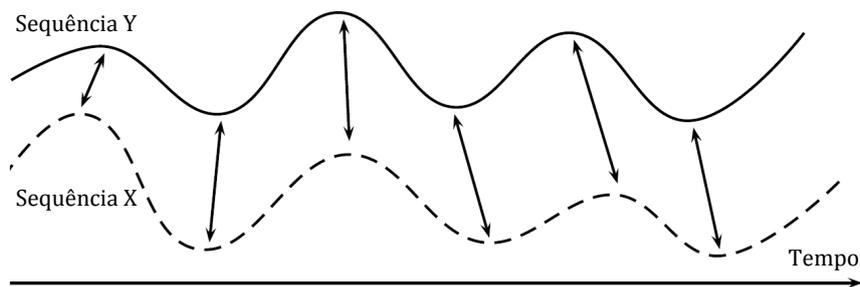


Figura 3.8: Alinhamento de duas sequências no tempo. Pontos alinhados são indicados pelas setas duplas. Fonte: [4].

Essas sequências podem ser sequências de características amostradas de forma equidistante. Assim, dado um espaço de características  $\mathcal{F}$ , com  $x_n, y_m \in \mathcal{F}$  para  $n \in [1 : N]$  e

$m \in [1 : M]$ , será necessário uma função de custo, também chamada de medida local de distância, definida como:

$$c : \mathcal{F} \times \mathcal{F} \rightarrow \mathbb{R}_{\geq 0} \quad (3.24)$$

Tipicamente,  $c(x, y)$  é pequena (baixo custo) se  $x$  e  $y$  são similares, e grande (alto custo) caso contrário [65]. Ao avaliar a função de custo local para cada par de elementos das sequências  $X$  e  $Y$ , obtém-se uma matriz de custos  $C \in \mathbb{R}^{M \times N}$  definida por  $C(n, m) = c(x_n, y_m)$ , conforme mostra a Figura 3.9.

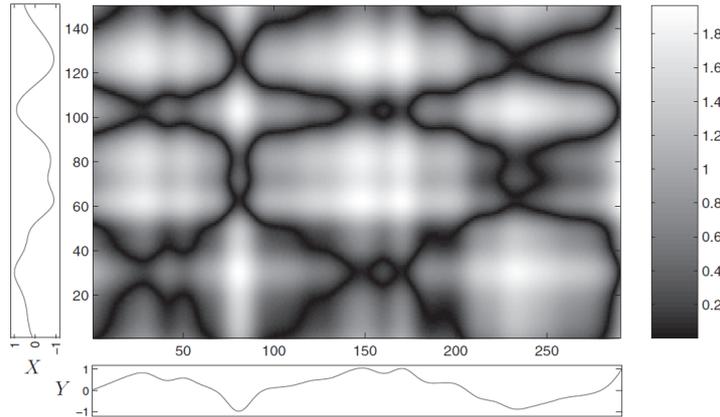


Figura 3.9: Matriz de custos da comparação de duas sequências reais  $X$  e  $Y$ , usando a distância de Manhattan como função de custo  $c$ . Regiões de baixo custo são indicadas pelas cores escuras e regiões de alto custo são indicadas por cores claras. Fonte: [4].

A meta é encontrar um alinhamento entre  $X$  e  $Y$  que tenha custo mínimo. Intuitivamente, um alinhamento ótimo passará pelos “vales” de custo mínimo dentro da matriz  $C$ , conforme mostra a Figura 3.10. Esse alinhamento é chamado de caminho de deformação. De maneira formal, um caminho de deformação  $(N, M)$  é uma sequência  $p = (p_1, p_2, \dots, p_L)$  com  $p_l = (n_l, m_l) \in [1 : N] \times [1 : M]$  para  $l \in [1 : L]$  satisfazendo as seguintes condições:

1. Condição de fronteira:  $p_1 = (1, 1)$  e  $p_L = (N, M)$ ;
2. Condição de monotonicidade:  $n_1 \leq n_2 \leq \dots \leq n_L$  and  $m_1 \leq m_2 \leq \dots \leq m_L$ ;
3. Condição de passo:  $p_{l+1} - p_l \in \{(1, 0), (0, 1), (1, 1)\}$  para  $l \in [1 : L - 1]$ .

O custo total  $c_p(X, Y)$  de um caminho de deformação  $p$  entre  $X$  e  $Y$  com respeito à métrica de custo local  $c$  é definido como

$$c_p(X, Y) = \sum_{l=1}^L c(x_{n_l}, y_{m_l}). \quad (3.25)$$

Um caminho de deformação ótimo entre  $X$  e  $Y$  é um caminho  $p^*$  que tem o custo mínimo dentre todos os custos possíveis. Assim, a distância DTW,  $DTW(X, Y)$ , entre  $X$

e  $Y$  é definida como o custo total de  $p^*$ :

$$\begin{aligned} DTW(X, Y) &= c_{p^*}(X, Y) \\ DTW(X, Y) &= \min\{c_p(X, Y) | p \text{ é um caminho de deformação } (N - M)\} \end{aligned} \quad (3.26)$$

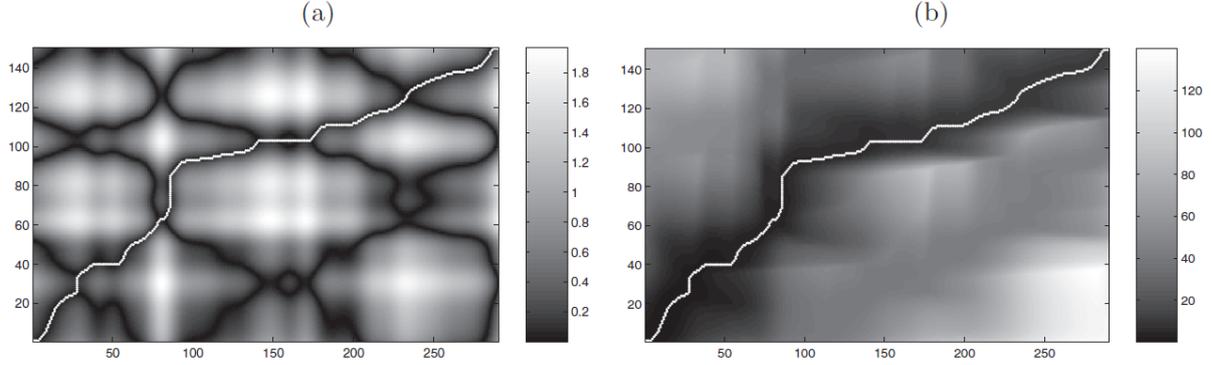


Figura 3.10: (a) Matriz de custo  $C$  e (b) matriz de custo acumulado  $D$  com um caminho de deformação ótimo  $p^*$  (caminho em branco). Fonte: [4].

### Propriedade utilizada no algoritmo da DTW

Para determinar o caminho ótimo  $p^*$  poderia testar-se todas as possibilidades de caminho entre  $X$  e  $Y$ . Tal procedimento, todavia, levaria a uma complexidade computacional que é exponencial nos comprimentos  $N$  e  $M$ . É possível definir um algoritmo que é  $\mathcal{O}(NM)$ , baseado em programação dinâmica [4]. Considere uma sequência de prefixos  $X(1:n) = (x_1, \dots, x_n)$  para  $n \in [1:N]$  e  $Y(1:m) = (y_1, \dots, y_m)$  para  $m \in [1:M]$  e seja

$$D(n, m) = DTW(X(1:n), Y(1:m)) \quad (3.27)$$

Os valores de  $D(n, m)$  definem uma matriz  $D$   $N \times N$ , também designada matriz de custo acumulado. Obviamente,  $D(N, M) = DTW(X, Y)$ . A matriz de custo acumulado  $D$  satisfaz as seguintes identidades:

$$\begin{aligned} D(n, 1) &= \sum_{k=1}^n c(x_k, y_1) \text{ para } n \in [1:N], \\ D(1, m) &= \sum_{k=1}^m c(x_1, y_k) \text{ para } m \in [1:M], \\ D(n, m) &= \min\{D(n-1, m-1), D(n-1, m), D(n, m-1)\} + c(x_n, y_m) \end{aligned} \quad (3.28)$$

para  $1 < n \leq N$  e  $1 < m \leq M$ . Em particular,  $DTW(X, Y) = D(N, M)$  pode ser computada com  $\mathcal{O}(NM)$  operações. O algoritmo para implementar a DTW tira vantagem dessa propriedade para calcular o caminho e o custo ótimo.

## Aplicações e usos da DTW

A DTW é utilizada como métrica de comparação em inúmeros trabalhos. É particularmente útil quando as sequências não tem o mesmo tamanho, visto que as métricas vistas nas seções anteriores, Euclidiana e Manhattan, não seriam aplicáveis diretamente. É interessante notar que dentre as propostas com DTW, nem todas abordam o algoritmo clássico da DTW, apresentado anteriormente. Variações incluem alterações na condição do tamanho do passo, dos pesos locais, restrições globais dos caminhos de deformação e aproximações com abordagens em multiescala.

Além de reconhecimento de fala, a DTW encontra inúmeras aplicações em uma grande variedade de campos, incluindo mineração de dados, recuperação de informações, bioinformática, engenharia química, processamento de sinais, robótica, computação gráfica, entre outros. Basicamente, qualquer dado pode ser transformado em uma sequência linear de características que podem ser analisadas com a DTW, o que inclui dados do tipo texto, vídeo, áudio, ou séries temporais gerais [4].

Quando a visão volta-se para o processamento de sinais, especialmente na área de biometria, é possível encontrar vários trabalhos que utilizam a DTW como métrica de dissimilaridade para a comparação de vetores de características. [66] utiliza a DTW em conjunto com a quantização de vetores para a tarefa de reconhecimento automático de assinaturas *online*. [28] também aborda a mesma tarefa com o uso de uma DTW com pesos alterados. Uma proposta baseada em uso de mapas auto-organizáveis para o reconhecimento automático de palavras cursivas isoladas *online* pode ser vista em [5], onde os autores utilizam a DTW clássica para comparação das sequências.

O próximo capítulo fará a descrição metodológica do trabalho proposto, apresentando a estrutura do sistema desenvolvido.

# Capítulo 4

## Metodologia

Esse capítulo abordará os detalhes metodológicos utilizados no trabalho apresentado. Serão discutidos pontos acerca da coleta do dados, a criação da base de dados, o pré-processamento dos dados, e a estratégia para extração de características. Além disso, serão abordados os métodos utilizados para caracterização dos perfis, classificação das amostras e os critérios para decisão do sistema.

### 4.1 A estrutura geral do sistema

Nesse trabalho, é proposto um sistema de verificação automática de escritores baseado em ou amostras de palavras cursivas isoladas ou assinaturas. Ambas as tarefas supõe um cenário *online*, onde há informações espaço-temporais disponíveis. Quanto ao reconhecimento de palavras cursivas isoladas, o cenário foi avaliado com o uso de dispositivo móvel.

A arquitetura geral do sistema é composta de duas tarefas principais: o cadastro e a verificação/autenticação. Na primeira, uma pessoa é registrada no sistema provendo  $M$  amostras de referências para criar um perfil. Na última, a pessoa será autenticada no sistema fazendo uma amostra e indicando a suposta identidade. A Figura 4.1 mostra o diagrama de fluxo de todas as principais tarefas realizadas pelo sistema proposto. Todas as etapas serão abordadas em detalhes nas próximas seções.

A etapa de aquisição de dados diz respeito ao processo de coleta das assinaturas, os meios para adquirir os dados e a posterior criação da base de dados. Já a etapa de pré-processamento está vinculada aos processamentos feitos nas informações antes de sua efetiva utilização.

A fase posterior refere-se à extração das características utilizadas para caracterizar as amostras dentro do sistema. O fluxo, em seguida, é determinado pela finalidade do

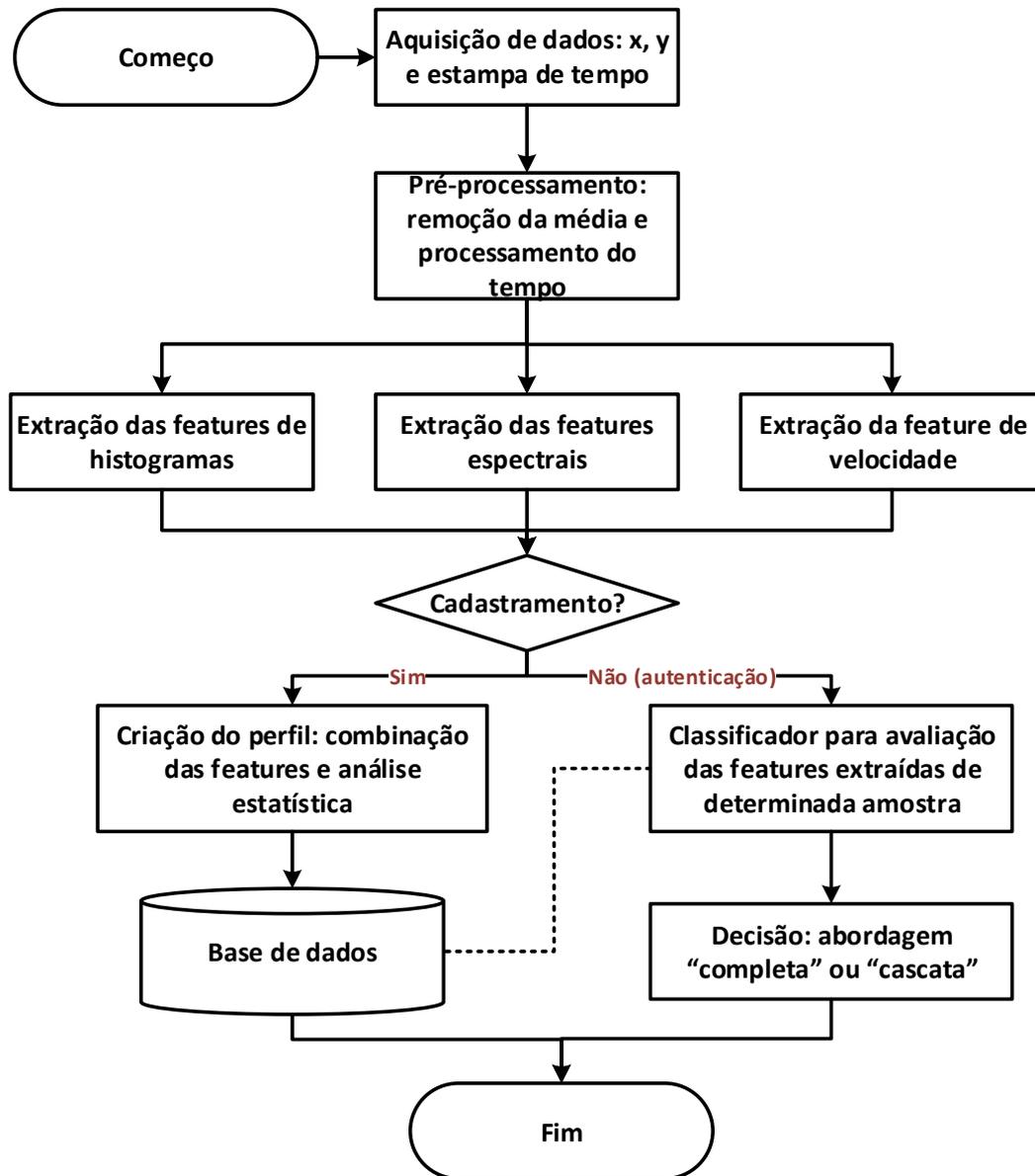


Figura 4.1: Diagrama de fluxo contendo as principais tarefas realizadas pelo sistema proposto.

que está ocorrendo no sistema naquele momento: se é para o cadastramento de um novo usuário ou se é para a autenticação/verificação de determinado usuário.

Se for escolhido o cadastramento de usuário, as  $M$  amostras de referência irão compor o perfil. Esse perfil caracterizará o usuário dentro do sistema, e será composto pelas características extraídas e por alguns dados de análises estatísticas, visando caracterizar

melhor as amostras cedidas pelo usuário. Em seguida, essas informações são guardadas na base de dados e o processo geral se encerra.

No caso de verificação/autenticação de amostras, o próximo passo é a classificação das características. Dada a amostra cedida e a suposta identidade, o classificador fará comparações e, com base nos resultados retornados pelo classificador, o processo de decisão será feito. A amostra pode ser rejeitada ou tida como verdadeira. Para o processo de decisão foram propostas duas abordagens. Maiores detalhes serão vistos nas seções posteriores.

## 4.2 A etapa de aquisição de dados

A aquisição de dados é o passo inicial de toda a arquitetura e envolve os dados que serão analisados pelo sistema. Nesse estudo, o principal foco é em palavras cursivas isoladas escritas a mão, capturadas por dispositivos sensíveis ao toque, sem o uso de canetas ativas. Também são avaliadas assinaturas escritas à mão.

Canetas ativas são aquelas capazes de adquirir outros tipos de informações, como por exemplo os ângulos de altitude e o azimute, ou mesmo informação de pressão com uma maior precisão. A Figura 4.2 mostra o significado geométrico do azimute e da altitude.

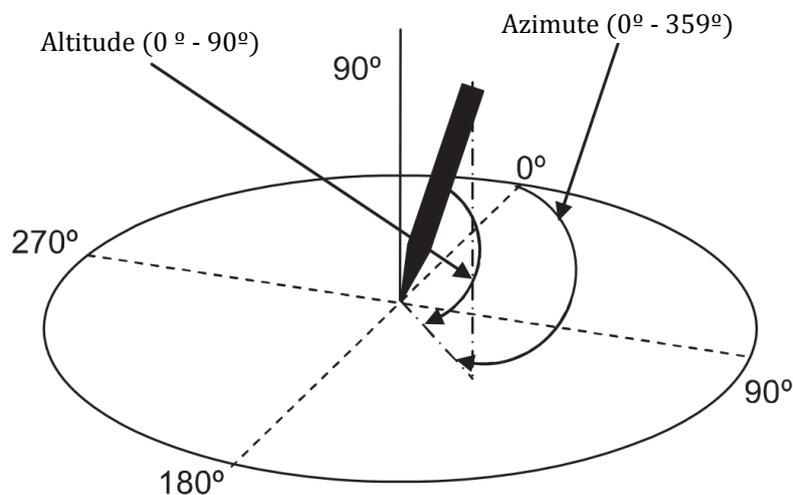


Figura 4.2: Ângulos de azimute e altitude de uma caneta em relação ao plano da superfície de escrita. Fonte: [5].

No melhor de nosso conhecimento, não há nenhuma base de dados de palavras cursivas isoladas em dispositivos móveis disponível. Dessa maneira, foi criada uma base de dados denominada LISA-01 (Laboratório de Imagens, Sinais e Áudio). Além da LISA-01, para abordar o cenário de assinaturas, foi utilizada a base de dados MCYT-100 [67]. Detalhes das bases de dados serão relatados no Capítulo 5.

### 4.3 Pré-processamento

A etapa de pré-processamento visa preparar os dados antes que eles sejam efetivamente considerados para análise no sistema. Uma prática bastante comum nessa fase é fazer a reamostragem dos dados a fim de remover pontos redundantes para agilizar as comparações. Também é usual remover a dependência do tempo. Todavia, a reamostragem também implica em significativa perda de informação visto que distorce a característica de velocidade da amostra. Outro problema com a reamostragem é que pontos críticos da amostra podem ser perdidos [28].

Embora passos de pré-processamento sejam bastante úteis no reconhecimento de objetos, eles podem diminuir a performance de sistemas de autenticação biométrica [7, 28]. Por essa razão, foi decidido reduzir ao mínimo o pré-processamento feitos nos dados. Dois processos foram realizados aqui, o processamento da estampa de tempo e a remoção da média dos dados.

Os dados coletados podem ser vistos como um vetor da forma

$$V = [estampaTempo_j, \hat{x}_j, \hat{y}_j] \quad (4.1)$$

onde  $1 \leq j \leq N$ , sendo  $N$  o número de pontos adquiridos para uma certa palavra/assinatura,  $estampaTempo_j$  é o tempo, em milissegundos, desde que o sistema operacional do aplicativo de captura de amostras foi iniciado, em que o ponto  $j$  foi capturado e o par  $(\hat{x}_j, \hat{y}_j)$  representa a componente horizontal e vertical do ponto, respectivamente.

