Jânio Coutinho Canuto

Eficiência da Análise Multifractal na Verificação de Assinaturas Dinâmicas

Dissertação de Mestrado apresentada à Faculdade de Engenharia Elétrica e de Computação da UNICAMP como parte dos requisitos exigidos para obtenção do título de Mestre em Engenharia Elétrica. Área de concentração: Telecomunicações e Telemática.

Banca Examinadora:
Prof. Dr. Jugurta Rosa Montalvão Filho – UFS
Prof. Dr. Lee Luan Ling – UNICAMP
Prof. Dr. Romis Ribeiro de Faissol Attux – UNICAMP

Este exemplar correspondente à redação final da Dissertação/Tese
defendida por: JÂNIO COUTINHO CANUTO
e aprovada através da Comissão Julgada em: 12/08/2010

orientador

Campinas, SP
2010

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA E ARQUITETURA - BAE - UNICAMP

C169e

Canuto, Jânio Coutinho
Eficiência da análise multifractal na verificação de assinaturas dinâmicas / Jânio Coutinho Canuto. -- Campinas, SP: [s.n.], 2010.

Orientador: Lee Luan Ling.
Dissertação de Mestrado - Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Biometria. 2. Multifractais. 3. Assinaturas. I. Lee, Luan Ling. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

Título em Inglês: Effectiveness of multifractal analysis for online signature verification
Palavras-chave em Inglês: Biometrics, Multifractals, Signature
Área de concentração: Telecomunicações e Telemática
Titulação: Mestre em Engenharia Elétrica
Banca examinadora: Jugurta Rosa Montalvão Filho, Romis Ribeiro de Faissol Attux
Data da defesa: 12/08/2010
Programa de Pós Graduação: Engenharia Elétrica

COMISSÃO JULGADORA - TESE DE MESTRADO

Candidato: Jânio Coutinho Canuto

Data da Defesa: 12 de agosto de 2010

Título da Tese: "Eficiência da Análise Multifractal na Verificação de Assinaturas Dinâmicas"

Prof. Dr. Lee Luan Ling (Presidente): _____

Prof. Dr. Jugurta Rosa Montalvão Filho: _____

Prof. Dr. Romis Ribeiro de Faissol Attux: _____

Resumo

A verificação de identidades de forma confiável é cada vez mais necessária em nossa sociedade amplamente interconectada. Nesse contexto, a verificação biométrica é uma proposta alternativa, e mais segura, aos métodos tradicionalmente utilizados, como senhas e cartões. A análise multifractal, por sua vez, tem sido usada com sucesso em diversas aplicações de processamento de sinais, além disso, diversos estudos mostram a presença de características multifractais em processos naturais. Este trabalho tem como objetivo analisar os sinais referentes às assinaturas dinâmicas, provenientes de equipamentos como PDAs e tablet-pcs, sob o prisma da teoria multifractal. É estudada a capacidade de discriminação da característica multifractal na detecção de falsificações de assinaturas, tanto quando usadas isoladamente quanto em conjunto com características tradicionais, num contexto de fusão de informação, com resultados equivalentes ao estado da arte deste tema. Além disso, é realizada uma quantificação, através da teoria da informação, desta capacidade discriminatória. Por fim, é apresentada uma aplicação alternativa da informação multifractal no contexto da biometria: a análise de qualidade das amostras.

Palavras-chave: biometria, multifractais, verificação de assinaturas, fusão de informação.

Abstract

Reliable identity verification is an increasing necessity in our largely networked society. On this topic, biometric verification is a safer alternative to the traditional methods, such as passwords and ID cards. On the other hand, multifractal analysis has been successfully used in a wide range of signal processing applications; moreover, many works show the occurrence of multifractal traits on biological processes. This work aims at analyzing dynamic signature signals collected through devices such as PDAs and tablet-pcs, from a multifractal perspective. A study of the multifractal features discriminative capabilities on signature forgery detection is realized on two scenarios: when it is the unique feature used by the system, and in tandem with traditional features on an information fusion scheme; with results as good as those found in the state of the art of this area. Furthermore, an information theoretic quantification of the discrimination capability is realized. Finally, an alternative application for such features is presented: the evaluation of samples quality.

Keywords: biometrics, multifractals, signature verification, information fusion.

Agradecimentos

À minha namorada, Camila, pelo carinho, apoio e compreensão ao longo destes anos em que estive ausente na maior parte do tempo; por me fazer feliz e não me deixar desistir...

Aos meus pais, Jânio e Ana Márcia, por todo carinho, esforço e suporte para que pudesse chegar até aqui; por sempre apostar em mim...

A todos de minhas famílias, pelo suporte e incentivo...

Ao professor Romis Ribeiro de Faissol Attux, pelo tempo e esforço dedicados à leitura e avaliação deste trabalho; pelas sugestões e pelos ensinamentos...

Ao professor e amigo Jugurta Rosa Montalvão Filho, pelo incentivo e dedicação à minha formação; também pelas valiosas discussões que tanto contribuíram na elaboração deste texto...

Ao meu orientador, professor Lee Luan Ling, pelo apoio, incentivo e pela oportunidade de desenvolver este trabalho...

Aos professores da Faculdade de Engenharia Elétrica da UNICAMP e da Universidade Federal de Sergipe com os quais tive contato, pelos valiosos ensinamentos...

À CAPES pelo suporte financeiro...

A todos os amigos e colegas...

... Obrigado.

Sumário

Lista de Figuras ..... **xi**

Lista de Tabelas ..... **xiii**

Trabalhos Publicados Pelo Autor ..... **xv**

1 Introdução **1**

2 Biometria **5**

2.1 Introdução 5

2.2 Anatomia e Operação de um Sistema Biométrico 7

2.3 Características Biométricas 11

2.4 Análise de Desempenho de um Sistema Biométrico 12

2.5 Vulnerabilidades e Limitações 18

2.6 Sistemas Multi-Biométricos 19

2.6.1 Níveis de Fusão 23

2.7 Reconhecimento de Assinaturas 24

3 Fractais **27**

3.1 Introdução 27

3.1.1 Dimensão de Hausdorff e Dimensão de Capacidade 28

3.1.2 Dimensão de Correlação 30

3.1.3 Dimensão de Informação 31

3.2 Monofractais 32

3.3 Multifractais 34

3.4 Espectro Multifractal 41

3.4.1 Espectro de Hausdorff 41

3.4.2 Espectro de Grandes Desvios 41

3.4.3 Espectro de Legendre 43

3.5 Considerações Finais 43

4 Análise Multifractal de Assinaturas 45
4.1 Introdução 45
4.2 Descrição das Bases de Dados 46
4.3 Estimação da Característica Fractal 51
4.4 Verificação Biométrica de Assinaturas 54
4.4.1 Método DIST 54
4.4.2 Método STAT 59
4.5 Resultados Experimentais 63
4.5.1 Fusão no Nível dos Dados 66
4.5.2 Fusão no Nível das Pontuações 67
4.5.3 Resultados dos Experimentos de Fusão 67
4.5.4 Quantificação da Informação 78

5 Conclusões 85
5.1 Aplicações Alternativas dos Expoentes de Hölder na Verificação de Assinaturas ... 86
5.1.1 Qualidade da Assinatura 86
5.1.2 Amostragem Adaptativa 86
5.2 Resumo 87