É mais útil utilizar uma informação de tempo onde  $tempo_0 = 0$ . Portanto,  $tempo$  é o vetor diferença do vetor  $estampaTempo$  com o primeiro valor desse mesmo vetor, o  $estampaTempo_0$ :

$$tempo_j = [estampaTempo_j - estampaTempo_0]. \quad (4.2)$$

Com essa abordagem, é possível obter um tempo absoluto.

Uma das características que serão utilizadas nesse trabalho são baseadas na análise de Fourier. Com o fim de obter invariância à translação para essas características, uma prática comum é extrair a média dos dados. Essa prática não afetaria nenhum dos outros dois grupos de características, as dos histogramas e a característica de velocidade. Portanto, dado o vetor  $\hat{x}$ , realiza-se a remoção da média, obtendo o vetor  $x$ :

$$x_j = \hat{x}_j - \frac{1}{N} \times \sum_{k=1}^N x_k, \quad (4.3)$$

onde  $1 \leq j \leq N$ , sendo  $N$  o total de pontos da amostra. O mesmo processo é aplicado a  $\hat{y}$ :

$$y_j = \hat{y}_j - \frac{1}{N} \times \sum_{k=1}^N y_k. \quad (4.4)$$

Obtem-se assim, um conjunto de coordenadas espaciais  $(x, y)$  com média 0. Dessa maneira, após o pré-processamento o vetor de dados é dado por:

$$V = [tempo_j, x_j, y_j]. \quad (4.5)$$

Esses foram os únicos passos de processamento inicial dos dados, antes de sua efetiva utilização no sistema. A próxima seção irá explicar o processo de extração de características.

## 4.4 Extração de características

Dois tipos de características podem ser usadas para a verificação de texto: funções ou parâmetros (locais ou globais). Quando uma função é usada, a assinatura é usualmente caracterizada em termos de uma função do tempo, cujos valores constituem o conjunto de características. Quando características de parâmetros são usadas, a amostra é caracterizada como um vetor de elementos onde cada elemento representa o valor de uma característica.

Embora características de funções permitam uma melhor performance do que os parâmetros, elas usualmente requerem técnicas pesadas para realizar a comparação. Como o cenário proposto nesse trabalho envolve o uso de dispositivos móveis, que não dispõe de grande capacidade computacional, o uso de parâmetros/características globais se mostra mais apropriado, visto que nesses casos não há necessidade de procedimentos de comparação complexos.

Dado o vetor de características  $V = [tempo_j, x_j, y_j]$ ,  $1 \leq j \leq N$ , é possível derivar outras informações importantes. Por exemplo, é possível calcular o ângulo/direção tangencial:

$$\theta_j = \tan^{-1} \frac{y_{j+1} - y_j}{x_{j+1} - x_j} \quad (4.6)$$

com  $-180^\circ \leq \theta_j \leq 180^\circ$ . Por exemplo, também é possível calcular uma magnitude associada a esse ponto, dada por:

$$mag_j = \sqrt{(x_{j+1} - x_j)^2 + (y_{j+1} - y_j)^2} \quad (4.7)$$

Um outro tipo de característica derivada é a velocidade em ambas as direções, horizontal e vertical, conforme segue:

$$(v_x)_j = \frac{x_{j+1} - x_j}{tempo_{j+1} - tempo_j} \quad (4.8)$$

e

$$(v_y)_j = \frac{y_{j+1} - y_j}{tempo_{j+1} - tempo_j} \quad (4.9)$$

As características globais utilizadas nesse sistema são os descritores de Fourier, os *bins* de histogramas de ângulos e uma característica de velocidade. As próximas seções irão relatar cada uma dessas características.

#### 4.4.1 Histogramas de Ângulos

Histogramas são características comumente utilizadas para capturar características estatísticas em várias tarefas de reconhecimento, por exemplo, reconhecimento de ações humanas [68, 69] e verificação *offline* de assinaturas [70]. Trabalhos aplicando histogramas para reconhecimento *online* de assinaturas podem ser vistos em [6, 29].

Nesse trabalho são propostos dois métodos para o uso de histogramas. Uma palavra ou assinatura *online* é representada não por um, mas por um conjunto de histogramas de ângulos. A Figura 4.3 ilustra o mapeamento de uma amostra em termos de seus ângulos e a distribuição dos pontos dentre os *bins* de um histograma.

#### O cálculo do histograma de ângulos para uma amostra

A criação dos histogramas é dada de maneira bastante intuitiva a partir dos ângulos. Considere o conjunto  $\theta = \{\theta_1, \theta_2, \dots, \theta_{N-1}\}$ , com  $N$  sendo o total de pontos da amostra.  $b^i$  representa  $i$ -ésimo *bin* do histograma, com  $1 \leq i \leq \Omega$ , dado que  $\Omega$  é o total de *bins* do histograma. O cálculo do histograma, dado um  $\theta_j$ , com  $1 \leq j \leq N - 1$ , é dado pela seguinte regra:

$$-\pi + 2\pi \frac{i-1}{\Omega} \leq \theta_j < -\pi + 2\pi \frac{i}{\Omega} \rightarrow b^i = b^i + 1. \quad (4.10)$$

Se os valores obtidos em cada  $b^i$  fossem utilizados diretamente como características do sistema, eles não seriam invariantes ao tamanho da amostra. Para evitar esses problemas, utiliza-se a distribuição dos ângulos. Dessa forma, dado um  $\hat{b}^i$ , com  $1 \leq i \leq \Omega$ , obtêm-se um  $b^i$  da seguinte forma:

$$b^i = \frac{\hat{b}^i}{N-1} \quad (4.11)$$

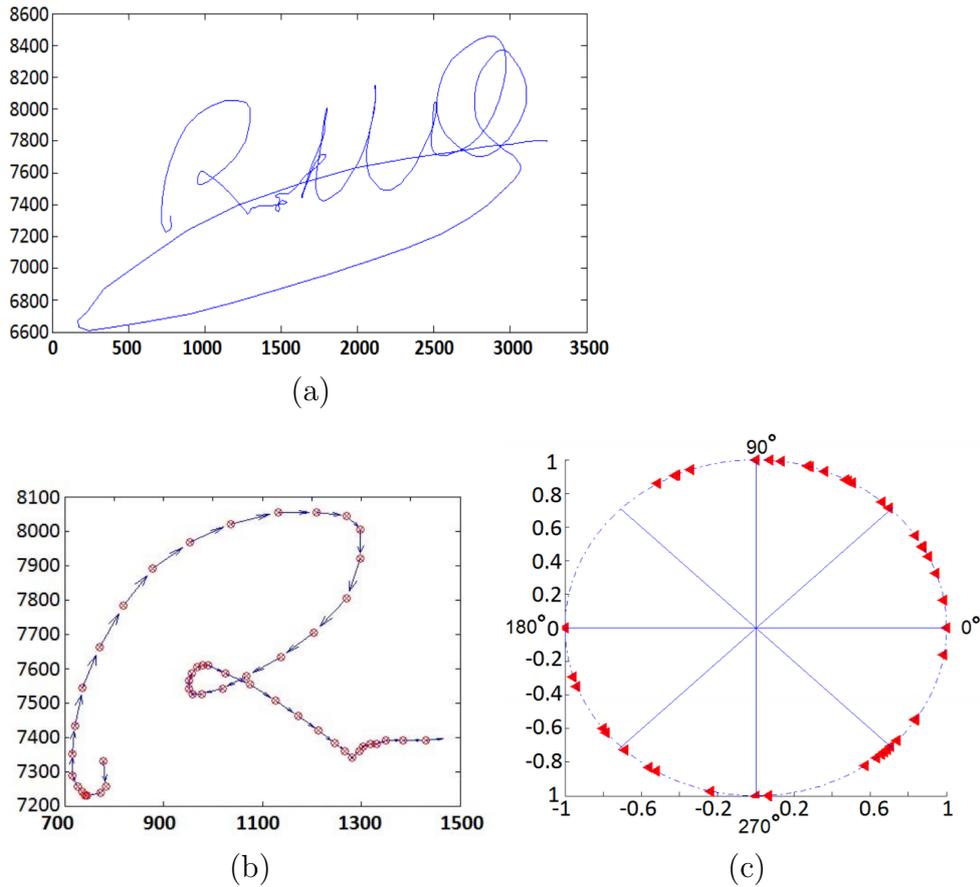


Figura 4.3: Processo de derivar um histograma de ângulos: (a) a assinatura original, (b) a sequência de vetores dos 60 primeiros pontos da assinatura e (c) a distribuição dos ângulos derivados do vetor em (b). Fonte: [6].

onde  $N$  é o total de pontos da amostra e  $0 \leq b^i \leq 1$ . Com isso obtêm-se um conjunto de valores de *bins* que são invariantes ao tamanho da amostra. O diagrama de fluxo correspondente a esse procedimento é dado pela Figura 4.4.

### Calculando múltiplos histogramas

Na abordagem aqui desenvolvida, uma palavra/assinatura é representada por um grupo de histogramas de  $\theta$ , variável cujos valores estão limitados a  $[-\pi, \pi]$ . O conjunto de histogramas é definido como

$$H = \{h^1, h^2, \dots, h^K\} \quad (4.12)$$

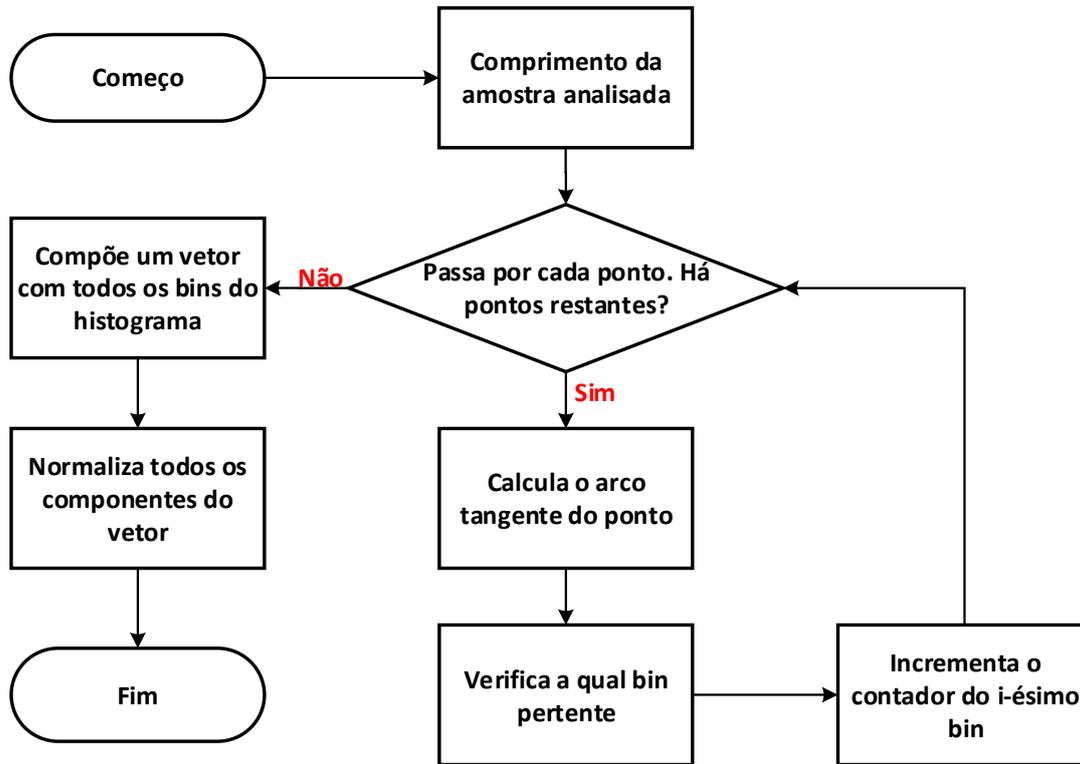


Figura 4.4: Diagrama de fluxo para o cálculo de um histograma de ângulos.

onde  $K \in \mathbb{N}^*$  é o total de histogramas e cada histograma  $h^i$ ,  $1 \leq i \leq K$ , consiste em uma contagem de ângulos divididos em  $b^i$  bins, com

$$b^1 < b^2 < \dots < b^K. \quad (4.13)$$

No restante desse documento, o histograma  $h^i$  será referido como o histograma de nível  $i$ . O sistema irá trabalhar com diferentes palavras/assinaturas de diferentes usuários. A ideia é que a representação ótima das palavras/assinaturas muda em cada cenário. Isso porque um número muito baixo de bins dentro dos histogramas pode não ser suficiente para discriminar a assinatura, enquanto um número muito alto pode deixar a representação instável. A questão é que esses valores de máximo e mínimo desejável podem variar para cada amostra.

A intenção, com o uso de vários histogramas, é tentar abarcar essa representação ótima para a amostra em questão. Essa multiplicidade de representações poderia melhorar a performance do sistema. A forma pela qual essas representações serão utilizadas será explicada nas seções subsequentes.

## 4.4.2 Descritores de Fourier

O outro tipo de características utilizadas nesse trabalho são baseadas na análise de Fourier. Os descritores obtidos são inspirados no trabalho de [7]. Análises espectrais são aplicadas nas palavras/assinaturas através do algoritmo da FFT (*Fast Fourier Transform*), que propõe uma maneira mais eficiente de calcular a DFT (*Discrete Fourier Transform*).

A vantagem de usar o domínio de Fourier é a possibilidade de representar uma amostra de forma compacta, usando um número fixo de coeficientes, o que mantém a ideia de realizar comparações de maneira ágil.

### Estimação de Drift

Considere o sinal que descreve as coordenadas espaciais (após a remoção da média) de uma amostra arbitrária:

$$V = [x_j, y_j]. \quad (4.14)$$

Em amostras que vão da esquerda para a direita,  $x_t$  tem *drift* na medida em que o tempo aumenta e o mesmo pode ser dito de assinaturas que vão do topo para base com relação ao  $y_t$ . Este desvio nas medidas ocorre devido à variabilidade natural na escrita de cada um [7]. Vários métodos podem ser aplicados para a remoção de *drift*, sendo possível encontrar pesquisas buscando realizar estas estimações nos mais variados tipos de sinais [71, 72, 73]. O desafio é caracterizar, da melhor forma possível, o *drift* em uma amostra, dado o fato de que este *drift* não necessariamente acontece da mesma forma em todas as amostras, mas que pode ter alguma característica geral.

Um dos modelos utilizados para a estimação de *drift* são as regressões lineares por mínimos quadrados. Dado um sinal discreto  $x$  de comprimento  $N$ , a versão com *drift* removido  $x'$  pode ser calculada como [7]:

$$x' = x - \beta \times (t - \bar{t}), \quad (4.15)$$

onde

$$\beta = \frac{\sum_{i=1}^N (x_i \times t_i) - N \times \bar{x} \times \bar{t}}{\sum_{i=1}^N (t_i^2) - N \times \bar{t}^2} \quad (4.16)$$

onde  $\bar{x}$  é a média do sinal  $x$ .

A estimação de *drift*, quando não feita de maneira adequada, pode deformar o sinal em questão. Embora na proposta original dos descritores [7] a remoção do *drift* ajudou

a melhorar o desempenho do sistema, o mesmo não foi comprovado neste trabalho. A análise da estimação de *drift* será detalhada no Capítulo 5.

## Aplicação da Transformada de Fourier

Qualquer função periódica pode ser expressada como séries de funções senoides de amplitudes variadas, as chamadas Séries de Fourier [7]. Se o sinal é periódico com frequência fundamental  $\omega$ , as frequências dos componentes senoidais, que compõe o sinal, são múltiplos inteiros de  $\omega$  e são chamados harmônicos. A Transformada de Fourier é usada para encontrar a amplitude do componente harmônico, que é chamado de espectro de frequência do sinal. Dessa forma, quando aplica-se a transformada o sinal é convertido do domínio do tempo/espaço para o domínio da frequência.

Considerando novamente os componentes espaciais do sinal de entrada  $V$ , definido por  $V = [x_j, y_j]$ , a ideia é aplicar a transformada de Fourier em cada um dos sinais,  $x$  e  $y$ , isoladamente. Dado o sinal discreto  $x$ , de comprimento  $N$ , calcula-se o sinal de coeficientes de Fourier

$$\mathcal{F}_k(x) = \frac{1}{N} \sum_{i=0}^{N-1} x_i \times e^{\frac{-2\pi i k t}{N}} \quad (4.17)$$

com  $0 \leq k \leq N - 1$ .  $k$  indica a frequência do harmônico em particular e  $e^{i\theta} = \cos(\theta) + i \times \sin(\theta)$ . O mesmo se aplica ao sinal  $y$ , obtendo

$$\mathcal{F}_k(y) = \frac{1}{N} \sum_{i=0}^{N-1} y_i \times e^{\frac{-2\pi i k t}{N}}. \quad (4.18)$$

Cada um dos elementos dos respectivos espectros é dado por um número complexo. Assim,  $\mathcal{F}_k(x) = a_k + i \times b_k$ . Dois componentes importantes ao utilizar descritores de Fourier são a magnitude e a fase, dados, respectivamente por:

$$|\mathcal{F}_k(x)| = \sqrt{a_k^2 + b_k^2} \quad (4.19)$$

e

$$\theta_k = \arctan^{-1} \frac{b_k}{a_k}. \quad (4.20)$$

É importante citar que o algoritmo utilizado para calcular a DFT nesse caso foi a FFT, um algoritmo bastante eficiente, com complexidade assintótica de  $\mathcal{O}(N \log N)$  para cada vetor de características.

## Obtendo invariância a translação, rotação e escala

Os coeficientes de Fourier extraídos são normalizados para obter os Descritores de Fourier, usando alguns passos de normalização. Em particular, os coeficientes de Fourier obtidos pela aplicação da Transformada de Fourier ao contorno da amostra  $(x_j, y_j)$  devem ser normalizados para obter invariância sobre a translação, rotação e escala da forma original.

Especificamente, a translação corresponde a adicionar um termo constante para cada ponto da forma original e afeta somente o primeiro coeficiente de Fourier. Assim, ao descartar  $\mathcal{F}_0(x)$ , pode-se obter invariância à translação no restante dos coeficientes.

A rotação da forma original resulta em mudança de fase em cada um dos coeficientes de Fourier. A invariância à rotação pode ser obtida ao usar somente a informação de magnitude da Transformada de Fourier. Alternativamente, cada coeficiente pode ser normalizado tal que a fase de um dos coeficientes (por exemplo  $\mathcal{F}_0(x)$ ) seja 0. Isso é equivalente a assumir uma rotação cônica que dá fase zero para  $\mathcal{F}_0(x)$ .

Por último, a invariância à escala da forma corresponde a multiplicar todos os valores de coordenadas da forma por um fator constante e resulta em todos os coeficientes de Fourier serem multiplicados pelo mesmo fator. Dessa maneira, a normalização por escala pode ser alcançada dividindo cada coeficiente da magnitude por uma das componentes, tipicamente  $|\mathcal{F}_1(x)|$ .

Uma amostra *online* deve mostrar comparação adequada às amostras de referência tanto na forma quanto nas propriedades dinâmicas a fim de ser aceita. Com os passos de normalização descritos acima, descartar o  $\mathcal{F}_0(x)$  e usar a magnitude do restante dos coeficientes, é possível obter a invariância à translação e rotação. Entretanto, a invariância à escala é um pouco mais complicada devido à informação implícita do tempo.

Se uma assinatura só é escalada no espaço, enquanto mantém o tempo de duração, dividir cada magnitude dos coeficientes por  $|\mathcal{F}_1(x)|$  alcança a normalização da escala. Todavia, para um caso mais geral, envolvendo variações na escala e no tempo, há duas possibilidades consideradas mais robustas do que a apresentada. A primeira é dividir cada coeficiente pela magnitude total do espectro de Fourier

$$magTotal = \sum_{k=0}^{N-1} |\mathcal{F}_k(x)| = \sum_{k=0}^{N-1} \sqrt{\mathcal{F}_k(x) \times \mathcal{F}_k^*(x)} \quad (4.21)$$

onde  $N$  é o comprimento da amostra,  $|\mathcal{F}_k(x)|$  é a magnitude do coeficiente complexo  $\mathcal{F}_k(x)$  e  $\mathcal{F}_k^*(x)$  é o conjugado complexo de  $\mathcal{F}_k(x)$ .

A segunda possibilidade de normalização é pela energia total do espectro de fourier, dada por

$$eTotal = \sum_{k=0}^{N-1} |\mathcal{F}_k(x)|^2. \quad (4.22)$$

O trabalho apresentado em [7] fez um comparativo dessas abordagens com os 3 métodos de normalização comumente aplicados e encontrou que a normalização pela amplitude total do espectro é melhor do que a normalização por  $|\mathcal{F}_1(x)|$  e pela energia total.

Usando a definição da Equação 4.21, os descritores de Fourier são dados por:

$$(F_x)_k = \frac{\mathcal{F}_k(x)}{magTotal} \quad (4.23)$$

para o sinal  $x$  e

$$(F_y)_k = \frac{\mathcal{F}_k(y)}{magTotal} \quad (4.24)$$

para o sinal  $y$ , com  $1 \leq k \leq \frac{N}{2}$ . Repare que  $k$  vai de 1 a  $\frac{N}{2}$  visto que metade dos coeficientes podem ser descartados devido à simetria do espectro da Transformada de Fourier.

## Normalizando o Comprimento dos Descritores

Devido à variação natural no processo de escrita/assinatura, amostras verdadeiras do mesmo usuário quase nunca terão o mesmo comprimento. A variação do comprimento resulta de representações do domínio de Fourier com número variável de componentes, caracterizando, assim, vetores de características de comprimentos variáveis.

Embora seja plausível eliminar os componentes de alta frequência, deixando somente os  $p$  primeiros coeficientes de Fourier, quando as assinaturas são de comprimentos diferentes, esses componentes não vão corresponder às mesmas frequência.

A fim de obter um número igual de descritores de Fourier que correspondam às mesmas frequências, é realizado o *padding* de cada amostra a ser comparada (as amostras de referência + amostra *query*) com zeros, para alinhar o comprimento com a maior assinatura no conjunto. Esse processo é realizado de maneira prévia à aplicação da transformada de Fourier. Esse processo é chamado de *padding* com zeros e não afeta a amplitude dos coeficientes de Fourier, mas impacta na resolução da frequência [7].

Uma perspectiva é que se há grande variação no comprimento das amostras e, circunstancialmente o comprimento padrão é pequeno em relação à media dos comprimentos das *queries* futuras, haverá uma deformação dessas amostras visto que deverão ser feitos truncamentos nesse caso. Portanto, ao invés de considerar o comprimento da maior amostra, esse *padding* poderia ser feito em um comprimento superior.

Dado o maior comprimento das amostras de referência,  $M$ , o *padding* poderia se feito em relação a um valor proporcional a  $M$ , dado por:

$$M' = a \times M \quad (4.25)$$

onde  $a \in \mathbb{N}$  e  $a > 1$ . Detalhes a respeito da estimação do parâmetro  $a$  serão mostrados no Capítulo 5.

### Suavização dos Descritores de Fourier

É feita a suavização dos descritores  $(F_k)_x$  e  $(F_k)_y$ . Esse processo é realizado ponderando dois descritores consecutivos. Isso para contabilizar as variações normais de tempo entre amostras genuínas que poderiam resultar em oscilações em harmônicos vizinhos.

Assim, dado  $(F_k)_x$ , pode-se obter os novos coeficientes por:

$$(F_x)_k = \frac{1}{2} \times ((F_x)_k + (F_x)_{k+1}). \quad (4.26)$$

e o mesmo se aplica a  $(F_k)_y$ :

$$(F_y)_k = \frac{1}{2} \times ((F_y)_k + (F_y)_{k+1}). \quad (4.27)$$

A Figura 4.5 apresenta um exemplo de assinaturas, seu perfil em  $y$  e os descritores de Fourier obtidos.

### 4.4.3 Característica de Velocidade

Informação de velocidade também é usada. Já foi mostrado que as características de velocidade são discriminativas e úteis no que se refere à verificação de assinaturas [25].

A ideia é usar uma única característica que contenha informações de velocidade e aceleração, visto que o cenário a ser considerado envolve dispositivos com baixa capacidade computacional. Dado o vetor de entrada,  $V = [tempo_j, x_j, y_j]$ , com  $1 \leq j \leq N$ , tal que  $N$  é o comprimento da amostra em questão, pode-se calcular as componentes de velocidade

$$(v_x)_j = \frac{x_{j+1} - x_j}{tempo_{j+1} - tempo_j}, \quad (4.28)$$

$$(v_y)_j = \frac{y_{j+1} - y_j}{tempo_{j+1} - tempo_j}. \quad (4.29)$$

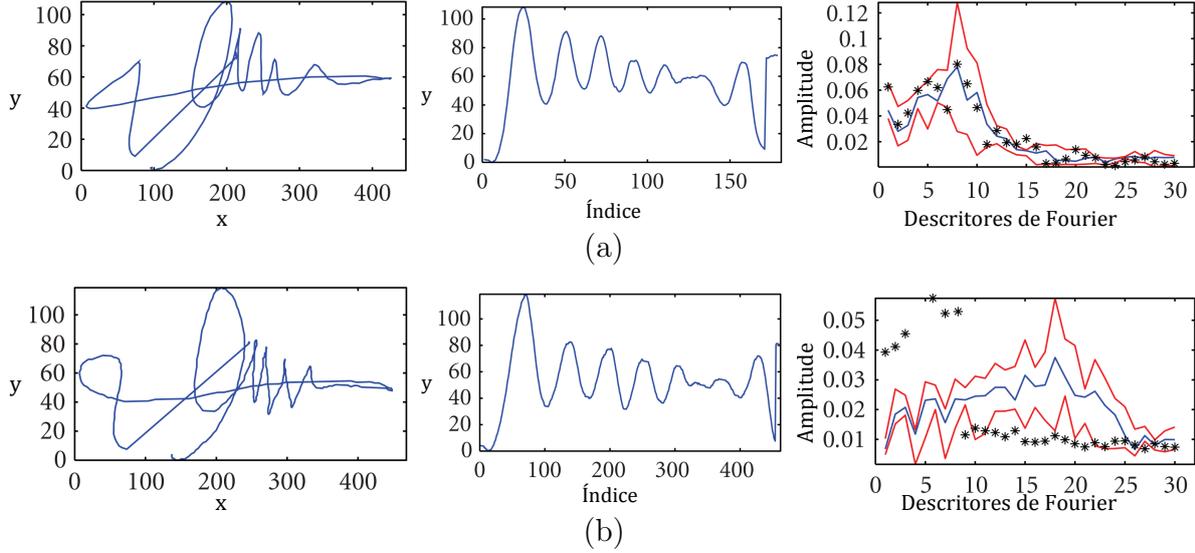


Figura 4.5: Um caso é mostrado para ilustração, usando somente os valores da coordenada  $y$ . Da esquerda para a direita: (a) Assinatura genuína, seu perfil de  $y$  e seus descritores de Fourier; (b) Assinatura falsificada, seu perfil de  $y$  e seus descritores de Fourier. Os descritores de Fourier das assinaturas genuínas e falsas (apresentadas como pontos) são sobrepostos, mostrando os valores mínimos e máximos dos descritores de Fourier das assinaturas de referência, enquanto a linha no meio denota a média das assinaturas de referência. Fonte: [7]

Com isso, a característica de velocidade pode ser calculada com o seguinte algoritmo, cujo fluxo é explicitado na Figura 4.6:

1. Calcula a velocidade média de ambas as componentes,  $x$  e  $y$ :

$$velMedia_x = \frac{1}{N} \sum_{j=1}^N (v_x)_j, \quad (4.30)$$

$$velMedia_y = \frac{1}{N} \sum_{j=1}^N (v_y)_j; \quad (4.31)$$

2. Com base nas velocidades médias, calcule a velocidade absoluta:

$$velAbs = \sqrt{velMedia_x^2 + velMedia_y^2}; \quad (4.32)$$

3. Calcula o vetor diferença (o qual está relacionado com a aceleração), com base nos vetores de velocidade:

$$(difVel_x)_k = (v_x)_{k+1} - (v_x)_k, \quad (4.33)$$

$$(difVel_y)_k = (v_y)_{k+1} - (v_y)_k; \quad (4.34)$$

4. Assim como o vetor velocidade, calcule as médias desses vetores diferença:

$$difVelMedia_x = \frac{1}{N} \sum_{j=1}^N (difVel_x)_j, \quad (4.35)$$

$$difVelMedia_y = \frac{1}{N} \sum_{j=1}^N (difVel_y)_j; \quad (4.36)$$

5. Calcule uma “velocidade absoluta” equivalente para os vetores diferença:

$$difVelAbs = \sqrt{difVelMedia_x^2 + difVelMedia_y^2}. \quad (4.37)$$

6. Obtenha a característica de velocidade:

$$caracVelocidade = a \times velAbs + b \times difVelAbs \quad (4.38)$$

onde  $a, b \in \mathbb{N}^*$  e  $a + b = 1$ .

Os valores adotados para  $a$  e  $b$  serão explicitados no Capítulo 5. Ao fim do processo de extração de características, uma amostra  $S$  é representada por  $S = \{caracVelocidade, H, F\}$ , onde  $caracVelocidade$  representa a métrica derivada da velocidade e aceleração, citada anteriormente,  $H$  representa o conjunto dos  $K$  níveis de histogramas e  $F$  representa os descritores de Fourier para  $x$  e  $y$ .

## 4.5 Composição do Perfil

O próximo passo é a criação do perfil do usuário no sistema. Esse é o caminho do diagrama de fluxo explicitado na Figura 4.1. Nesse contexto, o usuário que deseja se cadastrar no sistema deve fornecer  $M$  amostras de determinada assinatura/palavra para realizar o processo.

É interessante destacar que no caso de palavras, há a possibilidade de escolha de uma palavra específica e mesmo a situação em que o usuário deseja cadastrar mais do que uma palavra, de forma a utilizar determinadas palavras para determinados fins. Nesse caso, o usuário do sistema teria de fornecer  $M$  amostras de cada uma das palavras, visto que há necessidade de um número mínimo para fazer as devidas análises. As próximas seções irão detalhar o processo de criação dos perfis.

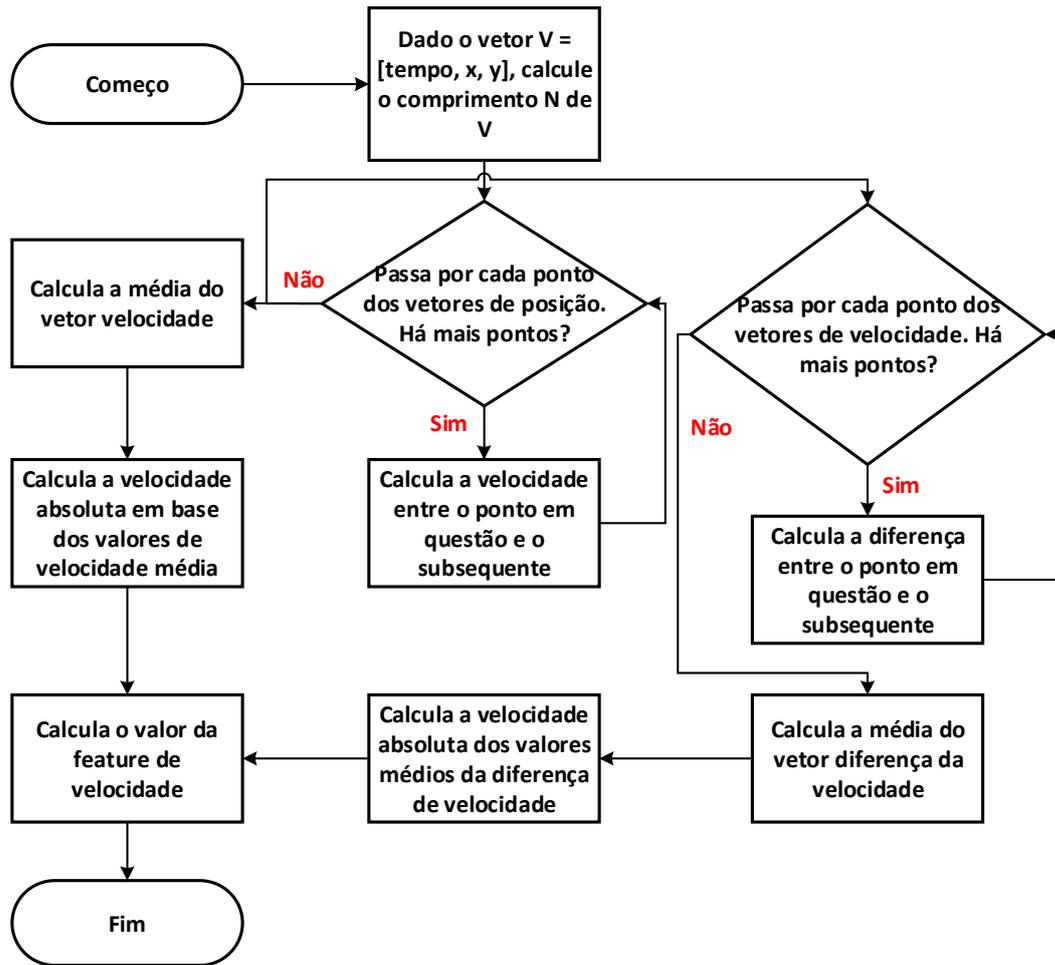


Figura 4.6: Diagrama de fluxo para o cálculo da característica de velocidade.

#### 4.5.1 Análise das Amostras do Perfil

Da etapa de extração de características, uma amostra, que inicialmente é representada pelo vetor de características inicial  $V = [tempo_j, x_j, y_j]$ ,  $1 \leq j \leq N$ , tal que  $N \in \mathbb{N}^*$ . A partir desse vetor, foram calculados os descritores de Fourier, representados por  $F = [(F_x)_j, (F_y)_j]$ ,  $1 \leq j \leq N$  utilizando o algoritmo da FFT. Além disso, foram gerados um conjunto de  $K$  histogramas, representados por  $H = [h_1, h_2, \dots, h_K]$ , onde  $K \in \mathbb{N}^*$ , com os números de bins  $B = [b_1, b_2, \dots, b_K]$  com  $a < b_1 < b_2 < \dots < b_K < b$ , com  $a, b \in \mathbb{N}^*$ .

Considerando o total de  $M$  amostras providas pelo usuário, tem-se as seguintes informações, para  $1 \leq m \leq M$ , correspondente a cada uma das amostras:

- $F = [F_1, F_2, \dots, F_M]$  tal que  $F_m = [(F^m_x), (F^m_y)]$ ;

- $H = [H_1, H_2, \dots, H_M]$  tal que  $H_m = [h^1_m, h^2_m, \dots, h^K_m]$ ;
- $caracVelocidade = [caracVelocidade_1, caracVelocidade_2, \dots, caracVelocidade_m]$ .

O diagrama de fluxo para a análise de perfil pode ser visto na Figura 4.7. Os cálculos realizados nesse procedimento podem ser descritos nos seguintes passos:

1. Calcula o valor médio da característica de velocidade entre as  $M$  amostras:

$$velocidadePerfil = \frac{1}{M} \sum_{m=1}^M caracVelocidade_m \quad (4.39)$$

com  $1 \leq m \leq M$ .

2. Considere cada subconjunto  $H^i = [h^i_1, h^i_2, \dots, h^i_M]$ , com  $1 \leq i \leq K$ , tal que  $K$  é o total de histogramas criados para cada amostra e  $M$  é o total de amostras, representando o grupo de histogramas com  $b^i$  bins. Para cada subgrupo:

- (a) Calcular uma matriz  $matrizD^i$ , de tamanho  $M \times M$ , cujos elementos são dados por:

$$matrizD^i_{m,n} = ||h^i_m, h^i_n|| \quad (4.40)$$

com  $1 \leq m, n \leq M$  e  $||a, b||$  representa alguma métrica/kernel de avaliação do sistema entre os dois histogramas. Cada linha  $m$  dessa matriz representa a comparação do histograma de  $b^i$  bins da  $m$ -ésima amostra com relação a todas as outras amostras, inclusive consigo mesma.

- (b) Calcular o vetor coluna  $minimoVetor^i$ , de comprimento  $M$  dado pela seguinte regra:

$$minimoVetor^i_m = \frac{1}{M-1} \left( \sum_{n=1}^{m-1} matrizD^i_{m,n} + \sum_{n=m+1}^M matrizD^i_{m,n} \right). \quad (4.41)$$

Cada linha do  $minimoVetor^i$  representa a média da respectiva linha em  $matrizD^i$  desconsiderando as autocomparações. Ele provê a informação de quão similar uma amostra é com relação à cada uma das outras.

- (c) Calcular algumas métricas:

- i. A média dos valores mínimos de cada linha de  $matrizD^i$  desconsiderando as autocomparações, representado por:

$$valorMinimo^i = \frac{1}{M} \sum_{j=1}^M (\minimo\{matrizD^i_{j,1}, matrizD^i_{j,2}, \dots, matrizD^i_{j,j-1}, matrizD^i_{j,j+1}, matrizD^i_{j,j+2}, \dots, matrizD^i_{j,M}\}). \quad (4.42)$$

Em outras palavras, essa equação calcula a média das distâncias dos vizinhos mais próximos de cada amostra.

- ii. A média dos valores máximos de cada linha de *matrizD* desconsiderando as autocomparações, representado por:

$$\begin{aligned} \text{valorMaximo}^i = \frac{1}{M} \sum_{j=1}^M (\text{máximo}\{ & \text{matrizD}^i_{j,1}, \text{matrizD}^i_{j,2}, \dots, \\ & \text{matrizD}^i_{j,j-1}, \text{matrizD}^i_{j,j+1}, \text{matrizD}^i_{j,j+2}, \dots, \text{matrizD}^i_{j,M}\}). \end{aligned} \quad (4.43)$$

Dessa forma, calcula-se a média das distâncias dos vizinhos mais distantes de cada amostra.

- iii. O valor mínimo do vetor *minimoVetor*, que, por sua vez, provê informação de quão similar uma amostra é em relação ao conjunto de referência.

$$\begin{aligned} \text{valorARef}^i = \text{mínimo}\{ & \text{minimoVetor}^i_1, \text{minimoVetor}^i_2, \dots, \\ & \text{minimoVetor}^i_M\} \end{aligned} \quad (4.44)$$

Com isso, tem-se o valor de dissimilaridade da amostra que é a amostra mais próxima de todas as outras em média. Além disso, calcula-se *posARef<sup>i</sup>*, que representa qual é essa amostra de valor mínimo.

Após esses passos obtêm-se um vetor de análises para os histogramas de nível *i*, *P<sup>i</sup><sub>H</sub>*, definido como

$$P^i_H = \{\text{valorMinimo}^i, \text{valorMaximo}^i, \text{valorARef}^i, \text{posARef}^i, H^i\} \quad (4.45)$$

onde *H* representa o conjunto de todos os histogramas das *M* amostras.

3. Considere o conjunto de descritores de Fourier  $F = [F_1, F_2, \dots, F_M]$  com  $F_m = [(F^m_x), (F^m_y)]$ ,  $1 \leq m \leq M$ . Um processamento similar ao Item 2 da lista atual (aplicado a histogramas) é feito. A diferença é que neste caso só há um grupo de descritores para cada amostra, então não será necessário realizar um processamento em *K* níveis diferentes.

- (a) Calcula-se uma *matrizD*,  $M \times M$ , para os componentes *x, y* separadamente:

$$\text{matrizD}^x_{m,n} = \|\text{F}^m_x, \text{F}^n_x\| \quad (4.46)$$

$$\text{matrizD}^y_{m,n} = \|\text{F}^m_y, \text{F}^n_y\| \quad (4.47)$$

com  $1 \leq m, n \leq M$  e  $\|a, b\|$  representa alguma métrica para avaliação da diferença entre os dois espectros. A contribuição total de  $x$  e  $y$  é considerada para o cálculo final:

$$\text{matriz}D_{m,n} = \text{matriz}D_{m,n}^x + \text{matriz}D_{m,n}^y. \quad (4.48)$$

O conteúdo da matriz  $\text{matriz}D_{m,n}$  é semelhante à matriz correspondente dos histogramas. De fato, deste ponto em diante, todos os passos são iguais aos adotados com os histogramas.