Referências Bibliográficas 89

Apêndice A 103

Lista de Figuras

Figura 1.1: Faturamento anula das empresas de biometria no período 2005-2010 (em bilhões de Dólares)..... 1
Figura 1.2: Fatia de mercado por tipo de biometria 3
Figura 2.1: Biometria - "Sua identidade é você"..... 6
Figura 2.2: Diagrama de fluxo de um sistema biométrico genérico 8
Figura 2.3: Visualização dos conceitos de variância intra e inter classe..... 13
Figura 2.4: Visualização do processo de tomada de decisão e dos conceitos de FAR e FRR. 15
Figura 2.5: FAR e FRR em função do limiar $\boldsymbol{\tau}$ para as distribuição da Figura 2.4. A EER é o ponto de cruzamento das curvas..... 15
Figura 2.6: Curva ROC representando a FAR e a FRR da Figura 2.5 A EER é o valor da curva que cruza a diagonal (FAR = FRR) 16
Figura 3.1: Fractais (a) conjunto de Mandelbrot; (b) conjunto de Julia; (c) floco de neve de Koch; (d) triângulo de Sierpinski 28
Figura 3.2: Coleção de conjuntos de cobertura 29
Figura 3.3: Aproximação por elementos infinitesimais (a) Curva aproximada por segmentos de reta de comprimento $\boldsymbol{l}$ (b) Círculo aproximado por quadrados de lado $\boldsymbol{l}$ 30
Figura 3.4: Processos monofractais com diferentes valores do parâmetro de Hurst..... 34
Figura 3.5: Condição de Hölder para diferentes valores de $\boldsymbol{\alpha}$ 36
Figura 3.6: WTMM de um movimento Browniano no instante $\boldsymbol{t0 = 250}$ Pontos azuis são máximos locais, pontos circulados em branco formam a linha de máximos 40
Figura 3.7: Estimação do expoente de Hölder pontual utilizando *wavelets* (a) utilizando WTMM; (b) sem utilizar WTMM. 40
Figura 4.1: Exemplos de assinaturas da base SVC2004. Assinaturas genuínas na parte superior e falsificações na parte inferior..... 48
Figura 4.2: Exemplos de assinaturas da base MCYT-100 Assinaturas genuínas na parte superior e falsificações na parte inferior..... 48
Figura 4.3: Sobreposição de assinaturas genuínas 50
Figura 4.4: Sobreposição de assinaturas falsas e genuínas 50
Figura 4.5: Comparação do espectro multifractal. Assinatura x fBm..... 51

Figura 4.6: Espectro multifractal de diversas assinaturas 52
Figura 4.7: Estimação da função Hölder pelos três métodos propostos..... 53
Figura 4.8: Máximos locais da transformada *wavelet* do sinal senoidal da Figura 4.7..... 53
Figura 4.9: Comparações de dois sinais de comprimentos diferentes realizadas pelo método da reamostragem e DTW 55
Figura 4.10: Superposição das curvas da Figura 4.9 após a sincronização através da reamostragem e do DTW..... 55
Figura 4.11: Restrições de busca local para o DTW implementado 57
Figura 4.12: Matriz de distância acumulada e caminho ótimo para as curvas da Figura 4.9... 58
Figura 4.13: Efeito da inserção do tempo no espaço dos sinais..... 62
Figura 4.14: Estimação da densidade de probabilidade utilizando o método STAT. 62
Figura 4.15: Tempo de execução em função do número de amostras para o método da oscilação 77
Figura 4.16: Tempo de execução em função do número de amostras para o método *wavelet* 77
Figura 4.17: Problema da classificação em um contexto bayesiano 79
Figura 4.18: Problema da classificação como um problema de decodificação..... 79
Figura 4.19: Limite de Fano e Limite de Hellman-Raviv 80

Lista de Tabelas

Tabela 2.1: Probabilidade de erro em função do tamanho da base de dados em uma tarefa de identificação..... 10
Tabela 2.2: Características das modalidades biométricas 12
Tabela 4.1 - Resultados de Verificação Utilizando 1 Variável (Base SCV2004)..... 64
Tabela 4.2 - Resultados de Verificação Utilizando 1 Variável (Base MCYT-100)..... 64
Tabela 4.3: Combinações de dados de um mesmo tipo para a base SVC2004..... 68
Tabela 4.4: Combinações de dados de um mesmo tipo para a base MCYT-100..... 69
Tabela 4.5: Estado da arte para a base MCYT (unibiométrico)..... 70
Tabela 4.6: Resultados da SVC2004 utilizando apenas as informações de posição (X e Y)... 72
Tabela 4.7: Resultados da SVC2004 utilizando todas as informações disponíveis..... 72
Tabela 4.8: Resultados dos algoritmos DIST e STAT seguindo o procedimento da SVC2004 (base de treino apenas) 73
Tabela 4.9: Fusão da informação multifractal com os dados originais (base SVC2004) 73
Tabela 4.10: Fusão da informação multifractal com os dados originais (base MCYT-100) ... 74
Tabela 4.11: Estado da arte para a base MCYT (multibiométrico) 75
Tabela 4.12: Estimação da incerteza sobre a classe dada a pontuação. Variáveis consideradas isoladamente 81
Tabela 4.13: Estimação da incerteza sobre a classe dada a pontuação. Combinações de variáveis..... 82
Tabela 4.14: Remoção de incerteza em bits promovida pela fusão de dados 82

Trabalhos Publicados Pelo Autor

1. J. C. Canuto, L. L. Lee. “Effectiveness of Hölder Function for Online Signature Forgery Detection” In: *Proceedings of the ISSNIP Biosignals and Biorobotics Conference*. pp. 199-203. Vitória. Brasil. 3-5 de Janeiro. 2010.

Capítulo 1

Introdução

A verificação da identidade de um indivíduo de forma confiável é cada vez mais necessária em nossa sociedade amplamente interconectada, seja para adentrar um país, fazer uma transação comercial, ou mesmo uma investigação criminal. Nesse contexto, o advento da verificação biométrica corresponde a uma alternativa aos métodos tradicionalmente empregados, como senhas ou cartões. O objetivo da biometria é determinar a identidade de um indivíduo através da análise de características fisiológicas ou comportamentais do mesmo: deste modo, podemos dizer que a biometria é uma forma de identificação que não pode ser roubada, como ocorre com cartões e documentos, ou esquecida, como senhas; determina-se a identidade da pessoa com base em "quem ela é" e não em "o que ela sabe" ou "o que ela possui" [1].

Esta é uma área de pesquisa bastante interessante, tanto do ponto de vista dos desafios científicos quanto do ponto de vista financeiro, pois seu mercado tem crescido bastante nos últimos anos. Na Figura 1.1, apresentamos o faturamento anual das principais empresas de biometria nos período de 2005 a 2010 [2].