(b) Calcula *minimoVetor*:

$$\text{minimoVetor}_m = \frac{1}{M-1} \left( \sum_{n=1}^{m-1} \text{matriz}D_{m,n} + \sum_{n=m+1}^M \text{matriz}D_{m,n} \right) \quad (4.49)$$

(c) Calcula *valorMinimo*, *valorMaximo*, *valorARef* e *posARef*:

$$\text{valorMinimo} = \frac{1}{M} \sum_{j=1}^M (\text{mínimo}\{\text{matriz}D_{j,1}, \text{matriz}D_{j,2}, \dots, \text{matriz}D_{j,j-1}, \text{matriz}D_{j,j+1}, \text{matriz}D_{j,j+2}, \dots, \text{matriz}D_{j,M}\}). \quad (4.50)$$

$$\text{valorMaximo} = \frac{1}{M} \sum_{j=1}^M (\text{máximo}\{\text{matriz}D_{j,1}, \text{matriz}D_{j,2}, \dots, \text{matriz}D_{j,j-1}, \text{matriz}D_{j,j+1}, \text{matriz}D_{j,j+2}, \dots, \text{matriz}D_{j,M}\}). \quad (4.51)$$

$$\text{valorARef} = \text{mínimo}\{\text{minimoVetor}_1, \text{minimoVetor}_2, \dots, \text{minimoVetor}_M\}. \quad (4.52)$$

Obtêm-se, com a conclusão do Passo 3, o vetor de análises  $P_F = \{\text{valorMinimo}, \text{valorMaximo}, \text{valorARef}, \text{posARef}, F\}$ .

Ao final de todo o processo de análise de amostras para a composição do perfil, um perfil é representado pelo conjunto

$$P = \{\text{velocidadePerfil}, P_H, P_F\}, \quad (4.53)$$

com  $P_H = \{P_H^1, P_H^2, \dots, P_H^K\}$ .

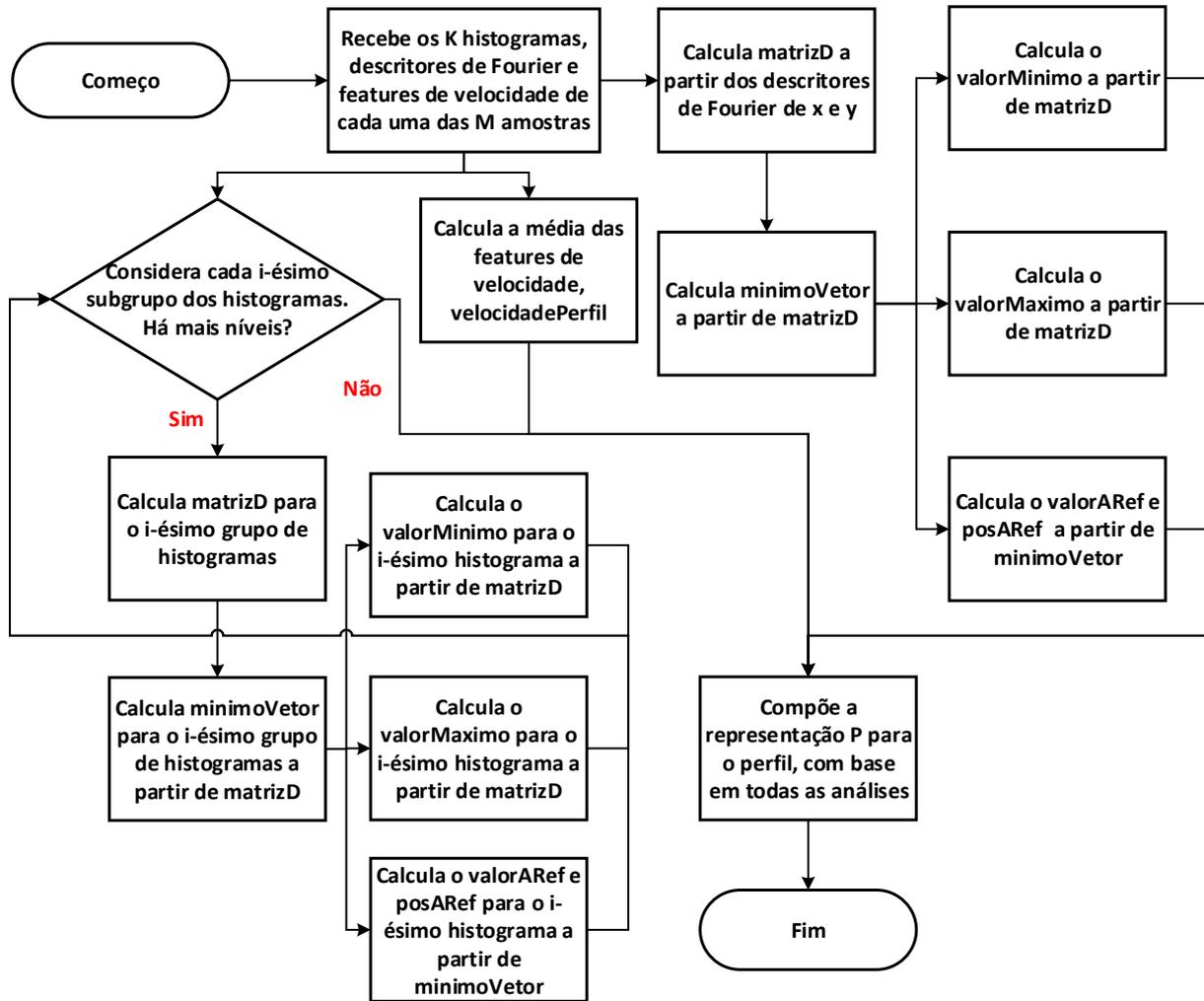


Figura 4.7: Diagrama de fluxo representando os passos de análise de perfil. Ao receber o bloco de características das  $M$  amostras, são expandidos três fluxos em paralelo: um para calcular a característica de velocidade, outro para fazer amostras em cada um dos níveis dos histogramas e um último para realizar as mesmas análises nos descritores de Fourier.

## 4.6 Classificador

Esse estágio é realizado durante a tarefa de verificação e é inspirado em [28]. Para realizar a verificação/autenticação do usuário no sistema, o usuário deve prover uma amostra de assinatura ou de palavra (daquela que seja adequada para a autenticação em questão), além da identidade com a qual está querendo ser verificado.

Nessa situação, o sistema irá buscar o perfil dessa identidade, representado por  $P = \{velocidadePerfil, P_H, P_F\}$ . Assim, dado uma amostra  $S =$ , cujas características já foram extraídas, sendo representada por  $S = \{caracVelocidade_S, H_S, F_S\}$ , que se queria comparar com o perfil  $P$ , faz-se os seguintes passos:

1. Calcula a diferença entre a característica de velocidade da amostra  $S$  com a média correspondente calculada para o perfil  $P$ :

$$difVelocidade_S = |caracVelocidade_S - velocidaPerfil|, \quad (4.54)$$

onde  $|a|$  é o módulo de  $a$  tal que  $a \in \mathbb{R}$ .

2. Considera cada subconjunto de histogramas do perfil  $H^i = [h^i_1, h^i_2, \dots, h^i_M]$ , com  $1 \leq i \leq K$ , tal que  $K$  é o total de histogramas criados para cada amostra e  $M$  é o total de amostras, representando o grupo de histogramas com  $b^i$  bins e o histograma correspondente da amostra  $S$ ,  $h^i_S$ :

- (a) Calcula o vetor coluna  $difAmostra^i$ , de comprimento  $M$ , cujos elementos são dados por:

$$difAmostra^i_j = ||h^i_S, h^i_j||, \quad (4.55)$$

onde  $||a, b||$  representa a métrica/kernel de comparação entre os elementos  $a$  e  $b$ . O vetor  $difAmostra$  representa a diferença da amostra em questão com relação a todas as amostras do perfil  $P$  no nível  $i$ .

- (b) Seleciona três valores específicos, em posições iguais ou diferentes, do vetor  $difAmostra$ , dados por:

$$minimoValor^i_S = minimo\{difAmostra_1, difAmostra_2, \dots, \quad (4.56)$$

$$difAmostra_M\}$$

$$maximoValor^i_S = maximo\{difAmostra_1, difAmostra_2, \dots, \quad (4.57)$$

$$difAmostra_M\}$$

$$valorAREf^i_S = difAmostra_{\{posAREf^i_P\}}, \quad (4.58)$$

onde  $minimoValor^i_S$  representa o vizinho mais próximo da amostra  $S$  com relação ao perfil  $P$ ,  $maximoValor^i_S$ , o vizinho mais distante e  $valorAREf^i_S$  é a comparação da amostra com a amostra considerada a “amostra de perfil” da seção anterior, cujo indicação é dada por  $valorAREf^i_P$

- (c) Normaliza os valores  $minimoValor^i_S, maximoValor^i_S, valorAREf^i_S$ , obtidos pela comparação com a amostra  $S$ , pelos respectivos valores calculados na

análise do perfil  $P$ :

$$valorARef'_{iS} = \frac{valorARef^i_S}{valorARef^i_P} \quad (4.59)$$

$$minimoValor'_{iS} = \frac{minimoValor^i_S}{minimoValor^i_P} \quad (4.60)$$

$$maximoValor'_{iS} = \frac{maximoValor^i_S}{maximoValor^i_P} \quad (4.61)$$

3. Correspondentemente, para o conjunto de descritores de Fourier do perfil  $P$ ,  $F_P = \{F_P^1, F_P^2, \dots, F_P^M\}$ , tal que  $F_P^m = [(F_P^x)^m, (F_P^y)^m]$ , e os descritores correspondentes da amostra  $S$ ,  $F_S = [F_S^x, F_S^y]$  adota-se os seguintes passos:

(a) Calcula  $difAmostra$ , de comprimento  $M$ , dado por:

$$difAmostra_j^x = ||F_S^x, (F_P^x)_j|| \quad (4.62)$$

$$difAmostra_j^y = ||F_S^y, (F_P^y)_j|| \quad (4.63)$$

$$difAmostra_j = difAmostra_j^x + difAmostra_j^y \quad (4.64)$$

onde  $||a, b||$  representa uma métrica de comparação de  $a$  e  $b$ . O vetor  $difAmostra$  representa a diferença da amostra  $S$  com relação as amostras do perfil  $P$ , considerando as análises espectrais realizadas.

- (b) Seleciona  $valorMinimos_S$ ,  $valorMaximos_S$  e  $valorARef_S$  de maneira análoga à realizada nos histogramas, evidenciado nas Equações 4.56, 4.57 e 4.58.
- (c) Normaliza os valores obtidos no item anterior pelos respectivos no perfil, conforme evidenciado pelas equações 4.60, 4.61 e 4.59, obtendo  $valorMinimo'_S$ ,  $valorMaximo'_S$  e  $valorARef'_S$ .

A intenção dessa estratégia, selecionar e realizar comparações com o vizinho mais próximo, o vizinho mais distante e com a amostra central do perfil (a quem tem a menor distância média em relação às outras amostras), é caracterizar a variabilidade dentro das amostras do perfil. O papel da normalização pelos valores do perfil, que caracterizam a sua variabilidade, é para permitir a utilização de um limiar de decisão único para o sistema na etapa de decisão, que será descrita mais a frente.

### 4.6.1 Combinação das características

Os valores calculados anteriormente são:

1. *diffVelocidade*, referente à diferença da característica de velocidade da amostra com o perfil;
2. *minimoValor<sup>i</sup><sub>S</sub>*, *maximoValor<sup>i</sup><sub>S</sub>*, *valorARef<sup>i</sup><sub>S</sub>*, representando a maior distância de *S* em relação ao perfil *P*, menor distância e distância em relação à amostra central de *P*, respectivamente, normalizadas, para o *i*-ésimo nível de histogramas;
3. De forma correspondente ao item anterior, *minimoValor'<sub>S</sub>*, *maximoValor'<sub>S</sub>*, *valorARef'<sub>S</sub>*, calculados a partir da comparação dos descritores de Fourier da amostra *S* em relação ao perfil *P*.

O próximo passo envolve combinar todos esses valores, de forma a obter uma única métrica que possa ser julgada na etapa de decisão, referenciada na seção subsequente. A abordagem que inspirou o classificador adotado [28], aplicava a técnica de Análise de Componentes Principais (PCA, do inglês, *Principal Component Analysis*). Assim, o vetor 3D obtido, correspondente às três distâncias calculadas para os descritores de Fourier e histogramas poderia ser reduzido para um único componente.

Para os histogramas, uma distância final pode ser calculada por:

$$distanciaHist^i_S = k^i_1 \times minimoValor^i_S + k^i_2 \times maximoValor^i_S + k^i_3 \times valorARef^i_S, \quad (4.65)$$

onde  $k^i_1, k^i_2, k^i_3 \in \mathbb{R}$  e *i* representa o *i*-ésimo nível de histogramas. Esse conjunto de valores é escolhido de forma a maximizar a variância do conjunto original de 3 componentes. Analogamente, pode-se calcular uma métrica única para os 3 componentes de distância obtidos com as comparações dos descritores de Fourier:

$$distanciaFourier_S = k_1 \times minimoValor'_S + k_2 \times maximoValor'_S + k_3 \times valorARef'_S. \quad (4.66)$$

A ideia seria usar alguns perfis, amostras verdadeiras e falsificações e treinar um procedimento baseado em PCA que estime esses valores. Porém, após várias simulações, foi observado que não há grande variação nos valores dos coeficientes e que eles podem ser fixados sem perdas consideráveis de performance.

A esse ponto das análises, ainda existem 3 métricas como resultado da comparação da amostra *S* com o perfil *P*: *diffVelocidade<sub>S</sub>*, *distanciaFourier<sub>S</sub>* e *distanciaHist<sup>i</sup>* representando a diferença da característica de velocidade, a distância da comparação dos

descritores de Fourier e a distância da comparação com os histogramas de nível  $i$ . Ainda é realizada uma combinação desses três valores para obter o valor final.

A estratégia utilizada é uma combinação linear desses valores. Aqui, entretanto, é necessário citar que são analisadas duas formas de adotar a técnica, citadas abaixo:

1. "Abordagem completa": nesse cenário, as informações obtidas de todos os níveis de histogramas,  $distanciaHist^i$ , com  $1 \leq i \leq K$  são combinadas em uma única métrica  $distanciaHist$  por média:

$$distanciaHist = \frac{1}{K} \sum_{i=1}^K distanciaHist^i. \quad (4.67)$$

2. "Abordagem cascata": nessa situação, cada valor  $distanciaHist^i$ ,  $1 \leq i \leq K$ , é usado numa análise de vários níveis, como um espécie de cascata.

Para o caso do Item 1 da lista anterior, a distância final é dada por:

$$distanciaFinal_S = k_4 \times distanciaHist_S + k_5 \times distanciaFourier_S + k_6 \times difVelocidade_S. \quad (4.68)$$

Para a situação representada no Item 2 da lista subjacente, a distância final é definida para cada um dos  $K$  níveis:

$$distanciaFinal^i_S = k_4^i \times distanciaHist^i_S + k_5^i \times distanciaFourier_S + k_6^i \times difVelocidade_S. \quad (4.69)$$

Detalhes sobre a estimação dos parâmetros serão explicitados no Capítulo 5. A próxima seção abordará maiores informações a respeito do uso dessas duas abordagens, a completa e a cascata, para a tarefa de decisão.

## 4.7 Decisão

O estágio de decisão é o último passo que o sistema adota na tarefa de verificação/autenticação. Consiste basicamente de uma comparação entre a métrica obtida no estágio de classificação para a amostra  $S$  e um limiar especificado.

Os passos de normalização adotados no estágio de classificação tem como fim colocar essa métrica em um valor estável e padrão para todos os usuários, de forma que seja possível estimar limiar global que seja ótimo. Conforme citado, há duas abordagens diferentes que são avaliadas nesse trabalho: a abordagem "completa" e a "cascata".

Os seguintes passos são adotados:

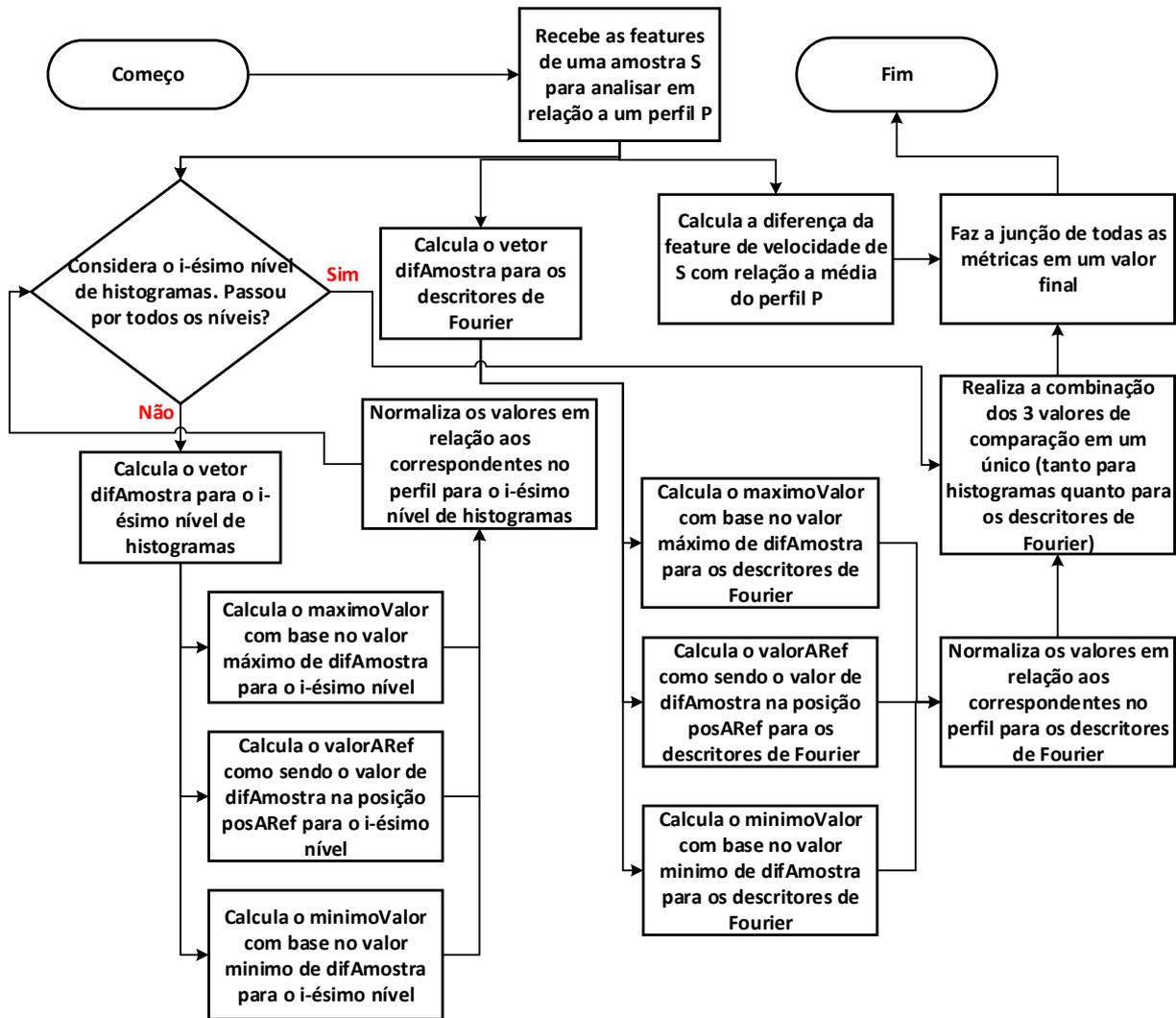


Figura 4.8: Diagrama de fluxo para os passos adotados no classificador.

1. Abordagem completa: estipula-se um limiar global  $L \in \mathbb{R}$  de forma que todas os valores de  $distanciaFinal$  para uma amostra  $S$  são comparados. O seguinte teste é feito:
  - (a) Se  $distanciaFinal \leq L \rightarrow$  a amostra  $S$  é considerada verdadeira;
  - (b) Se  $distanciaFinal > L \rightarrow$  a amostra  $S$  é considerada uma falsificação.
2. Abordagem cascata: o mesmo limiar global  $L \in \mathbb{R}$  pode ser utilizado nessa estratégia. A partir dele, pode-se gerar vários limiares correspondentes ao  $i$ -ésimo nível de histogramas,  $L^i$ , com a seguinte regra:

$$L^i = a^i \times T \quad (4.70)$$

onde tem-se a seguinte relação

$$0 < a^1 < a^2 < \dots < a^i < \dots < a^K = 1. \quad (4.71)$$

O teste, dado um nível  $i$  para os histogramas, é adotado da seguinte forma:

- (a) Se  $distanciaFinal \leq L^i \rightarrow$  a amostra  $S$  é considerada verdadeira;
- (b) Se  $i < K$  e  $distanciaFinal > L^i \rightarrow$  o teste é repetido para o nível  $i = i + 1$ .
- (c) Se  $i = K$  e  $distanciaFinal > L^i \rightarrow$  a amostra é considerada uma falsificação.

Quando considera-se o método completo une-se informações de vários níveis de análises para realizar a decisão. A ideia é que com diferentes números de *bins* pode-se maximizar a variação inter-usuário.

Ja ao considerar o método cascata, a ideia é começar com um menor número de *bins*. Para esse primeiro passo, leva-se em consideração um valor inferior ao limiar global de referência  $T$ . Isso porque com um menor número de *bins* há a possibilidade de haver discriminação insuficiente para a amostras e, assim, a intenção é diminuir as chances de aceitar uma amostra falsamente. Na medida em que os números de *bins* sobem, há um aumento no limiar de aceitação. A ideia, basicamente, é ter um método com um desempenho adequado e ser computacionalmente mais rápido que o anterior.

Uma análise dos resultados desses métodos será descrita no Capítulo 5. Também será feita uma avaliação com o uso de histogramas simples para fins de comparação.

## 4.8 Análise da Complexidade Assintótica para o Sistema Proposto

Esta seção abordará a análise da complexidade assintótica do sistema proposto. Como o cenário a ser avaliado é o de dispositivos móveis, é necessário considerar que a técnica não pode ser demasiadamente complexa. O sistema proposto é baseado na extração de características globais. É realizada uma análise no domínio da frequência para compor o vetor de descritores de Fourier e uma análise espacial para criar os histogramas. Uma característica de velocidade também é computada.

Considerando um conjunto de  $N$  pontos amostrados para capturar a palavra/assinatura, a complexidade assintótica para calcular a característica de velocidade é  $\mathcal{O}(N)$ . No caso dos histogramas, a complexidade também é  $\mathcal{O}(N)$ . Isto porque a composição de um único histograma pode ser feita em tempo linear e, mesmo que sejam computados vários histogramas, o número deles é constante. Por sua vez, a criação dos descritores de Fourier envolve o uso do algoritmo da FFT, que tem complexidade assintótica  $\mathcal{O}(N \times \log N)$ . O

algoritmo também é aplicado um número constante de vezes. Todos os cálculos auxiliares ou são feitos em tempo constante ou em tempo linear. Assim, a complexidade assintótica do sistema é  $\mathcal{O}(N \times \log N)$ .

O próximo capítulo evidenciará os experimentos realizados para avaliar a performance do sistema proposto neste trabalho.

# Capítulo 5

## Resultados Experimentais

Este capítulo abordará os principais experimentos realizados além dos resultados obtidos. Serão relatados os detalhes da criação da base de dados LISA-01 e informações a respeito da base de dados MCYT-100. Após, serão descritos os experimentos com histogramas, seguidos pelas análises de performance do sistema completo. Ao final, será realizada uma comparação entre o uso de caneta e dedo para a composição de perfis, abordando a criação de uma pequena base de dados para avaliação deste cenário.

### 5.1 Bases de dados

Esta seção descreverá detalhes da composição da base de dados LISA-01, que aborda amostras de palavras cursivas isoladas escritas com a ponta dos dedos. Além disso, também serão descritas características da base de dados MCYT-100, relativa a amostras de assinaturas.

#### 5.1.1 A base de dados LISA-01

A base de dados LISA-01 consiste de um conjunto de amostras de palavras cursivas isoladas capturadas sem o uso de nenhum dispositivo auxiliar, simplesmente com o dedo. A opção por montar a base de dados dessa maneira foi com a intenção de ilustrar um cenário onde não é necessário o uso de nada além do dispositivo móvel.

As pessoas são acostumadas a escrever com o uso de canetas. Naturalmente, o uso de uma caneta para a captura dos dados acarretaria em uma maior afinidade das pessoas para a escrita. Esse fato poderia levar a um maior grau de discriminação das palavras. Todavia, há um interesse maior em simular um cenário minimalista.

Essa base de dados está publicamente disponível no endereço <http://www.cic.unb.br/~fbvidal/htdb>. Para ter acesso, basta seguir os procedimentos explicitados na página.

### Interface para recolhimento dos dados

A dispositivo utilizado para captura os dados da base LISA-01 foi um *tablet* da fabricante *Samsung*, modelo *Galaxy Tab 7.0 Plus* (GT-P6200L), com Sistema Operacional Android versão 4.2, que conta com uma tela de 7 polegadas e frequência de amostragem de 60Hz. O software utilizado foi desenvolvido por conta própria, com o auxílio da ferramenta de desenvolvimento *Eclipse Juno*, cuja interface é apresentada na Figura 5.1



Figura 5.1: Interface desenvolvida para o recolhimento dos dados da base LISA-01.

As únicas informações atribuídas a cada usuário foram um número único, começando por 1 e seguindo de forma crescente nos números naturais, e uma referência da palavra que estava sendo escrita.

O dispositivo periodicamente extrai informações a respeito do estado do toque. O escritor deveria esboçar a palavra na área em branco. O conjunto de todas essas extrações formam uma sequência de dados no tempo. À medida que o usuário desliza o dedo sobre a superfície uma representação gráfica é apresentada. Um ponto  $(x_j, y_j)$  amostrado no tempo  $t_j$  é ligado por uma reta ao ponto  $(x_{j+1}, y_{j+1})$  amostrado no tempo  $t_{j+1}$ , dado que este é o tempo atual. Repetindo esse processo  $\forall t_j$  tem-se a representação gráfica da palavra.

## Protocolo de aquisição

O protocolo de aquisição para a criação da base de dados seguiu os seguintes parâmetros:

- 50 escritores aleatórios foram selecionados, com idades variando entre 20 e 50 anos e proporções semelhantes de pessoas do sexo masculino e feminino;
- Cada escritor teve de escrever 4 palavras específicas, sendo 10 repetições de cada uma;
- As palavras deviam ser escritas de forma natural, sendo predominantemente cursivas.

Dessa forma, a base de dados conta com 2000 amostras. As palavras escolhidas para compor a base de dados são “*december*”, “*intelligence*”, “*love*” e “*pattern*”. Não houveram critérios específicos para a escolha das palavras. A única intenção foi escolher pelo menos uma palavra relativamente grande (nesse caso a palavra “*intelligence*” encaixaria nesse cenário) e uma palavra pequena (a palavra “*love*” é bastante curta).

O sentido de predominantemente cursivas, do último item, refere-se ao fato de que não era obrigatório que o usuário escrevesse de forma totalmente contínua, mas que tomasse cuidado para não se tornar uma escrita em letra de forma. Portanto, era uma restrição de caráter bem amplo, visto que cumprindo isso, o usuário tinha total liberdade para escrever. Essa liberdade acaba se tornando um problema, pois aumenta a possibilidade de oscilações na forma com a qual o usuário escreve. A Figura 5.2 mostra um exemplo de 3 amostras de cada palavra para determinado usuário.

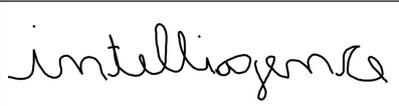
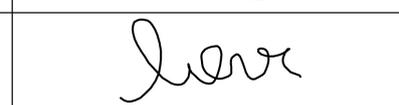
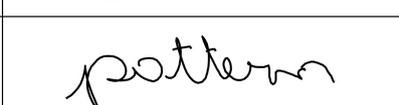
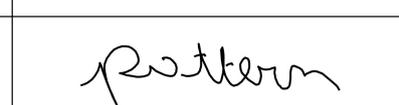
|   |   |  |
|---|---|--|
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

Figura 5.2: Conjunto de 3 amostras de cada uma das palavras da LISA-01: “*december*”, “*intelligence*”, “*love*” e “*pattern*”. Cada grupo de palavras pertence a um usuário distinto.

Observando a Figura 5.2 é possível observar que por mais que o usuário escreva de maneira “igual”, haverá variações naturais ao escrever. Esse é um dos grandes desafios ao lidar com análise de palavras.

## Os dados recolhidos

Dois tipos de dados foram armazenados na base LISA-01: as características dinâmicas do processo de escrita e a imagem da palavra. As características dinâmicas são úteis para simular sistemas em cenários de reconhecimento *online*, enquanto as imagens podem ser utilizadas em sistemas de reconhecimento *offline*. Os dois tipos de informações podem ser usadas em propostas híbridas - *online* e *offline*.

Os dados dinâmicos, armazenados nos arquivos, podem ser representados pelo seguinte vetor:

$$[V]^N = [estampaTempo_j, nrStroke_j, tamanho_j, x_j, y_j]^N \quad (5.1)$$

com  $1 \leq j \leq N$ ,  $N, j \in \mathbb{N}$ , tal que  $N$  representa o comprimento, em quantidade de pontos, da amostra em questão.

O primeiro campo, *estampaTempo* representa o valor, em milissegundos, em que o  $j$ -ésimo ponto foi amostrado, tendo como tempo 0 o momento em que o sistema operacional foi iniciado. O campo *nrStroke* representa um contador da quantidade de toques até o momento. Por exemplo, inicialmente o valor é 1. A partir do momento em que o usuário tirou o dedo em algum momento e tocou novamente, ele passa a ser 2, e assim sucessivamente.

O *tamanho* representa a estimativa feita pelo sistema operacional, em termos de pixels, da área de toque. Essa informação está vinculada à pressão, visto que quanto maior for a pressão aplicada no dedo, maior será a área de toque estimada. A dupla  $(x, y)$  representa as coordenadas espaciais tomando como  $(0, 0)$  o canto superior esquerdo da tela do dispositivo.

Embora todas essas informações estejam disponíveis na base de dados, os únicos dados que efetivamente foram utilizados foram a *estampaTempo* e  $(x, y)$ . Isto porque a variável *nrStroke* é simplesmente uma variável que representa qual o número de determinado traço da amostra, sendo útil para isolar todos os traços da palavra, em técnicas que fazem uso desta características. Já a variável *tamanho* está relacionada com a pressão, características que não foi considerada aqui por conta de sua variabilidade em relação ao meio de escrita.

### 5.1.2 A base de dados MCYT-100

Para o segundo cenário a ser avaliado, o de assinaturas escritas a mão, foi utilizada a base de dados MCYT-100, apresentada em [67]. Trata-se de uma das bases de dados mais utilizadas em trabalhos envolvendo reconhecimento biométrico de escritores baseado em assinaturas [66, 28, 6, 29, 17, 18]. O dispositivo utilizado para captura foi o *tablet* gráfico WACOM, modelo INTUOS A6 USB. A resolução do *tablet* é 2540 linhas por polegada (100 linhas / mm) e a precisão é de  $\pm 0.25$ mm. A frequência de amostragem dos sinais adquiridos é de 100Hz.

Cada usuário produziu 25 assinaturas genuínas e 25 assinaturas falsificadas foram produzidas para cada usuário. Essas assinaturas falsificadas foram produzidas pelos 5 usuários subsequentes ao usuário em questão. Os falsificadores poderiam observar as imagens estáticas da assinatura para tentar imitá-las. Antes de tentar forjar, o falsificador tinha 10 tentativas de copiar para se familiarizar com a assinatura. Dessa forma, falsificações proficientes foram obtidas. A Figura 5.3 mostra exemplos de amostras da base de dados MCYT-100 [67].

As informações disponíveis na base de dados são:

- Informações espaciais,  $(x_t, y_t)$ , no intervalo  $[0 - 12700]$  e  $[0 - 9700]$ , respectivamente;
- A pressão  $p_t$  aplicada pela caneta, no intervalo  $[0 - 1024]$ ;
- O ângulo azimute,  $\gamma_t$ , da caneta em relação ao *tablet*, no intervalo  $[0 - 3600]$ , correspondendo a  $0^\circ - 360^\circ$ ;
- O ângulo de altitude,  $\phi_t$ , da caneta em relação ao *tablet*, no intervalo  $[300 - 900]$ , correspondendo a  $30^\circ - 90^\circ$ .

Assim como na base LISA-01, as únicas informações utilizadas foram no trabalho foram as coordenadas especiais e o tempo de amostragem.

## 5.2 Parâmetros adotados nos experimentos

Esta seção irá retratar alguns detalhes dos parâmetros utilizados para realizar os testes.

### 5.2.1 Composição do perfil e simulações

O primeiro ponto a descrever é o número de amostras utilizado para compor o perfil,  $M$ . O valor  $M$  foi definido como 5. Dessa forma, esse é o número de amostras de referência, sendo um valor muito utilizado na literatura [7, 33, 35]. Todavia, serão evidenciadas

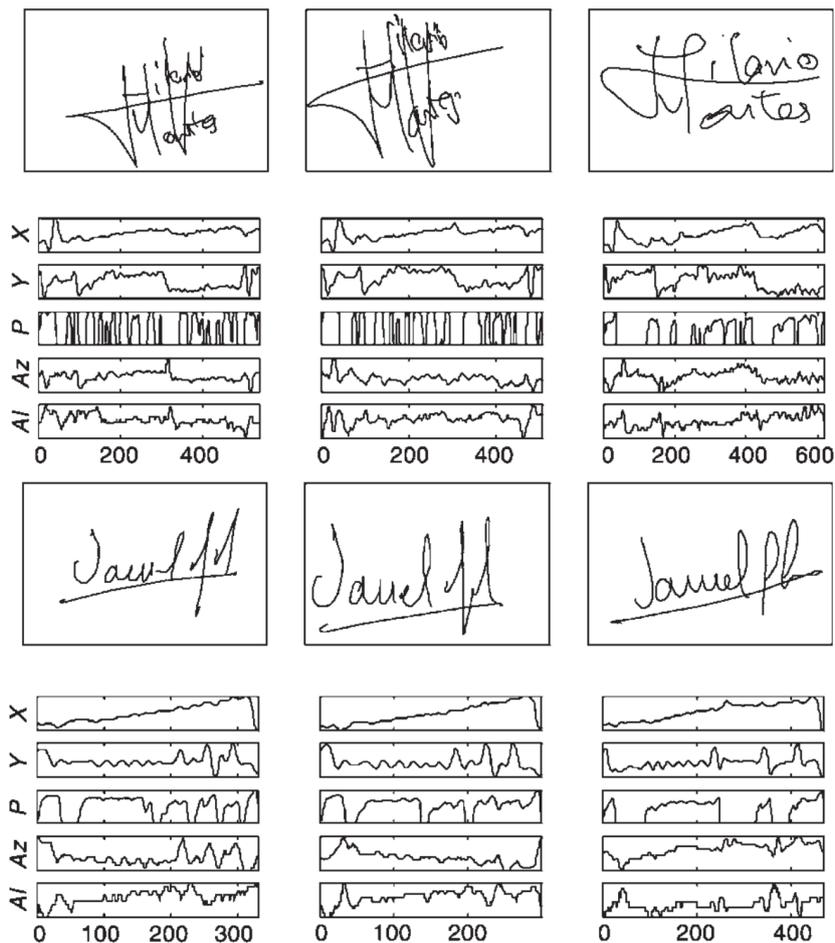


Figura 5.3: Exemplos de amostras de 2 usuários na base de dados MCYT-100. Para cada usuário, as duas amostras mais à esquerda são assinaturas verdadeiras, enquanto que a amostra mais à direita é uma falsificação proficiente. Em cada caso, os *plots* gráficos apresentados correspondem às informações dinâmicas armazenadas na base de dados. Fonte: [7].

algumas análises do impacto em variar essa quantidade de amostras para a eficácia do sistema.

As simulações foram repetidas 5 vezes em cada base de dados, sorteando as amostras usadas para compor os perfis e deixando as remanescentes para teste. Essa é uma prática comum a fim de avaliar melhor o sistema, obtendo uma performance média do mesmo [74, 7].

## 5.2.2 Utilização das amostras das bases de dados

Na base de dados MCYT-100, 25 falsificações qualificadas estão disponíveis para cada um dos usuários, que foram criadas a partir da observação dos outros usuários da base

com uma amostra verdadeira. Dessa maneira, nós avaliamos o sistema considerando que o falsificador tem acesso à assinatura de referência para tentar burlar o sistema. Esse é um cenário onde a EER é maior do que em um cenário onde as falsificações são compostas a partir de amostras de outros usuários ou de tentativas de falsificações onde o fraudador não tinha acesso à assinatura original.

Quando considera-se a base de dados LISA-01, não há falsificações para cada um dos usuários. Todavia, por se tratar de palavras cursivas, é interessante notar que elas naturalmente tem restrições de legibilidade e forma. Naturalmente, versões diferentes de uma mesma palavra terão formas parecidas. Dessa maneira, também sortearam-se as amostras de outros usuários como falsificações para um usuário em questão.

Há ainda uma outra vertente ao lidar com palavras. Pode-se sortear as falsificações como amostras da mesma palavra, ou amostras de palavras avulsas. No primeiro caso, espera-se que as taxas de erro sejam maiores visto que as palavras terão mais semelhança. Esses são dois cenários válidos porque há a flexibilidade de um usuário escolher palavras específicas para realizar determinadas ações.

Dessa forma, se o fraudador tem conhecimento de qual palavra deve ser usada em determinado momento, a situação será abordada pelo primeiro cenário. Todavia, se o fraudador não tiver esse conhecimento, provavelmente tentará escrever uma amostra aleatória, que pode ou não ser correspondente à mesma palavra. A Figura 5.4 mostra um grupo de palavras de um usuário, seguido por um grupo de amostras da mesma palavra de usuários diferentes e mais um grupo de palavras diferentes de usuários diferentes. É possível ver a semelhança entre as palavras iguais, mostrando que mesmo sortear as falsificações, no caso de palavras cursivas, é possível ter um cenário interessante.

### 5.2.3 As características

Conforme apresentado na Seção 4.4.1, é considerado um conjunto de histogramas  $H = \{h^1, h^2, \dots, h^K\}$  e um conjunto de *bins*  $B = \{b^1, b^2, \dots, b^K\}$ , tal que  $b^1 < b^2 < \dots < b^K$ . A estipulação desse conjunto  $B$ , que representa o grupo de histogramas foi realizada de maneira empírica. A performance do sistema foi observada em simulações com várias tentativas para o conjunto  $B$ , até que foi encontrado um conjunto estável de *bins*. Dessa forma, foi estimado que  $B = \{6, 8, 10, 12, 14, 16\}$  é um valor ideal para o número de *bins*, tanto nos testes na MCYT-100, na LISA-01 e na base de comparação de entradas por dedo e por caneta, para o sistema completo. É interessante notar que este grupo, quando são considerados somente os histogramas como características, não retorna os melhores valores. Todavia, as fusões com as outras características levam a diferenças com relação ao uso de histogramas sozinhos e, no caso do sistema completo, o grupo supracitado retorna os melhores valores.

(a)

(b)

(c)

Figura 5.4: Exemplos de amostras de um usuário e possíveis tentativas de falsificação: (a) três amostras da palavra “deceember” de um usuário, (b) conjunto de amostras da palavra “deceember” de usuários diferentes (c) conjunto de amostras de palavras e usuários diferentes.

A Equação 4.38 apresenta o último passo para obter a característica de velocidade,  $caracVelocidade = a \times velAbs + b \times difVelAbs$ , onde os parâmetros  $a, b$  são utilizados para realizar a fusão das informações diferentes. Após a realização de vários testes, foi visto que  $a = 0,6$  e  $b = 0,4$  são os valores ideais para realizar a junção das duas informações.

No que concerne ao uso dos descritores de Fourier, há a questão de quantos descritores utilizar de cada espectro. Isso porque a maior parte da informação da amostra, neste caso, está contida nos descritores de menor frequência. A utilização dos descritores com maior frequência contém bastante ruído, o que pode prejudicar o reconhecimento. Depois de realizar algumas simulações, foi estipulado que os 30 primeiros descritores de cada um dos espectros seriam utilizados. Embora um menor número de descritores, por exemplo 15 ou 20, também tenham EER adequada, o uso de 30 é um bom compromisso entre a performance e o custo computacional. Usar um maior número de descritores não apresentou melhoras.

## 5.2.4 Parâmetros do classificador

Na Seção 4.6, as Equações 4.65 e 4.66 evidenciam como o classificador une as informações de comparação com a amostra mais próxima, amostras mais distante e amostra

de perfil para compor a métrica de cada uma das distâncias, de Histogramas e de Fourier, respectivamente. O valor dessas constantes, tanto para os histogramas quanto para Fourier, representam o mesmo valor.

A estimação destas constantes foi feita realizando a simulação de 5 usuários da base, com 5 amostras verdadeiras e 5 falsificações, para ambas as bases, MCYT-100 e LISA-01. Aplicou-se a técnica de Análise de componentes principais (PCA, do inglês, *Principal Component Analysis*) nos valores de comparação em questão e estimou-se o conjunto de bases que maximiza a variância máxima.