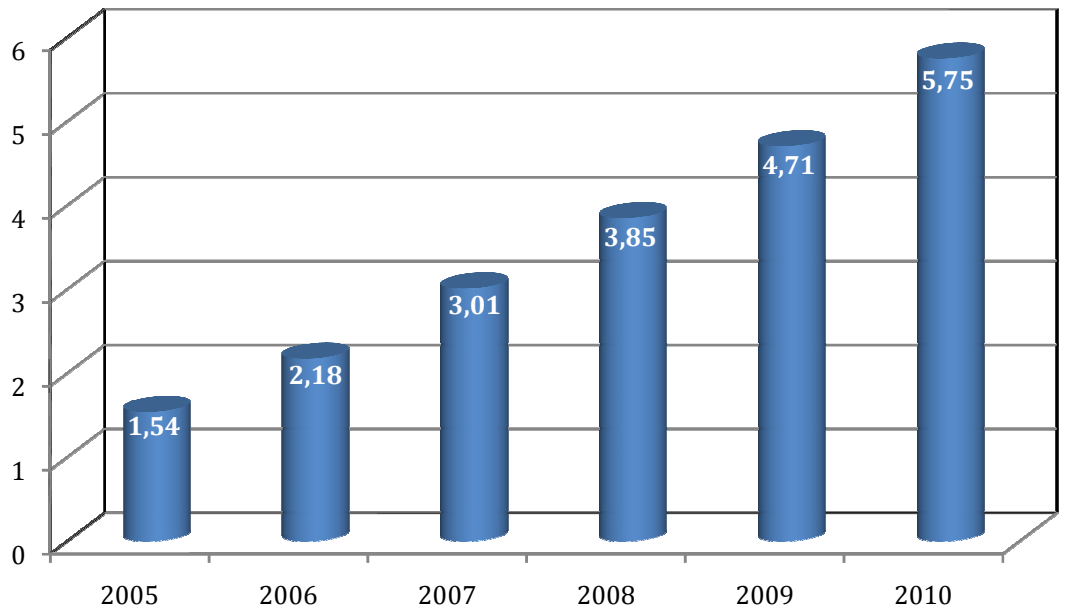


Figura 1.1: Faturamento anula das empresas de biometria no período 2005-2010 (em bilhões de Dólares)

Do ponto de vista científico, a verificação biométrica pode ser vista como um problema de classificação. No entanto, existem diversos aspectos particulares à biometria que tornam a pesquisa nesta área um pouco mais desafiadora:

1. Tipicamente, dispõe-se de poucas amostras de referência para treinamento dos classificadores;
2. O sistema deve ser rápido o suficiente para não frustrar os usuários: seria indesejável esperar 5 minutos para adentrar um recinto, completar uma transação financeira ou ligar o seu carro;
3. O sistema deve ser robusto, pois, além da variação natural dos traços biométricos, uma interação incorreta do usuário com o sistema pode degradar significativamente o sinal capturado. Não bastasse este fato, como tratam-se em geral de sistemas de segurança, eles estarão sujeitos a ataques de falsificação por pessoas que desejem obter acesso ao objeto ou local protegido, de forma ilícita;
4. Como trata-se da análise de dados fisiológicos e/ou comportamentais, é possível utilizar conhecimentos da biologia e/ou da psicologia para melhorar o desempenho dos sistemas;

Estes são apenas alguns dos fatores que, combinados, fazem da biometria uma área altamente interdisciplinar, com diversos problemas instigantes para serem resolvidos, desde a aquisição dos dados até a decisão propriamente dita [3].

As assinaturas, apesar de deterem apenas uma pequena fatia do mercado biométrico, como pode ser visto na Figura 1.2, são especialmente interessantes pelo fato de já serem utilizadas por diversas entidades financeiras, legais e governamentais como meio de autenticação. Além disso, dispositivos com capacidades de escrita natural, como PDAs e *tablet-pcs*, têm se tornado cada vez mais populares, o que faz da assinatura uma escolha óbvia nestes casos [4]. Em tais equipamentos eletrônicos, a assinatura é coletada como uma série de pontos no tempo, e não uma imagem estática, como é o caso da verificação tradicional de assinaturas. Neste trabalho, focaremos em assinaturas dinâmicas, coletadas através de dispositivos eletrônicos adequados, como os que acabamos de citar.

Por atrair pesquisadores dos mais variados campos, uma ampla gama de métodos de processamento de sinais, reconhecimento e classificação de padrões podem ser encontrados na

literatura referente à biometria. No entanto, quase nenhuma atenção foi dada a uma técnica relativamente recente: a análise multifractal.

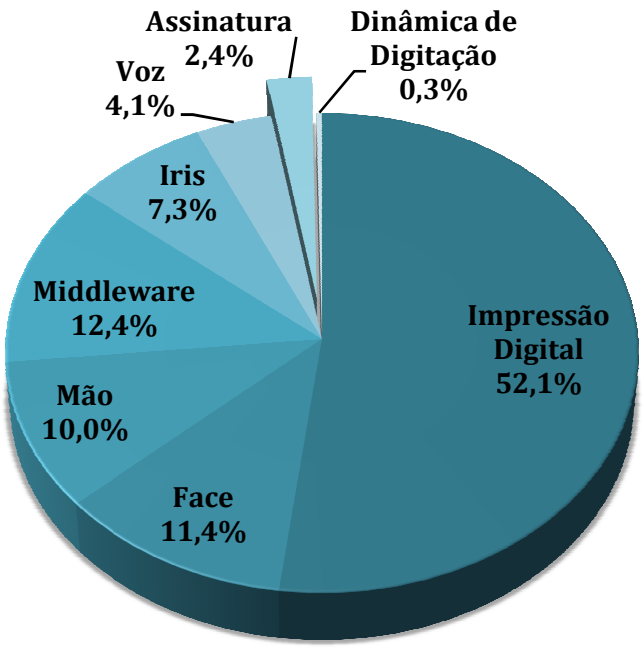


Figura 1.2: Fatia de mercado por tipo de biometria

O estudo dos fractais teve suas bases definidas nas décadas de 60 e 70 por Benoît Mandelbrot [5], e desde então tem sido usado com sucesso na modelagem de diversos fenômenos, como o tráfego em redes de computadores [6], o mercado financeiro [7] e até a análise de imagens artísticas [8]. Além disso, diversos estudos mostram a presença de comportamentos fractais em fenômenos naturais e biológicos, fato este que é utilizado na computação gráfica para a geração de relevo e vegetação artificiais, análise das séries temporais referentes ao batimento cardíaco [9], a distribuição de vasos sanguíneos [10] e até a forma de caminhar das pessoas [11].

Esta dissertação tem como objetivo analisar as assinaturas dinâmicas sob o prisma do formalismo multifractal, avaliando a sua viabilidade e desempenho na detecção de falsificações. Também verificamos se há ganhos na utilização da informação multifractal em conjunto com as características tradicionalmente utilizadas nesta área.

O texto encontra-se organizado da seguinte forma:

- No **Capítulo 2**, apresentamos uma visão geral do campo da biometria, definindo a terminologia utilizada na área, os principais problemas e limitações, a forma como é avaliado o desempenho de um sistema e conceitos sobre sistemas multi-biométricos. No fim deste capítulo, apresentamos com um pouco mais de detalhes o problema da verificação de assinaturas.
- O **Capítulo 3** contém as bases da teoria multifractal. Apesar de ser um assunto relativamente complicado, procuramos, sempre que possível, aliar o formalismo matemático a noções intuitivas dos conceitos apresentados, para facilitar o entendimento dos mesmos.
- No **Capítulo 4**, descrevemos os experimentos realizados e discutimos os resultados obtidos. Utilizamos aqui dois métodos de verificação de assinaturas dinâmicas, um baseado no algoritmo de distorção dinâmica do tempo (DTW – *Dynamic Time Warping)* e outro na modelagem estatística dos dados através do modelo de mistura de gaussianas (GMM – *Gaussian Mixture Model*).
- Por fim, no **Capítulo 5**, tecemos conclusões e propomos possíveis extensões ao uso da informação multifractal. Além disso, resumimos, de forma objetiva, os resultados obtidos e as contribuições advindas desta dissertação.

Capítulo 2

Biometria

2.1 Introdução

A palavra *biometria* vem da junção de dois vocábulo gregos [*βίος* (vida) e *μετρικός* (medida)], e significa medida da vida, ou, o que reflete melhor seu uso na atualidade, é a medida de dados biológicos. Apesar de também ser utilizado em outros contextos, como a medicina e a biologia, este termo têm se tornado cada vez mais comum no dia-a-dia como sinônimo de verificação de identidade.