Os experimentos foram repetidos 5 vezes. Os valores retornados pela PCA não tiveram grande variação. Dessa forma, é possível estipular valores fixos sem que haja perdas consideráveis na performance do sistema. Os valores das constantes foram adotados como sendo a média dos valores retornados por PCA. Assim,  $k_1 = 0.3870$ ,  $k_2 = 0.3705$ ,  $k_3 = 0.2425$  para todos os experimentos realizados nesse capítulo. É interessante relatar que diferente de algumas propostas [25, 14, 5, 33, 35, 34, 28], o classificador em questão não precisa de treinamento.

Já o cálculo do valor final para a distância, explicitado na Equação 4.68 envolve a consideração da escala dos valores retornados por cada uma das características. A união de múltiplos valores de características diferentes é objeto de estudo de alguns trabalhos [75, 32, 13]. Aqui, foi utilizado uma junção baseada na combinação linear dos valores de dissimilaridade retornados por cada característica. Depois de vários testes empíricos, foi verificado que as constantes podem ser definidas como  $k_4 = 0.29$ ,  $k_5 = 0.57$ ,  $k_6 = 0.14$ .

### 5.2.5 Limiares e Avaliação de desempenho

A Seção 4.7 apontou os detalhes referentes à etapa de decisão do sistema. Para o uso de um único histograma, por padrão utiliza-se um limiar global  $L \in \mathbb{R}$ . O mesmo se aplica à abordagem completa, onde os resultados das análises dos vários níveis dos histogramas são unidos numa única medida de dissimilaridade. Todavia, para a abordagem cascata, foi apontado que, a partir do limiar global  $L$ , poder-se-ia compor um conjunto de limiares  $L^i = a^i \times L$ , tal que  $0 < a^1 < a^2 < \dots < a^i < \dots < a^K = 1$ .

Várias simulações foram realizadas, até que foi obtida uma regra para composição dos limiares. Foi estipulado que o conjunto de valores de limiar corresponderiam aos pontos da reta que passa pelos pontos 0.8 e 1. Dessa forma, o limiar  $a^1 = 0.8$ , e o valor cresce linearmente, dada a quantidade de histogramas do conjunto  $H$ . Esta estratégia foi a que trouxe os melhores resultados para a EER do sistema nesta abordagem. Isto porque foi o ponto de equilíbrio entre evitar que histogramas com menor número de *bins* levem à erros de reconhecimento no sistema e, ao mesmo tempo, permitir que amostras que já sejam suficientemente parecidas desde o primeiro nível de histogramas possam ser aceitas.

Os resultados foram medidos usando a Taxa de Erro Igual (EER). A EER é o ponto de interseção entre a Taxa de Aceitação Falsa (FAR) e a Taxa de Rejeição Falsa (FRR), representando o equilíbrio entre essas duas importantes métricas de avaliação de desempenho. A EER é uma medida amplamente utilizada para mensurar a performance de sistemas biométricos [7, 33, 35, 28, 34, 74, 5, 14].

## 5.3 Experimento 1 - Análise dos Histogramas

Um dos objetivos deste trabalho é avaliar o uso de histogramas no reconhecimento automático de escritores *online*, visto que o uso dessas características não recebeu atenção da comunidade, embora seja bem utilizada em outras áreas. Ao adotar histogramas, há um problema de obter uma representação estável para a amostra. Isso porque, um baixo número de *bins* pode não ser suficientemente discriminativo e um alto número de *bins* pode trazer bastante instabilidade para a representação das amostras. Ambas as situações são um problema, e podem causar deterioração na performance de um sistema.

Sob esse ponto de vista, a questão que surge é que o número de *bins* para qualquer tipo de histograma, que seja suficientemente discriminativo e dê estabilidade para a representação da amostra é variável. Ao utilizar histogramas, é comum fixar o número de *bins* para o sistema, conforme apresentado em [6], que apresenta um grande número de histogramas unidimensionais e bidimensionais, mas cada um com um número fixo. Ao fixar esse número de *bins*, considerando a questão da representação ótima, certamente ele não se mostrará adequado para todos os usuários.

### 5.3.1 Número ótimo de *bins* dos histogramas

Para elucidar a questão da variabilidade no número de *bins*, foi realizado um cenário hipotético. Fez-se um extravasamento de várias possibilidades de *bins* para cada usuário e supôs-se que o limiar era ótimo, ou seja, a EER tinha sido atingida para aquele usuário. O número de *bins* que melhor representou o usuário em questão foi escolhido. Ou seja, esse seria o cenário em que a extração e seleção de características, a classificação e a decisão foram perfeitamente realizadas, alcançando a taxa de erro mínima possível.

As simulações foram repetidas somente 1 vez. O procedimento de sortear as amostras para composição do perfil e deixar as outras como teste foi seguido. O intervalo de limiares considerados foi [0,1;3] com passo de discretização de 0,01. Foram testadas as duas métricas para as comparações do classificador, a métrica Euclidiana e a distância de Manhattan. A Figura 5.5 mostra o resultado das contagens de números ótimos de *bins*, considerando todos os usuários, dado um intervalo de *bins* que foram considerados para ambas as métricas de comparação utilizadas.

É possível notar que há uma distribuição dos *bins* que retornam a melhor EER para cada usuário por todo o intervalo, com ambas as métricas. Pode-se alegar que este é um fator dinâmico, e que muitos aspectos da execução podem interferir na representação que é ótima para o usuário, como por exemplo as amostras para composição do perfil e as amostras de teste, o classificador adotado e o método de decisão. Seja qual for o motivo apontado, este é um fato a ser considerado.

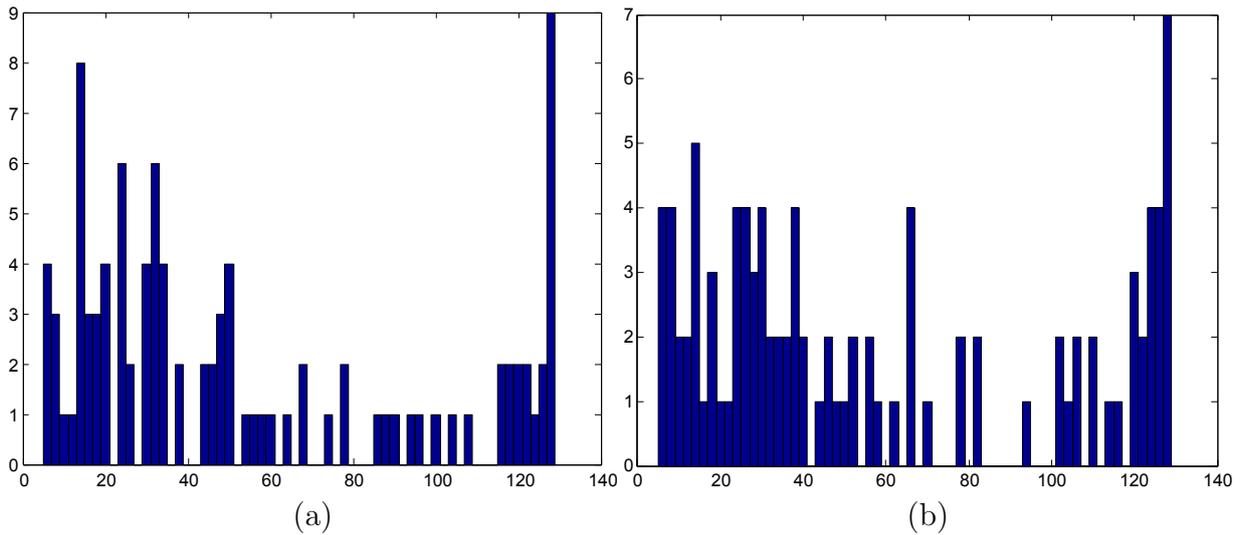


Figura 5.5: A distribuição dos números de *bins* que melhor discriminam cada usuário do sistema, na base de dados MCYT-100: (a) análise feita com a distância euclidiana, (b) análise feita com a distância de Manhattan

### 5.3.2 Avaliação do sistema com as técnicas de histogramas

Posteriormente, foi feita a análise do sistema somente com as características de histogramas, escolhendo um número de *bins* fixo para fazer a simulação. As EER's obtidas são mostradas na Figura 5.6.

É possível observar que a melhor EER, de 14,8%, é obtida com 22 *bins* para a distância euclidiana. Já para a distância de Manhattan, a representação que traz as melhores taxas de EER para o sistema é de 16 *bins*, com uma taxa de 14,2%. É importante citar que ao atribuir um valor arbitrário para o número de *bins* em um sistema real, há razoável possibilidade desse valor não ser o ideal. Um outro ponto é a respeito da instabilidade da taxa de erro ao escolher um número arbitrário de *bins*. Esse é mais um exemplo de que aumentar ou diminuir o número de *bins* pode não ser algo interessante, remontando a questão de que há uma representação ótima para cada amostra.

A Tabela 5.1 evidencia os resultados encontrados com as simulações, para as duas abordagens, em relação ao cenário de  $\{6, 8, 10, 12, 14, 16\}$  *bins*.

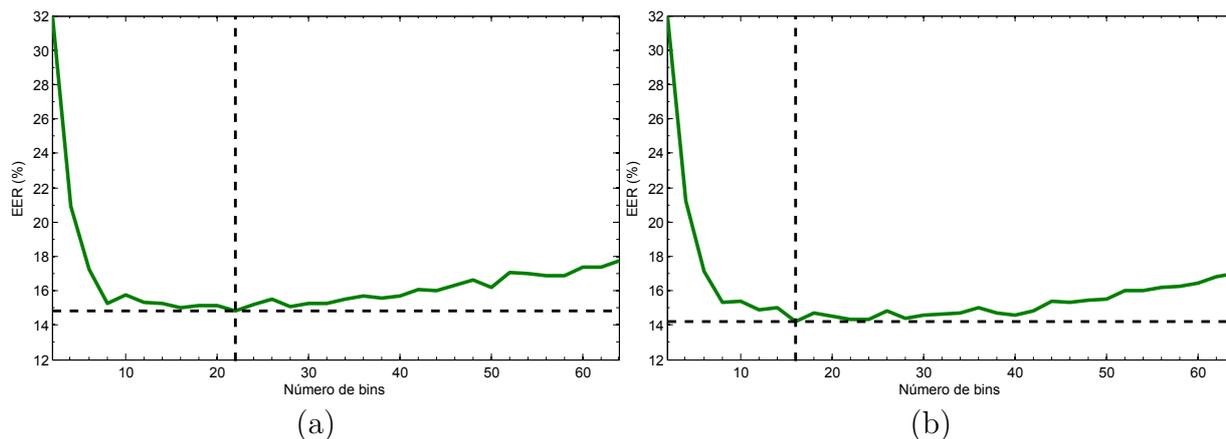


Figura 5.6: EER para determinado número de *bins* para a execução do sistema na base de dados MCYT-100: (a) simulação com distância euclidiana, (b) simulação com distância de Manhattan.

Tabela 5.1: Análise das três estratégias com o uso de histogramas como características, considerando a distância Euclidiana e de Manhattan como métricas de comparação, na MCYT-100, para o conjuntos de bins  $\{6, 8, 10, 12, 14, 16\}$ , tanto para a abordagem completa quanto para a abordagem cascata.

|                             | <b>Manhattan</b> | <b>Euclidiana</b> |
|-----------------------------|------------------|-------------------|
| <b>Número de bins ótimo</b> | 14,2%            | 14,8%             |
| <b>Abordagem completa</b>   | 12,4%            | 12,8%             |
| <b>Abordagem cascata</b>    | 13,6%            | 13,6%             |

Ao analisar a relevância da métrica para a comparação, é possível ver que não há diferenças substanciais em utilizar a distância Euclidiana e a distância de Manhattan. Todavia, em média, a distância de Manhattan possibilitou resultados de erro levemente menores em todos os cenários. A pertinência de determinada métrica para a análise das características é um fato que, por si só, enseja análise. Todavia, foge ao escopo deste trabalho. Desta forma, todos os testes subsequentes consideram somente a distância de Manhattan como métrica escolhida.

A tabela 5.2 mostra os resultados em ambas as bases de dados com a distância de Manhattan, variando os conjuntos de *bins*, em relação ao melhor cenário de histograma isolado para ambas as bases. Os ganhos obtidos nesses cenários podem ser úteis, visto que a escolha de conjuntos de histogramas pode trazer uma representação mais estável, dadas as oscilações no uso de um único número de *bins*. Também é possível ver que a abordagem completa traz melhores resultados que a abordagem cascata, visto que está é tida como mais rápida que a anterior.

Um ponto interessante é que embora outros conjuntos de *bins* se mostrem melhores,

utilizando somente histogramas como características, ao testar o sistema com os outros conjuntos de características, de fato foi o cenário explicitado na Tabela 5.1,  $\{6, 8, 10, 12, 14, 16\}$ , que trouxe os melhores resultados gerais. De qualquer forma, os outros resultados apresentados na Tabela 5.2 são úteis visto que mostram as melhorias trazidas com os métodos desenvolvidos neste trabalho.

Tabela 5.2: Análise das estratégias com o uso de histogramas como características, utilizando a distância de Manhattan, para ambas as bases de dados.  $[b^i; p; b^f]$  representa a faixa de *bins* utilizada, iniciando em  $b^i$  e finalizando em  $b^f$ , com passo de  $p$  nos valores.

| Análise           | Bins     | EER (%)  |         |
|-------------------|----------|----------|---------|
|                   |          | MCYT-100 | LISA-01 |
| Melhor histograma | 16       | 14.2     | 11.6    |
| Completa          | [6;2;16] | 12.5     | 9.5     |
| Completa          | [6;2;30] | 11.7     | 9       |
| Completa          | [6;2;34] | 14.2     | 9       |
| Cascata           | [6;2;16] | 13.6     | 10.3    |
| Cascata           | [6;2;30] | 12.7     | 10.3    |

## 5.4 Experimento 2 - Performance do sistema completo

Esta seção abordará os experimentos realizados com o sistema completo, envolvendo os histogramas, descritores de Fourier e a característica de velocidade. Os experimentos consideram como métrica de comparação a distância de Manhattan e são realizados tanto na base de dados LISA-01 quanto na base de dados MCYT-100.

### 5.4.1 Relevância do *padding* para os descritores de Fourier

A realização do *padding* nos espectros, explicitada na Seção 4.4.2 tem o fim de obter amostras de comprimento igual, de forma que ao comparar dois elementos do espectro, eles representam valores da mesma frequência. A ideia é realizar esse *padding* em relação ao comprimento da maior amostra do perfil [7]. Isso é bastante eficaz ao receber amostras menores do que esse valor. Todavia, ao considerar amostras de comprimento maior, é necessário realizar truncamento dessas amostras, o que pode piorar os resultados.

Para isso, dado o maior comprimento das amostras de referência,  $M$ , considerou-se fazer o *padding* em relação a um valor proporcional a  $M$ , dado por  $M' = a \times M$ , com  $a \in \mathbb{N}$

e  $a > 1$ . A Figura 5.7 mostra os valores de EER para as simulações do sistema variando o valor de  $a$ , tanto para a abordagem completa, quanto para a abordagem cascata, em ambas as bases de dados. Nesse cenário não foi realizada estimativa de *drift* descrita na Seção 4.4.2. O impacto desse método será analisado na Seção 5.4.3.

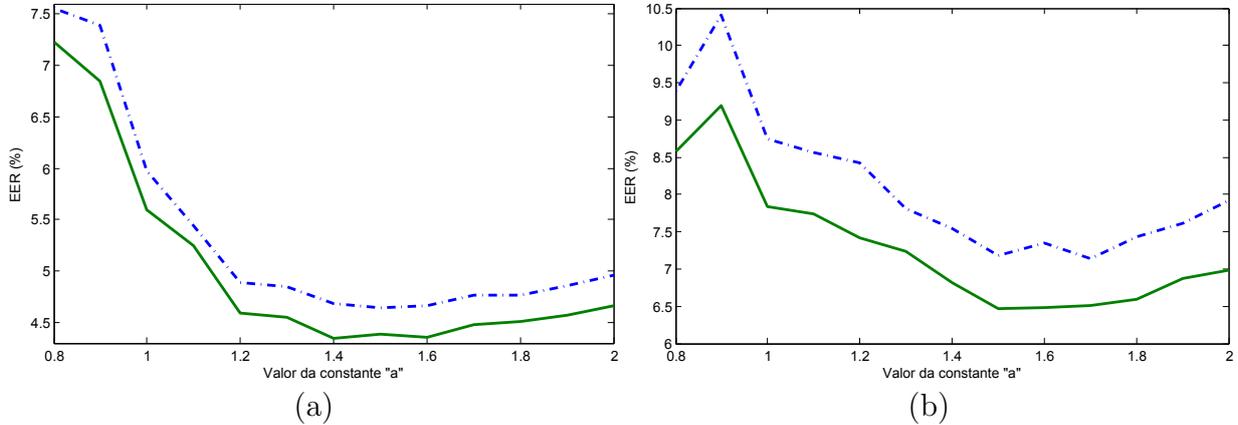


Figura 5.7: Gráficos da influência do tamanho para *padding* a EER do sistema, onde as linhas pontilhadas representam a abordagem cascata e as outras representam a abordagem completa: (a) resultados encontrados na MCYT-100, (b) resultados obtidos na LISA-100.

Embora não tenha sido considerado o valor de  $a < 1$ , foi esboçado nos gráficos para demonstrar a relevância que o truncamento das sequências tem para a performance do sistema. É possível observar que as taxas melhoram, em todos os casos, no intervalo [1.4;1.6]. Como valor médio para todos os cenários, foi escolhido  $a = 1.5$ .

## 5.4.2 Comparação com o estado da arte

Esta seção retratará as comparações do sistema com o estado da arte em ambas as bases, MCYT-100 e LISA-01. É importante salientar que uma comparação direta entre métodos e sistemas para reconhecimento biométrico de texto não é fácil de realizar [7, 33]. Isso porque no geral há grandes diferenças nos cenários, detalhes metodológicos e bases de dados. Todavia, as taxas de EER podem dar uma visão relativa do desempenho do sistema.

### MCYT-100

A Tabela 5.3 apresenta a comparação de alguns trabalhos de reconhecimento *online* de assinatura aplicados na base de dados MCYT-100.

Embora a técnica proposta não tenha os melhores resultados de EER, os resultados são comparáveis aos de estado da arte. É importante notar que a técnica desenvolvida neste trabalho não requer nenhuma fase complexa de treinamento do sistema [34], análise

Tabela 5.3: Abordagens de verificação automática de assinatura na base de dados MCYT-100, onde  $x$ ,  $y$ ,  $t$ ,  $p$ ,  $az$  representam as coordenadas em  $x$ , em  $y$ , o tempo, a pressão e o azimute, respectivamente. Além destas características, estava disponível a altitude da caneta.

| <b>Proposta</b>                          | <b>Características</b>                                | <b>EER</b>   |
|--|---|--------------|
| Histogramas + Quantizador [6]            | $x$ , $y$ , $t$ , $p$                                 | 4.02%        |
| Variâncias [17]                          | $x$ , $y$   | 4.49%        |
| Mistura de Gaussianas + Deformações [76] | Todas disponíveis                                     | 10.86%       |
| Cadeias de Markov [43]                   | Todas disponíveis                                     | 3.37%        |
| Classificador Polinomial [34]            | Todas disponíveis                                     | 4.22%        |
| Fourier [7]                              | Todas disponíveis                                     | 10.89%       |
| DCT + Wavelets [77]                      | $x$ , $y$ , $az$                                      | 9.80%        |
| <b>Abordagem completa</b>                | <b><math>x</math>, <math>y</math>, <math>t</math></b> | <b>4.39%</b> |
| <b>Abordagem cascata</b>                 | <b><math>x</math>, <math>y</math>, <math>t</math></b> | <b>4.64%</b> |

de um grande número de histogramas bidimensionais com técnicas de quantização [6], uso de características locais [43], ou mesmo o uso de características obtidas com canetas especiais [34]. A Figura 5.8 mostra a relação entre a FAR/FRR e os limiares utilizados.

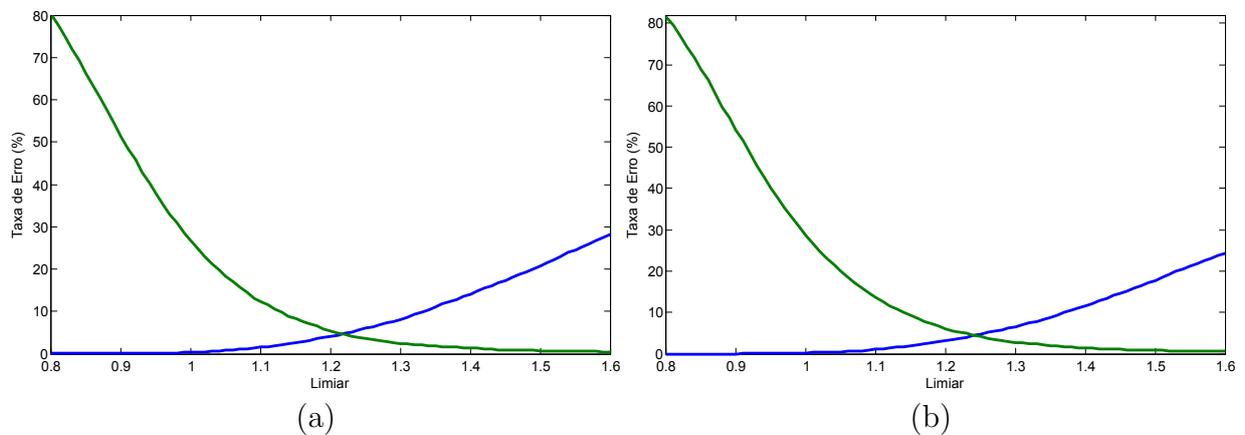


Figura 5.8: Gráfico representando as variações na FAR e FRR em termos do limiar para a base de dados MCYT-100. O ponto de interseção representa a EER: (a) abordagem cascata, (b) abordagem completa.

É possível notar que o comportamento do sistema em relação às mudanças no limiar para ambas as abordagens foi bastante semelhante, embora o valor de EER para a abordagem completa foi ligeiramente menor que o da abordagem cascata.

## LISA-01

Como a base de dados LISA-01 foi desenvolvida neste trabalho, para realizar comparação com algum outro trabalho, foi necessário a implementação e simulação na base. Para escolher o trabalho foi levado em consideração a simplicidade e a não existência de treinamento, duas características do sistema aqui desenvolvido. Dessa forma, o trabalho escolhido foi a abordagem baseada puramente em Fourier que mais recentemente foi aplicada a base de dados MCYT [7]. A Tabela 5.4 apresenta os resultados na LISA-01, considerando o valor EER como o valor médio retornado por cada palavra.

Tabela 5.4: Resultados das simulações para a base de dados LISA-01: palavra 1 - “december”, palavra 2 - “intelligence”, palavra 3 - “love” e palavra 4 - “pattern”.

|                           | Falsificações    | Palavra 1    | Palavra 2    | Palavra 3    | Palavra 4    | Média        |
|---------------------------|------------------|--------------|--------------|--------------|--------------|--------------|
| Fourier [7]               | Igual            | 17,42%       | 18,58%       | 18,14%       | 14,62%       | 17,19%       |
|                           | Diferente        | 16,24%       | 14,00%       | 12,62%       | 11,48%       | 13,58%       |
| <b>Abordagem completa</b> | <b>Igual</b>     | <b>7,18%</b> | <b>5,84%</b> | <b>8,52%</b> | <b>4,36%</b> | <b>6,47%</b> |
|                           | <b>Diferente</b> | <b>4,80%</b> | <b>3,78%</b> | <b>6,04%</b> | <b>3,92%</b> | <b>4,63%</b> |
| <b>Abordagem cascata</b>  | <b>Igual</b>     | <b>8%</b>    | <b>6,20%</b> | <b>9,06%</b> | <b>5,48%</b> | <b>7,19%</b> |
|                           | <b>Diferente</b> | <b>5,74%</b> | <b>4,26%</b> | <b>6,66%</b> | <b>4,16%</b> | <b>5,21%</b> |

É possível notar uma melhora quando as falsificações são feitas por palavras diferentes em relação ao caso quando são feitas por palavras iguais, o que é algo esperado. Além disso, há uma boa melhora entre a abordagem proposta e o uso de Fourier, o que indica que as características utilizadas nesse trabalho são satisfatoriamente complementares.

Há uma ligeira relação entre o comprimento da palavra utilizada e as taxas de erro. Em todas as simulações, a palavra “love”, a menor dentre as palavras, teve taxas de erro razoavelmente maiores do que as outras palavras. Todavia, não parece haver uma relação linear entre o comprimento das palavras e as taxas de erro, o que indica que outros fatores influenciam o fato de uma palavra ser boa ou não para ser utilizada como traço biométrico. A Figura 5.9 mostra a relação entre a FAR e FRR com o limiar para as propostas apresentadas, considerando a palavra “love”.

Novamente é possível notar que a abordagem cascata e completa tem curvas parecidas. Em ambos os casos, os limiares da EER estiveram no entorno do intervalo  $[1, 1 - 1, 2]$ .

### 5.4.3 Análise da estimação de *drift* na performance

Em trabalhos anteriores, foi proposto um modelo de estimação e remoção de *drift* para as amostras de texto [7], que melhorava o desempenho para o reconhecimento biométrico

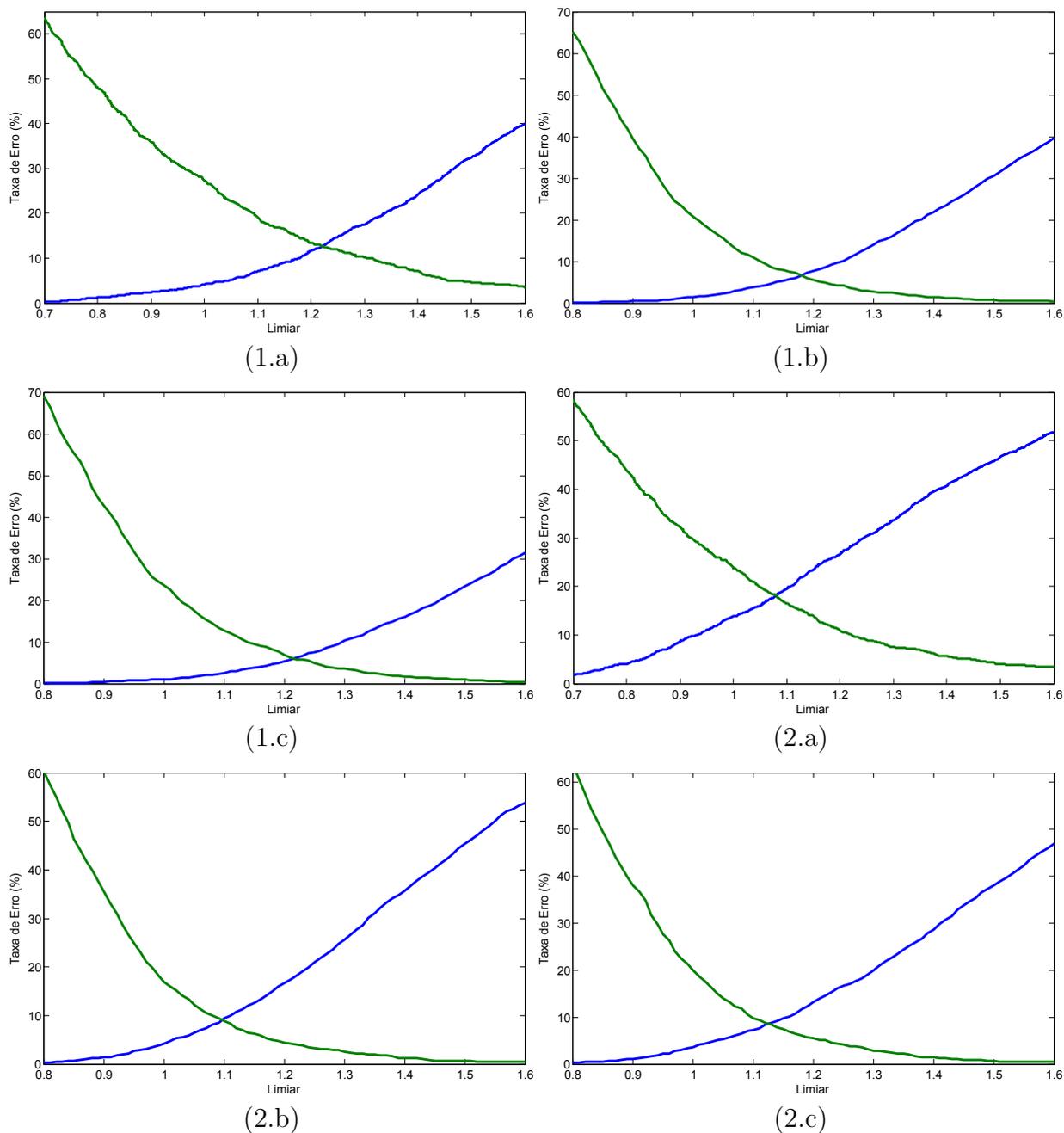


Figura 5.9: Gráfico representando as variações na FAR e FRR em termos do limiar para a base de dados LISA-01. O ponto de interseção representa a EER. (1) e (2) representam o caso onde as amostras são de palavras diferentes e da mesma palavra, respectivamente. (a), (b) e (c) representam a proposta de Fourier, a abordagem cascata e a abordagem completa, respectivamente.

na base de dados MCYT-100. A estimação de *drift* é foco de estudo de vários trabalhos de processamento de sinais na literatura.

Este modelo se baseia em um método de regressão linear por mínimos quadrados e foi descrito na Seção 4.4.2. Foi aplicado este modelo no sistema proposto para avaliar

a pertinência de se utilizar a estimação de *drift* nos cenários avaliados. A Tabela 5.5 apresenta os resultados.

Tabela 5.5: Análise do impacto da estimação de *drift* para a eficiência do Sistema na MCYT-100 e LISA-01. Para a LISA-01, as falsificações são compostas de amostras da mesma palavra.

|                    | Com remoção de drift |          | Sem recomeço de drift |          |
|--------------------|----------------------|----------|-----------------------|----------|
|                    | LISA-01              | MCYT-100 | LISA-01               | MCYT-100 |
| Abordagem completa | 7,33%                | 5,08%    | 6,47%                 | 4,39%    |
| Abordagem cascata  | 8,25%                | 5,26%    | 7,18%                 | 4,64%    |

Dessa forma, o modelo de estimação de *drift* não se mostrou adequado para o sistema em questão, visto que as taxas pioraram em todos os casos. Isso mostra que o modelo de regressão linear por mínimos quadrados não é amplamente eficaz.

## 5.5 Experimento 3 - Comparação da escrita por dedo e caneta

Esta seção descreverá os resultados de um experimento que visou simplesmente analisar se há superioridade em termos de discriminação com os usuários utilizando uma caneta passiva ao invés do dedo. Isso porque as pessoas escrevem com o uso de canetas e lápis, e, sob essa perspectiva há a ideia de que utilizando uma caneta, as amostras seriam mais estáveis e melhores para o reconhecimento biométrico.

Os procedimentos e parâmetros para as simulações realizadas aqui são os mesmos utilizados em todos os outros experimentos e análises. A base de dados utilizada conta com 10 escritores, que escreveram as mesmas palavras da LISA-01 cada uma com 10 repetições. Os doadores das amostras foram solicitados a prover essas amostras utilizando tanto a ponta do dedo, quanto uma caneta simples para *tablets*. Há, dessa forma, um total de 400 amostras nessa base de dados.

As falsificações foram compostas apenas com amostras da mesma palavra. Os resultados são apresentados na Tabela 5.6. As taxas de erro são ligeiramente menores utilizando o dedo. As curvas de FAR e FRR em termos do limiar são exibidas na Figura 5.10 para a palavra 4 (*pattern*) considerando a abordagem cascata, que mostrou maior variabilidade nos resultados.

É importante salientar que a criação de uma base de comparação maior é necessária para ter resultados mais conclusivos a respeito desse fato. De qualquer forma, a pouca

Tabela 5.6: Resultados da comparação de amostras escritas com o dedo e com a caneta, utilizando ambas as propostas descritas neste trabalho.

|                    | Meio de entrada | EER   |
|--------------------|-----------------|-------|
| Abordagem cascata  | Caneta          | 6,42% |
|                    | Dedo            | 5,32% |
| Abordagem completa | Caneta          | 5,02% |
|                    | Dedo            | 4,92% |

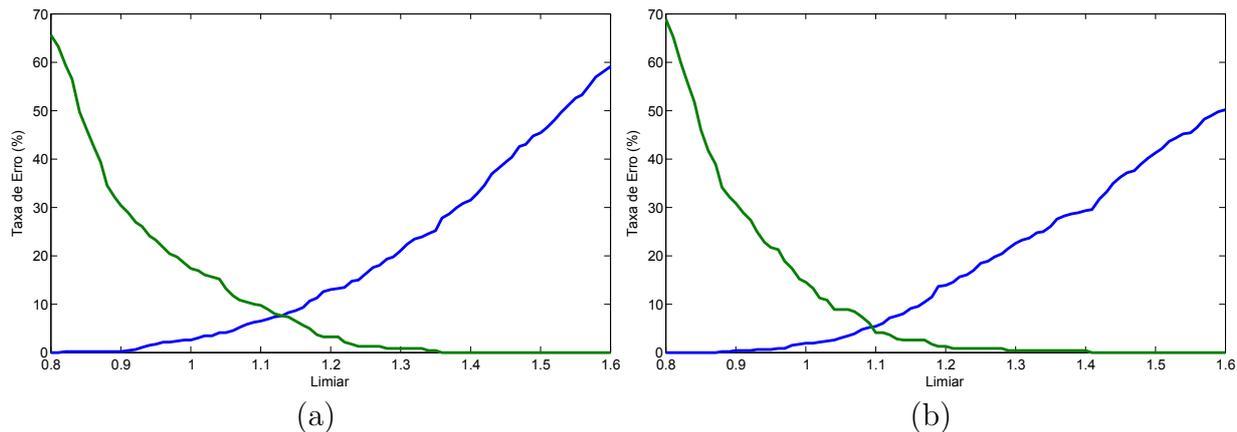


Figura 5.10: Gráfico representando as variações na FAR e FRR em termos do limiar para a base de comparações. O ponto de interseção representa a EER: (a) entrada utilizando a caneta, (b) entrada utilizando o dedo .

diferença apresentada é um fator positivo, visto que abre mais esse leque de flexibilidade para utilizar sistemas baseados em palavras isoladas.

# Capítulo 6

## Conclusão

A possibilidade de identificar pessoas, associando atributos pessoais, tem sido algo importante para a sociedade devido ao grande aumento da população nos últimos anos. O reconhecimento biométrico vem apoiar esta identificação de indivíduos mensurando características da pessoa. E, por este motivo, apresenta vantagens sobre outros métodos comumente usados.

O objetivo deste trabalho foi o desenvolvimento de um sistema de verificação multibiométrica aplicado a palavras cursivas isoladas e assinaturas em um cenário que considera o uso de dispositivos móveis comuns. Os dois pontos chave do trabalho recaem no fato de que palavras isoladas não receberam muita atenção da comunidade científica, assim como a avaliação do reconhecimento em cenários de dispositivos móveis, visto que a maioria dos trabalhos e base de dados usam mesas digitalizadoras e dispositivos especiais para captura.

O Capítulo 2 retratou detalhes a respeito das biometrias, diferenciando as abordagens unibiométricas das multibiométricas, sendo este último o ponto relevante do trabalho. As abordagens multibiométricas se mostram mais robustas no cenário de reconhecimento, embora requeiram maior poder computacional. Especial atenção foi dada a este ponto no trabalho desenvolvido, visto que o cenário visado envolve o uso de dispositivos com poder computacional limitado.

O Capítulo 3 apresentou alguns conceitos inerentes ao trabalho. Num primeiro momento caracteriza a importante área para o processamento de sinais, a Análise de Fourier, da qual são obtidos os descritores aqui apresentados. O algoritmo da FFT é apresentado justificando a sua eficiência para avaliação computacional. Posteriormente os histogramas são apresentados. O uso de histogramas se mostrou algo de interesse visto que foi pouco analisado no campo de reconhecimento *online* de escritor. Por último, as métricas de comparação são apresentadas, das quais a distância Euclidiana e de Manhattan são utilizadas. Ambas podem ser eficientemente calculadas.

A metodologia, referente ao Capítulo 4, apresentou todos os módulos do sistema proposto. Primeiro, o módulo de entrada de dados, que engloba o dispositivo para aquisição dos dados e o método de entrada. O pré-processamento, etapa posterior, foi mantido ao mínimo no trabalho apresentado, visto que há alegações de que certos cálculos diminuem o poder discriminativo. Na etapa de extração de características, foi apresentado o método e todos os cálculos pertinentes à extração dos descritores de Fourier. Os procedimentos para normalização no tempo e espaço, assim como o alinhamento do comprimento dos espectros. A abordagem comum de histogramas, a abordagem completa e a abordagem cascata para os histogramas também foi descrita. Por último, no contexto de características, a criação de um valor global para representar a velocidade, visto que esta se mostra uma importante aliada no reconhecimento *online*.

Para o armazenamento do perfil, várias análises estatísticas foram realizadas e alguns parâmetros de usuário também foram estimados. Para o classificador adotado, uma junção por combinação linear dos traços foi adotada, além da estimação de pesos para os valores intermediários calculados, otimizando, assim, a classificação. No fim, foi obtido um classificador cujo processo é bastante rápido. A etapa de decisão apresentou uma abordagem por limiar global, cuja ideia foi adaptada para a abordagem cascata, construindo limiares intermediários a partir do limiar global. Uma análise de complexidade assintótica foi realizada, apontando que o sistema tem baixa complexidade computacional. O valor  $\mathcal{O}(N \times \log N)$  foi obtido.

O Capítulo 5 apresentou as bases de dados utilizadas, as análises realizadas e a discussão a respeito do que foi obtido. Detalhes da composição da base de dados LISA-01 foram relatados, visto que não havia base de dados desse tipo. O uso de um *tablet* comum foi adotado, sem o uso de canetas especiais, sendo os usuários solicitados a utilizar a ponta do dedo como meio de entrada. Para as assinaturas, foi utilizada uma base de dados já desenvolvida, a MCYT-100.

O primeiro experimento comparou as três abordagens com histogramas. As abordagens proposta, completa e cascata, se mostraram superiores ao uso de histogramas simples, sendo um método pertinente para uso. A abordagem completa trouxe resultados levemente melhores, em média, do que a abordagem cascata. Todavia, o uso da abordagem completa envolve o cálculo de vários níveis de histogramas sendo que na abordagem cascata os níveis com maior número de *bins* só são utilizados caso necessário. A relevância da distância de Manhattan e euclidiana também foi avaliada. Os resultados sugerem que não há grandes diferenças no uso de qualquer uma das duas, embora a distância de Manhattan tenha retornado taxas de erro um pouco inferiores.

A segunda etapa de experimentos avaliou a composição do sistema completo. A combinação dos descritores de Fourier com histogramas (em ambas as abordagens, completa

e cascata) e a velocidade trouxeram resultados satisfatórios, comparáveis aos de estado da arte. Isso mostra que o conjunto de características adotadas é bem complementar, visto que a diferença dos resultados em comparação com o uso simples de descritores de Fourier [7]. Os resultados com a MCYT-100 e a LISA-01 foram comparáveis, o que indica que as palavras cursivas podem ser utilizadas em cenários de avaliação. Além disso, um vínculo do tamanho da palavra com relação às taxas de erros também ficou evidente, embora esse fato não se apresente de maneira linear. Isso indica que outros fatores influenciam na escolha de palavras adequadas para compor os sistemas.

Por último, foi apresentado a criação de uma pequena base de dados para analisar a relevância do método de entrada. A questão levantada dizia respeito a utilização do dedo ou da caneta para compor as bases de dados. Apesar dos resultados levemente superiores com o uso de dedo, os valores apontam que ambos os métodos são viáveis, o que abre mais um leque de flexibilidade no uso de sistemas baseados em palavras isoladas.