Tradicionalmente, os métodos de autenticação são baseados em posse (e.g. cartões, distintivos) ou conhecimento (e.g. senhas, frases secretas). Esses sistemas possuem diversos problemas: cartões podem ser perdidos ou esquecidos, resultando em frustração para o usuário, ou ainda roubados e/ou clonados, causando brechas de segurança. As senhas, de modo semelhante, podem ser esquecidas, compartilhadas ou divulgadas de forma não-autorizada, resultando nos mesmos problemas citados anteriormente. Além disso, senhas simples podem ser facilmente adivinhadas, enquanto senhas muito complexas podem ser de difícil memorização para o usuário, o que geralmente culmina com o mesmo anotando-a em algum lugar, tornando ainda mais provável que alguém a descubra.

A biometria, por sua vez, é a ciência de estabelecer a identidade de um indivíduo com base em características fisológicas, tais como face, íris, impressão digital, odores e código genético; ou comportamentais, como a assinatura, a voz, a dinâmica de digitação ou o movimento ao andar [12] [13] [14]. Portanto, é um método alternativo àqueles tradicionais, sendo ela uma técnica de identificação que leva em conta não "o que você tem" ou "o que você sabe", mas sim "quem você é". Na Figura 2.1, apresentamos uma ilustração do conceito. Como muitos dos traços biométricos são "únicos" para cada indivíduo, a biometria provê um sistema de autenticação mais confiável que cartões, chaves ou senhas [15] [16] [17].

Os avanços da tecnologia, em especial as de comunicação e transporte, juntamente com a crescente preocupação acerca de fraudes e segurança, tornaram a verificação confiável da identidade de um indivíduo necessária em uma ampla gama de aplicações, o que vem

estimulando a atividade de pesquisa no campo da biometria [18]. Em decorrência desta necessidade, a utilização de autenticação biométrica em sistemas de segurança tem se mostrado uma tendência mundial [19] [20] [21] [22], sendo aplicada em sistemas de segurança em aeroportos, investigações criminais, substituição de senhas em computadores pessoais, acesso a estabelecimentos (e.g., academias, escolas, hospitais e parques temáticos), autenticação de transações financerias e prevenção de fraudes em eleições. Esta efervescência em torno do tema tem incentivado desenvolvimento e implementação de sistemas cada vez mais robustos e em maior escala.

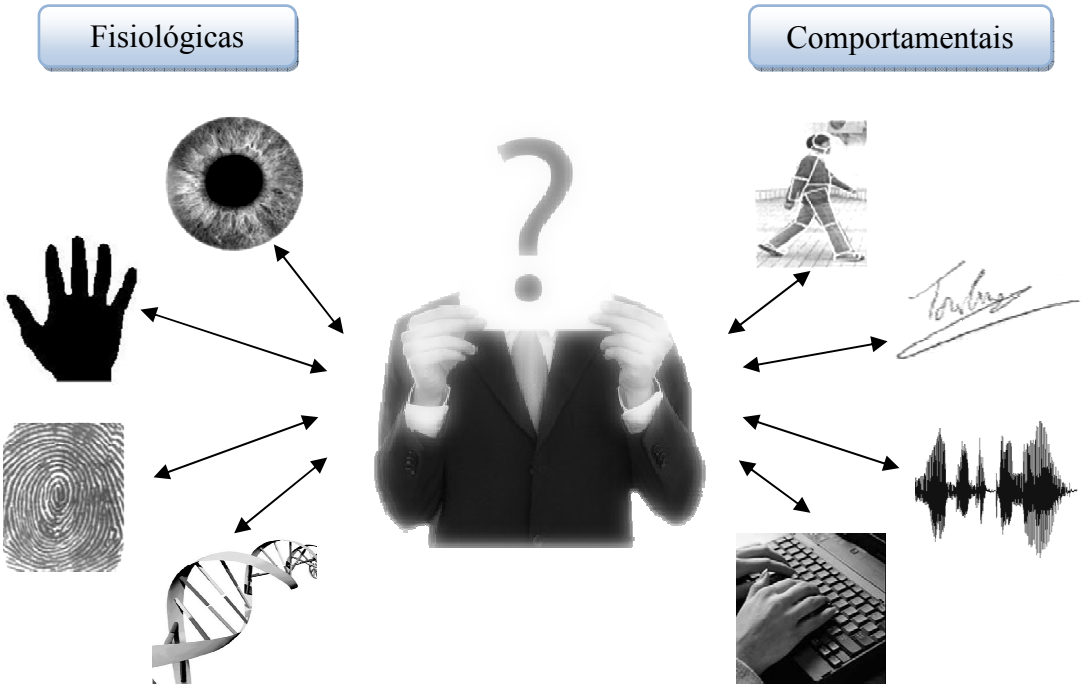


Figura 2.1: Biometria - "Sua identidade é você"

No entanto, esta tendência nem sempre é vista com bons olhos pela sociedade, o que pôs a biometria no centro de um grande debate com “questões orwellianas” referentes à privacidade [23]. Peguntas que ainda precisam de respostas definitivas, como: “As pessoas poderão ser rastreadas?”, “Seus perfis sociais e financeiros serão traçados?”, “Quais as consequências da exposição e/ou compartilhamento dos dados biométricos de um pessoa?”, têm causado desconforto em algumas pessoas quanto ao uso de soluções biométricas em aplicações de grande porte.

Alguns estudos já propõem medidas para reduzir estas inquietações [24] [25], como o uso de um cartão pessoal para armazenar e processar os dados biométricos, o que extingue a preocupação referente à criação das grandes bases de dados, mas traz de volta consigo parte dos problema das técnicas usadas anteriormente (perda/roubo/clonagem de cartões). Além de soluções tecnológicas, diversos governos e associações já estão empenhados em desenvolver leis e regulamentos para evitar a troca, apropriação ou processamento indevido de dados biométricos.

Desde os trabalhos pioneiros de Francis Galton, que, em 1892, propôs o primeiro sistema de identificação através de impressões digitais, até o uso atual da química na decodificação de códigos genéticos, a biometria sofreu grandes transformações do ponto de vista tecnológico, principalmente em decorrência do desenvolvimento dos computadores digitais, mas suas bases teóricas permanecem enraizadas na estatística e no reconhecimento de formas [26] [27]. De fato, apesar dos contínuos avanços, a biometria ainda apresenta desafios relevantes do ponto de vista científico, com questões abertas relativas aos modelos teóricos usados no pré-processamento dos sinais, na extração de características, na otimização de parâmetros das máquinas de classificação e na combinação de informações prvenientes de diversas fontes.

Além de todos estes fatores, os requisitos cada vez mais exigentes e as expectativas crescentes do público em geral acerca destes sistemas, que já figuram em filmes de ficção científica desde o fim da década de 60 (e.g. *2001: Uma Odisséia no Espaço*, filme dirigido por Stanley Kubrick em 1968) como métodos infalíveis de autenticação, tornam a pesquisa nesta área ainda mais desafiadora [28]

2.2 Anatomia e Operação de um Sistema Biométrico

Um sistema biométrico pode ser visto como um reconhecedor de padrões que a) adquire os dados biométricos de um indivíduo; b) realiza um pré-processamento; c) extrai características relevantes dos dados adquiridos; d) compara as características com referências existentes num banco de dados; e) executa uma ação baseada no resultado da comparação. Portanto, existem cinco módulos principais neste sistema, cada um dos quais com desafios suficientes para serem considerados uma área de pesquisa por si só: sensor, pré-processador, extrator de características e classificador. Na Figura 2.2, apresentamos um diagrama de fluxo de um sistema biométrico genérico.