## 6.1 Trabalhos Futuros

Há várias possibilidades de melhorias em relação ao trabalho proposto, sejam elas melhorias estruturas, sejam melhorias nas bases de dados:

- Aprimoramento das base de dados. No caso da LISA-01, poder-se-ia montar uma base de dados que tivesse falsificações qualificadas para cada usuário, o que representa um cenário de avaliação bem interessante. Além disso, a composição da base abordando situações adversas de entrada, tais como inclinação nos dados e rapidez de escrita também são pontos interessantes. O uso de palavras diferentes ou de um conjunto maior de palavras também é um ponto positivo para uma base de dados. Por último, a combinação de não só palavras cursivas isoladas, mas também assinaturas e palavras em letra de forma, numa mesma base de dados, é algo desejável.
- Base maior para comparação. Um ponto que pode ser melhor estudado diz respeito à relevância do método de entrada para a performance. Embora aqui tenha sido feita uma avaliação, é desejável a utilização de uma grande base de dados para fazer tal comparação. Os fatos descritos no item anterior, em uma base de entrada bimodal, proveriam melhores análises desse ponto.
- Estudo da relevância de palavras no reconhecimento. A LISA-01 só tinha 4 palavras para avaliação. Os resultados sugerem que o comprimento tem algum impacto notável no desempenho. Todavia, a composição de uma base com mais palavras

permitirá uma análise mais profunda e rigorosa da relevância de palavras para o reconhecimento.

- Variações na composição das características. Um ponto desejável envolve a utilização da Transformada de Fourier de Tempo Curto para tentar extrair mais características das amostras. A análise e proposta de outras estratégias para usar melhor as informações dos histogramas também é um ponto a melhorar. Por exemplo, a utilização de um método de estimação de número de *bins* para cada usuário, ou mesmo o aprimoramento da estratégia em cascata com a adoção de limiares superiores e inferiores são alguns exemplos de aprimoramentos.
- Melhorias no classificador. O classificador adotado pode ser melhorado, por exemplo considerando a comparação de outros pontos críticos do perfil, além do vizinho mais distante, mais próximo e a amostra de perfil. A utilização de outras métricas de comparação, como a distância de Mahalanobis e as métricas Riemannianas também constituem possíveis análises para o classificador.
- Aprimoramentos na fusão multibiométrica. O trabalho aqui apresentado adotou a fusão em nível de *score* por combinação linear das características. Todavia, aspectos multibiométricos tem sido bastante avaliados e outras técnicas e fusão de *score* ou mesmo em níveis diferentes podem ser analisadas e adotadas.
- Utilização de abordagem híbrida. A adoção do sistema aqui proposto em conjunto com um sistema baseado em características locais é um ponto a ser analisado. Por exemplo, o sistema global pode fazer uma primeira etapa de análises, enquanto o módulo de sistema local só é acessado quando as análises da primeira fazem não serem suficientemente distintivas.

# Referências

- [1] Anil K Jain, Arun A Ross, and Karthik Nandakumar. *Introduction to biometrics*. Springer, 2011. ix, 1, 2, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17
- [2] Donato Impedovo and Giuseppe Pirlo. Automatic signature verification: the state of the art. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 38(5):609–635, 2008. ix, 2, 3, 6, 11, 19, 20, 21, 22, 23, 24, 25, 36
- [3] W.K. Härdle and L. Simar. *Applied Multivariate Statistical Analysis*. Springer, 2012. ix, 30, 31, 32, 33, 34, 35
- [4] Meinard Müller. *Information retrieval for music and motion*. Springer, 2007. ix, x, 38, 39, 40, 41
- [5] Enric Sesa-Nogueras and Marcos Faundez-Zanuy. Biometric recognition using online uppercase handwritten text. *Pattern Recognition*, 45(1):128–144, 2012. x, 2, 3, 4, 5, 8, 11, 16, 20, 23, 24, 25, 41, 44, 77, 78
- [6] Napa Sae-Bae and Nasir Memon. Online signature verification on mobile devices. *Information Forensics and Security, IEEE Transactions on*, 2014. x, 3, 4, 5, 24, 25, 35, 36, 47, 48, 73, 78, 83
- [7] Yanikoglu Berrin and Kholmatov Alisher. Online signature verification using fourier descriptors. *EURASIP Journal on Advances in Signal Processing*, 2009. x, xi, 3, 4, 11, 24, 25, 30, 45, 50, 51, 53, 55, 73, 74, 78, 81, 82, 83, 84, 90
- [8] Lambert Schomaker. Advances in writer identification and verification. In *Document Analysis and Recognition, 2007. ICDAR 2007. Ninth International Conference on*, volume 2, pages 1268–1273. IEEE, 2007. 1, 6, 18
- [9] K Sasidhar, Vijaya L Kakulapati, Kolikipogu Ramakrishna, and K KailasaRao. Multimodal biometric systems-study to improve accuracy and performance. *arXiv pre-print arXiv:1011.6220*, 2010. 2, 6, 12, 13, 14
- [10] Santosh P Shrikhande and Hanumant S Fadewar. Personal identification using different biometrics: A review. In *International Journal of Engineering Research and Technology*, volume 3. ESRSA Publications, 2014. 2, 6, 8, 11, 16, 17, 18, 19
- [11] Helen van de Haar, Darelle van Greunen, and Dalenca Pottas. The characteristics of a biometric. In *Information Security for South Africa, 2013*, pages 1–8. IEEE, 2013. 2, 6, 10

- [12] Fathimath Sabena, Ali Dehghantanha, and Andrew P Seddon. A review of vulnerabilities in identity management using biometrics. In *Future Networks, 2010. ICFN'10. Second International Conference on*, pages 42–49. IEEE, 2010. 2, 6, 7, 8, 10
- [13] Tobias Scheidat, Claus Vielhauer, and Robert Fischer. Comparative study on fusion strategies for biometric handwriting. In *Proceedings of the thirteenth ACM multimedia workshop on Multimedia and security*, pages 61–68. ACM, 2011. 2, 6, 13, 77
- [14] Enric Sesa-Nogueras. Discriminative power of online handwritten words for writer recognition. In *Security Technology (ICCST), 2011 IEEE International Carnahan Conference on*, pages 1–8. IEEE, 2011. 2, 3, 4, 8, 11, 20, 23, 24, 25, 77, 78
- [15] Saba Mushtaq and AH Mir. Signature verification: A study. In *Computer and Communication Technology (ICCCT), 2013 4th International Conference on*, pages 258–263. IEEE, 2013. 3, 11, 20, 21, 22
- [16] Asghar Fallah, Mahdi Jamaati, and Ali Soleamani. A new online signature verification system based on combining mellin transform, mfcc and neural network. *Digital Signal Processing*, 21(2):404–416, 2011. 3, 35
- [17] Cangghih Puspo Wibowo, Pitak Thumwarin, and Takenobu Matsuura. On-line signature verification based on forward and backward variances of signature. In *Information and Communication Technology, Electronic and Electrical Engineering (JICTEE), 2014 4th Joint International Conference on*, pages 1–5. IEEE, 2014. 3, 73, 83
- [18] Mariano López-García, Rafael Ramos-Lara, Oscar Miguel-Hurtado, and Enrique Cantó-Navarro. Embedded system for biometric online signature verification. *Industrial Informatics, IEEE Transactions on*, 2014. 3, 73
- [19] Xiangqian Wu, Youbao Tang, and Wei Bu. Offline text-independent writer identification based on scale invariant feature transform. *Information Forensics and Security, IEEE Transactions on*, 2014. 3
- [20] Francisco Zamora-Martínez, Volkmar Frinken, S España-Boquera, MJ Castro-Bleda, Andreas Fischer, and Horst Bunke. Neural network language models for off-line handwriting recognition. *Pattern Recognition*, 47(4):1642–1652, 2014. 3
- [21] Zhong-hua Quan, De-shuang Huang, Xiao-lei Xia, Michael R Lyu, and Tat-Ming Lok. Spectrum analysis based on windows with variable widths for online signature verification. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, volume 2, pages 1122–1125. IEEE, 2006. 4, 24
- [22] CP Wibowo, P Thumwarin, and T Matsuura. On-line signature verification based on angles of pen-motion. In *32nd International Conference on Simulation Technology (JSST)*, 2013. 4, 24
- [23] Mohit Arora, Kulbir Singh, and Guneet Mander. Discrete fractional cosine transform based online handwritten signature verification. In *Engineering and Computational Sciences (RAECS), 2014 Recent Advances in*, pages 1–6. IEEE, 2014. 4, 24, 25

- [24] RK Bharathi and BH Shekar. Discriminative dct: An efficient and accurate approach for off-line signature verification. In *Signal and Image Processing (ICSIP), 2014 Fifth International Conference on*, pages 179–184. IEEE, 2014. 4, 24
- [25] S Rashidi, A Fallah, and F Towhidkhan. Feature extraction based dct on dynamic signature verification. *Scientia Iranica*, 19(6):1810–1819, 2012. 4, 11, 24, 54, 77
- [26] Lobna Haddad, Tarek M Hamdani, and Adel M Alimi. Improved neural based writer adaptation for on-line recognition systems. In *Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference on*, pages 1175–1180. IEEE, 2013. 4, 25
- [27] Julian Fierrez, Javier Ortega-Garcia, Daniel Ramos, and Joaquin Gonzalez-Rodriguez. Hmm-based on-line signature verification: Feature extraction and signature modeling. *Pattern recognition letters*, 28(16):2325–2334, 2007. 4, 25
- [28] Alisher Kholmatov and Berrin Yanikoglu. Identity authentication using improved online signature verification method. *Pattern recognition letters*, 26(15):2400–2408, 2005. 4, 11, 23, 25, 41, 45, 61, 64, 73, 77, 78
- [29] Julian Fierrez-Aguilar, Loris Nanni, Jaime Lopez-Peñalba, Javier Ortega-Garcia, and Davide Maltoni. An on-line signature verification system based on fusion of local and global information. In *Audio-and video-based biometric person authentication*, pages 523–532. Springer, 2005. 4, 6, 24, 35, 47, 73
- [30] Nilson D. Guerin Jr., Flávio de Barros Vidal, and Bruno Luigi Macchiavello. Handwritten text verification on mobile devices. In *Proceedings of International Conference on Computer Vision Theory and Applications (VISAPP)*, 2015. 5
- [31] Kenneth Revett. *Behavioral biometrics: a remote access approach*. John Wiley & Sons, 2008. 8, 11, 14, 17, 21, 22
- [32] Madhavi Gudavalli, AV Babu, SV Raju, and D Srinivasa Kumar. Multimodal biometrics—sources, architecture and fusion techniques: An overview. In *Biometrics and Security Technologies (ISBAST), 2012 International Symposium on*, pages 27–34. IEEE, 2012. 11, 13, 14, 77
- [33] Krzysztof Cpałka and Marcin Zalasiński. On-line signature verification using vertical signature partitioning. *Expert Systems with Applications*, 41(9):4170–4180, 2014. 11, 23, 25, 73, 77, 78, 82
- [34] Emanuele Maiorana, Patrizio Campisi, Daria La Rocca, and Gaetano Scarano. Use of polynomial classifiers for on-line signature recognition. In *Biometrics: Theory, Applications and Systems (BTAS), 2012 IEEE Fifth International Conference on*, pages 265–270. IEEE, 2012. 11, 77, 78, 82, 83
- [35] Marcin Zalasiński and Krzysztof Cpałka. New approach for the on-line signature verification based on method of horizontal partitioning. In *Artificial Intelligence and Soft Computing*, pages 342–350. Springer, 2013. 11, 23, 25, 73, 77, 78

- [36] Hiew Moi Sim, Hishammuddin Asmuni, Rohayanti Hassan, and Razib M Othman. Multimodal biometrics: Weighted score level fusion based on non-ideal iris and face images. *Expert Systems with Applications*, 41(11):5390–5404, 2014. 12, 13, 15, 17
- [37] Minakshi Gogoi and Dhruva Kr Bhattacharyya. Fusion of fingerprint and iris biometrics using binary ant colony optimization. In *Proceedings of the Third International Conference on Soft Computing for Problem Solving*, pages 601–613. Springer, 2014. 12, 13, 15, 16, 17
- [38] Robinson Jimenez Moreno, Andres L Jutinico Alarcon, Jorge Ramirez, and Carlos Andres Arredondo. Access control system using palm of the hand image processing. In *Image, Signal Processing, and Artificial Vision (STSIVA), 2013 XVIII Symposium of*, pages 1–4. IEEE, 2013. 16
- [39] Yang Cai, Xiaoyu Li, Zhenjiang Gong, and Tania Ros Codina. Speaker verification for multi-task interactions. *Interacting with Computers*, 26(2):135–144, 2014. 17
- [40] Roman V Yampolskiy and Venu Govindaraju. Behavioural biometrics: a survey and classification. *International Journal of Biometrics*, 1(1):81–113, 2008. 18
- [41] Farhana Javed Zareen and Suraiya Jabin. A comparative study of the recent trends in biometric signature verification. In *Contemporary Computing (IC3), 2013 Sixth International Conference on*, pages 354–358. IEEE, 2013. 22, 23
- [42] Hifzan Ahmed, Shailja Shukla, and Hari Mohan Rai. Static handwritten signature recognition using discrete random transform and combined projection based technique. In *Advanced Computing & Communication Technologies (ACCT), 2014 Fourth International Conference on*, pages 37–41. IEEE, 2014. 24
- [43] Bao Ly Van, Sonia Garcia-Salicetti, and Bernadette Dorizzi. On using the viterbi path along with hmm likelihood information for online signature verification. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 37(5):1237–1247, 2007. 25, 83
- [44] Marcos Faundez-Zanuy. On-line signature recognition based on vq-dtw. *Pattern Recognition*, 40(3):981–992, 2007. 25
- [45] Christian Gruber, Thiemo Gruber, Sebastian Krinninger, and Bernhard Sick. On-line signature verification with support vector machines based on lcss kernel functions. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 40(4):1088–1100, 2010. 25
- [46] Arti Shivram, Chetan Ramaiah, and Venu Govindaraju. A hierarchical bayesian approach to online writer identification. *Biometrics, IET*, 2(4):191–198, 2013. 25
- [47] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing (3rd Edition)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2006. 26, 27, 28, 29
- [48] P. Duhamel and M. Vetterli. Fast fourier transforms: A tutorial review and a state of the art. *Signal Processing*, 19(4):259 – 299, 1990. 29

- [49] Anika Sharin, Mahmudur Rahman Khan, Hafiz Imtiaz, Mirza Saquib Us Sarwar, and Shaikh Anowarul Fattah. An efficient face recognition algorithm based on frequency domain cross-correlation function. In *Electrical and Computer Engineering (ICECE), 2010 International Conference on*, pages 183–186. IEEE, 2010. 29
- [50] Michael M Bronstein and Iasonas Kokkinos. Scale-invariant heat kernel signatures for non-rigid shape recognition. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 1704–1711. IEEE, 2010. 29
- [51] Zheng Zhang and Mingqiang Yang. Area function fourier descriptors based on contour split. In *Communication Technology (ICCT), 2011 IEEE 13th International Conference on*, pages 823–826. IEEE, 2011. 29
- [52] Sonya Eini, Abdollah Chalechale, and Elham Akbari. A new fourier shape descriptor using smallest rectangle distance. In *Computer and Knowledge Engineering (ICCKE), 2012 2nd International eConference on*, pages 39–43. IEEE, 2012. 29
- [53] Li Dong, Jian Wang, Yuanman Li, and Yuan Yan Tang. Sector projection fourier descriptor for chinese character recognition. In *Cybernetics (CYBCONF), 2013 IEEE International Conference on*, pages 162–167. IEEE, 2013. 29
- [54] Bo Peng, Xizhang Wei, Bin Deng, Kangle Li, Haowen Chen, and Xiang Li. Sinusoidal frequency modulation fourier transform and micro-doppler spectrum estimation of vibrating vehicles. In *Signal and Information Processing (ChinaSIP), 2013 IEEE China Summit & International Conference on*, pages 626–630. IEEE, 2013. 29
- [55] Ismail A Ismail, Mohammed A Ramadan, Talaat S El-Danaf, and Ahmed H Samak. An efficient offline signature identification method based on fourier descriptor and chain codes. *International Journal of Biomedical Engineering and Technology*, 5(1):1–10, 2011. 30
- [56] A.I. Al-Shoshan. Handwritten signature verification using image invariants and dynamic features. In *Computer Graphics, Imaging and Visualisation, 2006 International Conference on*, pages 173–176, July 2006. 30
- [57] David Freedman and Persi Diaconis. On the histogram as a density estimator: L 2 theory. *Probability theory and related fields*, 57(4):453–476, 1981. 30
- [58] W. Härdle. *Nonparametric and Semiparametric Models*. Springer Series in Statistics. Springer Berlin Heidelberg, 2004. 31
- [59] David W Scott. Averaged shifted histograms: effective nonparametric density estimators in several dimensions. *The Annals of Statistics*, pages 1024–1040, 1985. 32
- [60] Bernt Schiele and James L Crowley. Object recognition using multidimensional receptive field histograms. In *Computer Vision—ECCV’96*, pages 610–619. Springer, 1996. 35

- [61] Yu Qiao, Jianzhuang Liu, and Xiaoou Tang. Offline signature verification using online handwriting registration. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–8. IEEE, 2007. 35
- [62] Winston Nelson, William Turin, and Trevor Hastie. Statistical methods for on-line signature verification. *International Journal of Pattern Recognition and Artificial Intelligence*, 8(03):749–770, 1994. 35
- [63] Manesh Kokare, BN Chatterji, and PK Biswas. Comparison of similarity metrics for texture image retrieval. In *TENCON 2003. Conference on Convergent Technologies for the Asia-Pacific Region*, volume 2, pages 571–575. IEEE, 2003. 36, 37
- [64] Vytautas Perlibakas. Distance measures for pca-based face recognition. *Pattern Recognition Letters*, 25(6):711 – 724, 2004. 36, 38
- [65] Stan Salvador and Philip Chan. Toward accurate dynamic time warping in linear time and space. *Intelligent Data Analysis*, 11(5):561–580, 2007. 38, 39
- [66] Marcos Faundez-Zanuy. On-line signature recognition based on vq-dtw. *Pattern Recognition*, 40(3):981 – 992, 2007. 41, 73
- [67] Javier Ortega-Garcia, J Fierrez-Aguilar, D Simon, J Gonzalez, M Faundez-Zanuy, V Espinosa, A Satue, I Hernaez, J-J Igarza, C Vivaracho, et al. Mcyt baseline corpus: a bimodal biometric database. *IEE Proceedings-Vision, Image and Signal Processing*, 150(6):395–401, 2003. 44, 73
- [68] Rizwan Chaudhry, Avinash Ravichandran, Gregory Hager, and René Vidal. Histograms of oriented optical flow and binet-cauchy kernels on nonlinear dynamical systems for the recognition of human actions. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 1932–1939. IEEE, 2009. 47
- [69] Thiago da Rocha, Flavio de Barros Vidal, and Alexandre Ricardo Soares Romariz. A proposal for human action classification based on motion analysis and artificial neural networks. In *Neural Networks (IJCNN), The 2012 International Joint Conference on*, pages 1–6. IEEE, 2012. 47
- [70] Giuseppe Pirlo and Donato Impedovo. Verification of static signatures by optical flow analysis. *Human-Machine Systems, IEEE Transactions on*, 43(5):499–505, 2013. 47
- [71] Yvo Pokern, Andrew M Stuart, and J Harry van Zanten. Posterior consistency via precision operators for bayesian nonparametric drift estimation in sdes. *Stochastic Processes and their Applications*, 123(2):603–628, 2013. 50
- [72] A Webb and B Fox-Kemper. Wave spectral moments and stokes drift estimation. *Ocean Modelling*, 40(3):273–288, 2011. 50
- [73] Dipl-Math Hilmar Mai. *Drift estimation for jump diffusions: time-continuous and high-frequency observations*. PhD thesis, Humboldt-Universität zu Berlin, 2012. 50

- [74] Saeid Rashidi, Ali Fallah, and Farzad Towhidkhah. Authentication based on signature verification using position, velocity, acceleration and jerk signals. In *Information Security and Cryptology (ISCISC), 2012 9th International ISC Conference on*, pages 26–31. IEEE, 2012. 74, 78
- [75] Naser Damer, Benedikt Fuhrer, and Arjan Kuijper. Missing data estimation in multi-biometric identification and verification. In *Biometric Measurements and Systems for Security and Medical Applications (BIOMS), 2013 IEEE Workshop on*, pages 41–45. IEEE, 2013. 77
- [76] Andreas Nautsch, Christian Rathgeb, and Christoph Busch. Bridging gaps: An application of feature warping to online signature verification. In *Security Technology (ICCST), 2014 International Carnahan Conference on*, pages 1–6. IEEE, 2014. 83
- [77] Loris Nanni and Alessandra Lumini. A novel local on-line signature verification system. *Pattern Recognition Letters*, 29(5):559–568, 2008. 83