Uma descrição detalhada de cada um destes módulos, com suas particularidades e tecnologias existentes, seria bastante extensa. Deste modo, apresentamos apenas uma visão geral do papel desempenhado por cada um deles no funcionamento do sistema.

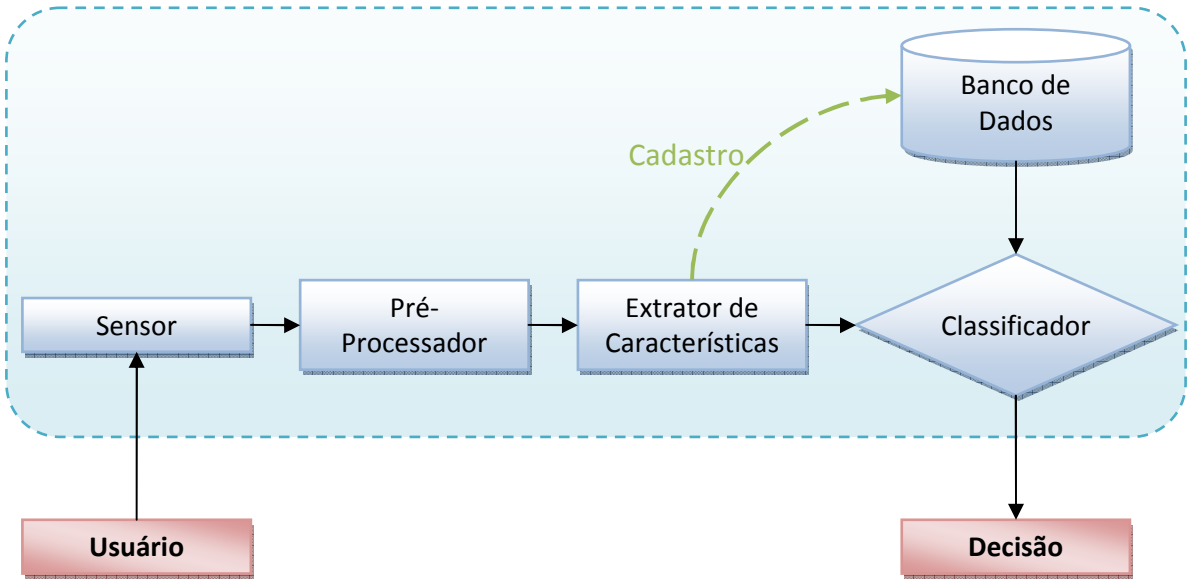


Figura 2.2: Diagrama de fluxo de um sistema biométrico genérico

a) Sensor: Este módulo define a interação homem-máquina, e é, portanto, fundamental para o desempenho do sistema como um todo. Uma interface mal projetada pode resultar em falhas de aquisição e baixa aceitação por partes dos usuários.

b) Pré-Processador: Após a aquisição, realiza-se um pré-processamento do sinal coletado, tipicamente através de técnicas para enfatizar características ou regiões do mesmo. Além disso, pode-se avaliar a qualidade do dado coletado e solicitar uma nova aquisição, caso a amostra obtida seja ruim.

c) Extrator de Características: Sua função é extrair um conjunto de características relevantes suficientes para identificar, com o mínimo de erro possível, um indivíduo. Durante a etapa de cadastro, as características são armazenadas em um banco de dados e são comumentes conhecidas como referências (*templates*).

d) Classificador: No modo de teste, as características obtidas no estágio anterior são comparadas às referências, gerando pontuações (*scores*). A depender da pontuação

obtida, uma decisão é tomada, que é a resposta do sistema. Vale notar que a qualidade dos dados de entrada afeta diretamente estas pontuações.

e) **Banco de Dados:** O desenvolvimento e manutenção do banco de dados é importante para a segurança do sistema e privacidade dos usuários. Em sistemas com um número elevado de usuários, é necessária a utilização de métodos eficientes de busca e agrupamento para tornar o sistema viável.

Conforme pode ser observado na descrição dos módulos e na Figura 2.2, durante a operação destes sistemas, podemos identificar duas fases distintas, o cadastro (*enrollment*) quando os dados de referência são armazenados, e o teste (*query*), quando uma nova entrada é comparada às referências [1] [17].

a) **Cadastro:** Esta etapa consiste em coletar os dados biométricos, extrair suas características e, em seguida, armazená-las no banco de dados ou cartão pessoal do usuário para que possam ser utilizadas como referência na fase de teste. O modelo de referência armazenado pode ser construído a partir de uma única amostra ou gerado a partir de múltiplas entradas. Alguns sistemas armazenam diversos modelos para compensar variações intrínsecas aos dados, como a pose e a iluminação em fotografias. Em geral o cadastro é um processo supervisado, onde um administrador do sistema verifica a identidade do usuário por meios não biométricos (e.g. documentos) antes da realização do mesmo.

b) **Teste:** Nesta fase da operação o usuário tem seus dados novamente coletados e comparados à referência previamente armazenada. Em seguida, um rótulo é atribuído à amostra recém coletada, em função da pontuação obtida como resultado da comparação.

A fase de teste pode ainda ser sub-dividida em dois tipos: autenticação e identificação. O modo de autenticação consiste em validar a identidade alegada pelo usuário. Neste caso, o sistema deve comparar a amostra adquirida com uma única referência e determinar se pertencem à mesma pessoa ou não. Este é o modo de operação mais comum em aplicações de controle de acesso por duas razões principais: a) a comparação do tipo $1:1$ é mais segura que a comparação com toda uma base de dados, pois a probabilidade de ocorrerência de erro, no último caso,

cresce, no mínimo, exponencialmente com o tamanho da base; b) como trata-se de um processo mais rápido (uma única comparação), é possível a utilização de algoritmos de extração de características e de comparação mais elaborados, com uma taxa de erro inferior e, possivelmente, um custo computacional maior por verificação.

No segundo modo de teste, a identificação, o usuário não precisa dizer quem é, o sistema deve determinar, dentre todos os usuários cadastrados no banco de dados, quem é a pessoa utilizando-o. Pode-se também, ao invés de retornar uma única identidade como resposta, selecionar o conjunto das amostras de referências mais semelhantes à amostra de entrada, juntamente com seus graus de similaridade. Este modo de operação é comumente utilizado em aplicações de vigilância, como a busca de um rosto em uma lista de procurados, e em aplicações de atendimento personalizado, como um sistema de atendimento telefônico onde o usuário é automaticamente reconhecido através da sua voz, por exemplo. Esta é uma relação do tipo $1:N$, o que torna necessária a utilização de algoritmos de alta velocidade e métodos de busca eficientes, para que o sistema funcione em um tempo aceitável.

Supondo a existência de N usuários na base de dados, que os usuários são independentes entre si (i.e. não há superposição, no espaço das características, entre usuários diferentes), e que a probabilidade de ocorrência de um erro qualquer (ver Seção 2.4), é P_E, é facil perceber que a probabilidade de executar todas as comparação sem incorrer em erros é dada por $(1 - P_E)^N$, ou seja, para uma base muito grande é virtualmente garantido que um erro ocorrerá. Note que, conforme o número de usuários aumenta, a probabilidade de superposição entre usuários cresce, de modo que P_E também aumentará. Na Tabela 2.1, apresentamos alguns valores para a probabilidade de erro em função do tamanho da base de dados, perceba que mesmo um sistema com erro de apenas 1% não é capaz de atuar de forma satisfatória em uma tarefa de identificação.

Tabela 2.1: Probabilidade de erro em função do tamanho da base de dados em uma tarefa de identificação

Tamanho da Base	$P_E = 1,00\%$	$P_E = 0,10\%$	$P_E = 0,01\%$
10	09,56%	1,00%	0,10%
100	63,40%	9,52%	1,00%
1000	99,99%	63,23%	9,52%
10000	100,00%[1]	99,99%	63,21%

[1] A probabilidade de erro tende a 100%, no entanto, o valor real para este caso é $1 - 2{,}25 \times 10^{-44}$.

2.3 Características Biométricas

O projeto de um sistema biométrico é um problema multifacetado, cada biométrica tem seus prós e contras e, portanto, a escolha de uma característica em particular depende de mais que seu desempenho na classificação [1] [29] [30]. Jain *et al.* [12] identificaram os principais fatores que influenciam a adequação de uma característica a uma dada aplicação biométrica:

a) **Universalidade:** Todos os usuários da aplicação devem possuir a característica;

b) **Unicidade:** A característica deve ser distinta o suficiente entre os indivíduos para que seja possível a realização de uma classificação correta;

c) **Permanência:** Idealmente, o traço biométrico deve ser invariante ao longo do tempo. Numa visão mais realista, admitem-se variações desde que estas não afetem significativamente o desempenho do sistema;

d) **Mensurabilidade:** A coleta dos dados biométricos deve causar o mínimo de desconforto possível ao usuário;

e) **Aceitabilidade:** Os usuários devem estar dispostos a apresentar a característica biométrica ao sistema;

f) **Circunvenção:** Refere-se à dificuldade de se reproduzir e, consequentemente, burlar o sistema.

Apesar de diversos sistemas biométricos já se encontrarem em uso, é bom notar que nenhuma biométrica é capaz de satisfazer todos estes requerimentos para todos os cenários de aplicação [16]. Portanto, a escolha da biometria é fortemente dependente da aplicação que se tem em vista, já que envolve não somente dificuldades técnicas como também sociais e culturais [16] [31] [17]. Na Tabela 2.2 estes requisitos são apresentados, em termos qualitativos, para diversas modalidades biométricas [1].

As modalidades biométricas podem ser divididas em duas classes: fisiológicas e comportamentais. As primeiras são baseadas em traços biológicos como impressão digital, formato da face, geometria da mão, padrão de veias da retina e padrão de rugas da íris. Os últimos consideram, como seu próprio nome sugere, traços comportamentais como os padrões acústicos da voz, assinaturas e a dinâmica de digitação [19] [1] [31] [17] [21]. Tipicamente, os

traços comportamentais são mais suscetíveis a variações ao longo do tempo, pois dependem não somente das condições físicas como também das condições psicológicas do indivíduo.

Além de escolher a característica adequada para a aplicação alvo, o projeto de um sistema biométrico deve considerar outros pontos desejáveis, como desempenho, custo e velocidade [32] [31].

Tabela 2.2: Características das modalidades biométricas

	Universalidade	Unicidade	Permanência	Mensurabilidade	Aceitabilidade	Circunvenção
Face	Alta	Baixa	Média	Alta	Alta	Baixa
Impressão Digital	Média	Alta	Alta	Média	Média	Alta
Geometria da Mão	Média	Média	Média	Alta	Média	Média
Dinâmica de Digitação	Baixa	Baixa	Baixa	Média	Média	Média
Veias da Mão	Média	Média	Média	Média	Média	Alta
Íris	Alta	Alta	Alta	Média	Baixa	Alta
Retina	Alta	Alta	Média	Baixa	Baixa	Alta
Termograma Facial	Alta	Alta	Baixa	Alta	Alta	Alta
Voz	Média	Baixa	Baixa	Média	Alta	Baixa
Código Genético	Alta	Alta	Alta	Baixa	Baixa	Baixa
Assinatura	Baixa	Baixa	Baixa	Alta	Alta	Baixa

2.4 Análise de Desempenho de um Sistema Biométrico

No contexto da biometria, raramente encontram-se duas amostras que resultem em um mesmo conjunto de características. Diversos fatores são responsáveis por essas diferenças, como imperfeições dos sensores, alterações na característica biométrica (e.g., pelos faciais alteram a aparência da face), mudanças no ambiente (e.g., iluminação) e variações na interação do usuário

com o sistema (e.g., mudança da orientação da mão ou abertura dos dedos em um sistema baseado na geometria da mão). De fato, a ocorrência de um casamento perfeito de todos estes fatores é tão rara que pode ser um indicativo de tentativa de fraude, no que se costuma chamar de ataque de repetição (*replay attack*), quando um usuário tenta usar uma cópia do traço biométrico apresentado por um usuário autêntico.

Esta variabilidade observada nas características extraídas de um mesmo indivíduo é chamada de variância intra-classe, enquanto as diferenças entre amostras de duas pessoas diferentes é chamada de variância inter-classe. Uma boa característica biométrica apresenta pequenas variações intra-classe e grandes variações inter-classe, ou seja, grupos compactos e bem separados, o que permite um funcionamento correto dos classificadores. Este conceito é ilustrado na Figura 2.3.

As pontuações obtidas no comparador podem ser de dois tipos: similaridade ou dissimilaridade, que, conforme seus nomes indicam, representam situações opostas. Pontuações de similaridade medem o grau de semelhança entre duas amostras, enquanto dissimilaridades, ou distâncias, representam o nível de disparidade entre as mesmas. No entanto, é sempre possível converter pontuações de similaridade em dissimilaridade e vice-versa. No restante deste texto, o termo "pontuação" se referirá a dissimilaridades, a menos que seja especificado o contrário.

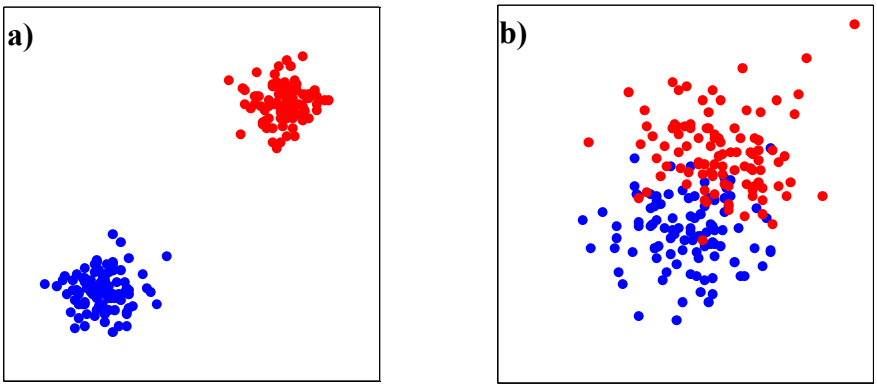


Figura 2.3: Visualização dos conceitos de variância intra e inter classe.

a) Baixa variância intra-classe e alta variância inter-classes: grupos compactos e bem definidos
b) Alta variância intra-classe e baixa variância inter-classes: grupos dispersos e sem uma fronteira bem definida

Obtidas as pontuações, o sistema decide se o usuário é genuíno ou impostor a partir de uma operação de limiarização: pontuações abaixo de um dado limiar τ são rotuladas como genuínas e, caso contrário, recebem o rótulo de impostoras.

Denominaremos as pontuações resultantes da comparação das características de um mesmo usuário de Pontuações Genuínas, e aquelas advindas da comparação das características de dois usuários distintos de Pontuações Impostoras. A decisão tomada pelo sistema pode incorrer em dois tipos de erro: uma Falsa Aceitação, quando uma pontuação impostora está abaixo do limiar τ, ou uma Falsa Rejeição, quando uma pontuação genuína é superior a τ.

Define-se a Taxa de Falsa Aceitação (FAR – *False Acceptance Rate*) de um sistema biométrico como a fração das pontuações impostoras inferiores ao limiar, ou seja, é uma estimativa da probabilidade de ocorrência deste tipo de erro. De forma análoga, a Taxa de Falsa Rejeição (FRR – *False Rejection Rate*) pode ser definida como a fração das pontuações genuínas que excedem o limiar adotado.

Os complementos destas duas probabilidades estimadas anteriormente são, respectivamente, a Taxa de Aceitação Corretas (CAR – *Correct Acceptance Rate*) e a Taxa de Rejeições Corretas (CRR – *Correct Rejection Rate*). Portanto, temos que $\mathrm{CAR} = 1 - \mathrm{FAR}$ e $\mathrm{CRR} = 1 - \mathrm{FRR}$.

Quando dispomos de um número suficientemente de pontuações, podemos estimar as Funções de Densidade de Probabilidade (pdf – *Probability Density Function*) dos dois conjuntos, genuíno e impostor, e deduzir analiticamente a FAR e a FRR. Sejam $p(s|genuíno)$ e $p(s|impostor)$ as pdf condicionais da pontuação s para os casos genuíno e impostor, respectivamente. Então, para um dado limiar:

$$\mathrm{FAR}(\tau) = \int_{-\infty}^{\tau} p(s|impostor)ds \tag{2.1}$$

$$\mathrm{FRR}(\tau) = \int_{\tau}^{\infty} p(s|genuíno)ds \tag{2.2}$$

Na Figura 2.4, apresentamos uma interpretação visual destes conceitos. É fácil perceber, a partir da observação desta ilustração, que uma mudança no limiar τ altera os valores das taxas de erro, mas, para um dado sistema, não é possível diminuir os dois tipos de erro simultaneamente. Na Figura 2.5 apresentamos a FAR e a FRR para as distribuições ilustradas na Figura 2.4 em função do limiar τ.

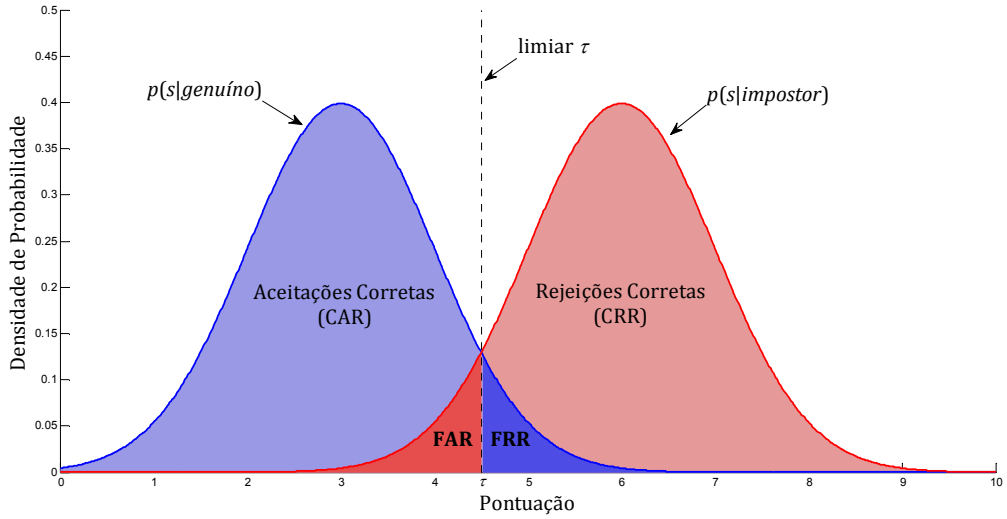


Figura 2.4: Visualização do processo de tomada de decisão e dos conceitos de FAR e FRR

É comum também a utilização de uma única curva para resumir o comportamento da FAR e da FRR em diversos valores de τ. A curva ROC (*Receiver Operating Characteristic*) [33] exibe os valores de FRR x FAR para diversos limiares diferentes. A curva ROC referente às curvas de FAR e FRR da Figura 2.5 é apresentada na Figura 2.6.

Também é possível resumir o desempenho do sistema utilizando um único número, como a EER (*Equal Error Rate*) ou o valor d'. A EER é o ponto da curva ROC no qual a FAR e a FRR são iguais, portanto, uma EER mais baixa é sinônimo de um melhor desempenho na classificação. A EER está indicada tanto na Figura 2.5 quanto na Figura 2.6.

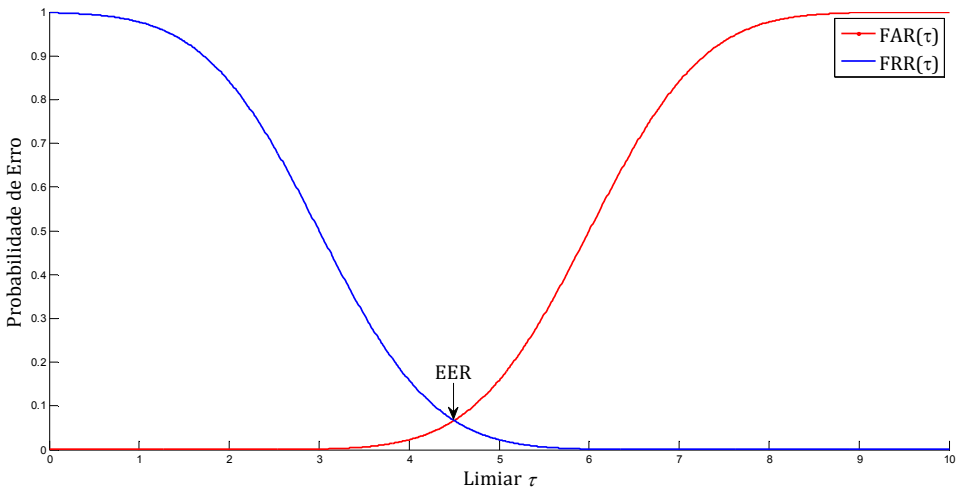


Figura 2.5: FAR e FRR em função do limiar τ para as distribuição da Figura 2.4. A EER é o ponto de cruzamento das curvas

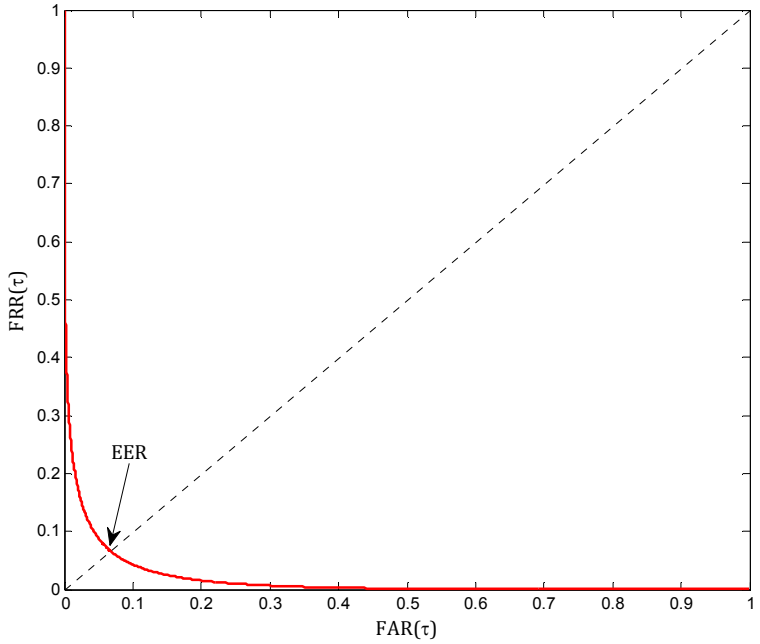


Figura 2.6: Curva ROC representando a FAR e a FRR da Figura 2.5
A EER é o valor da curva que cruza a diagonal (FAR = FRR)

O valor d', que citamos no parágrafo anterior, representa a distância entre as médias das distribuições de probabilidade genuína e impostora em unidades de desvio padrão, e é definido como [34]:

$$d' = \frac{\left|\mu_{genuíno} - \mu_{impostor}\right|}{\sqrt{\frac{1}{2}\left(\sigma^2_{genuíno} + \sigma^2_{impostor}\right)}} \tag{2.3}$$

onde μ e σ são média e desvio padrão, respectivamente. Quanto maior for d', maior a separação entre as pontuações genuínas e impostoras e, consequentemente, melhor o desempenho do sistema. Para as distribuição da Figura 2.3, por exemplo, (a) apresenta $d' = 14{,}14$ enquanto (b) tem $d' = 2{,}24$.

É interessante notar que a ocorrência de falsas aceitações e rejeições não é uniformemente distribuída na população de usuários. Como o leitor deve perceber até mesmo por experiência pessoal, existem pessoas mais fáceis de reconhecer do que outras, com características mais marcantas. Doddington *et al.* [35] identificam quatro categorias de usuários (conhecidas como zoológico de Doddington) baseados nestas diferenças inerentes:

a) **Ovelhas:** São usuários cujas características biométricas são bastante distintas das demais pessoas e com baixa variância intra-classe, o que resulta em baixo FAR e FRR.
b) **Cabras:** Usuários com alta variância intra-classe, causando, consequentemente, uma alta FRR.
c) **Cordeiros:** Estes possuem baixa variância inter-classe, suas características sobrepõe-se àquelas de outros usuários, de modo que são sujeitos a uma alta FAR.
d) **Lobos:** Usuários hábeis em manipular suas biométricas para simular aquelas dos usuários cadatrados. A presença deste tipo de usuário promove um aumento na FAR do sistema.

Além dos erros de classificação, há também os de aquisição (FTA – *Failure To Acquire*) e de cadastro (FTE – *Failure To Enroll*), que são, em geral, resultantes de falhas no sensor ou interação incorreta do usuário com o sistema. Caso exista no sistema um módulo de avaliação da qualidade da amostra, o limiar de qualidade afetará estes dois tipos de erro, além da FAR e FRR percebidas. Caso o módulo aceite apenas amostras de qualidade muito elevada, por exemplo, o sistema possuirá, aparentemente, baixas FAR e FRR, no entanto, os sucessivos erros de aquisição e cadastro resultantes deste rigoroso padrão de qualidade tornarão a experiência de uso do sistema desgastante para o usuário.

Todos os tipos de erro aqui discutidos são parâmetros importantes no projeto de um sistema biométrico, e, em geral, são eles que guiam as escolhas posteriores. Além disso, a depender da aplicação em questão, diferentes tipos de erro podem ser tolerados: em uma instalação militar, por exemplo, é geralmente preferível uma alta taxa de rejeição que de aceitação, pois as consequências do acesso não-autorizado podem ser bastante graves. Já em um parque de diversões, é mais importante manter o cliente satisfeito, e é preferível uma taxa de falsas aceitações superior à de falsas rejeições. Um mesmo sistema poderia ser utilizado em ambos os casos, porém com escolhas diferentes do limiar τ.

Além de taxas de erro e dos quesitos abordados acerca das características biométricas, fatores de cunho mais práticos, como custo, facilidade de uso, qualidade do sensor e velocidade também afetam a adequação do sistema à aplicação.

2.5 Vulnerabilidades e Limitações

Com a proliferação de soluções biométricas em diversas aplicações, é importante entender as vulnerabilidades e limitações destes sistemas. O desempenho de um sistema biométrico é afetado por diversos fatores, incluindo ruído, variações intra-classe e interação incorreta do usuário com o sistema [13]. A seguir, listamos as principais dificuldades encontradas no desenvolvimento de uma solução biométrica.

a) **Ruído:** A amostra coletada pode apresentar ruídos devido ao ambiente de operação, ao sensor ou até mesmo variações no traço biométrico. Este tipo de problema tende a aumentar a FRR do sistema.
b) **Não-Universalidade:** Pode ser impossível para o sistema coletar dados de alguns indivíduos da população alvo, o que resulta em erros FTE e FTA. Para tratar tais casos é necessária a inclusão de exceções no algoritmo, ou mesmo a troca de modalidade biométrica.
c) **Limite de Desempenho:** Um sistema biométrico não pode ser indefinidamente melhorado através de refinamentos sucessivos dos módulos de extração de características e comparação. Existe um limite implícito para o desempenho de qualquer sistema biométrico. Jain et al. [28] sugerem três razões para isso:
    I. **Informação Limitada:** A quantidade de informação disponível em uma biométrica é naturalmente restrita. A geometria da mão, por exemplo, possui uma capacidade de discriminação menor que as impressões digitais [36].
    II. **Representação Limitada:** O esquema ideal de representação para um traço biométrico deve reter toda a informação discriminatória do mesmo. Os sistemas de extração existentes, tipicamente baseados em modelos simplistas dos dados biométricos, não são capazes disto, resultando em uma inclusão de dados redundantes ou espúrios, e a exclusão de características relevantes. Em consequência deste fato, uma fração do espaço de caracaterísticas não pode ser tratado pelo sistema, resultando em erros de autenticação.
    III. **Comparador Limitado:** Dado um esquema de representação, um comparador ideal deveria modelar perfeitamente a relação invariante entre diferentes padrões

provenientes do mesmo usuário. Na prática, no entanto, os comparadores não possuem tal capacidade, resultando em baixo desempenho.

d) Ataques: Características comportamentais como voz [37] e assinatura [38] são vulneráveis a ataques por impostores capazes de imitar as características dos usuários legítimos. Características físicas, por sua vez, podem ser falsificadas através da utilização de representações artificiais da biométrica. As impressões digitais, por exemplo, podem ser falsificadas através da inscrição de estruturas semelhantes às reentrâncias da pele em materiais sintéticos tais como gelatina e massa-de-modelar [39], [40]. Ataques como estes podem minar a segurança provida pelo sistema e, consequentemente, mitigar seus benefícios [41].

Algumas das limitações dos sistemas biométricos que foram considerados até agora, que chamaremos de unibiométricos, podem ser superadas através do projeto de um sistema que utilize múltiplas fontes de informação biométrica. Isto pode ser realizado através da fusão, por exemplo, de múltiplos sistemas biométricos que utilizem diferentes características, ou múltiplos algoritmos de extração de características e classificação operando na mesma biométrica, dando origem ao que se designa multi-biometria [42] [43] [44].

Tais métodos de acúmulo de evidências podem resultar em uma melhoria no desempenho do sistema como um todo, além de aumentar a cobertura da população e dificultar a realização de ataques.

2.6 Sistemas Multi-Biométricos

A fusão de informação é um campo de pesquisa que atrai o interesse de estudiosos de diversas áreas distintas, incluindo a previsão do tempo [45], o rastreamento de objetos [46] e a navegação de robôs [47]. O princípio da resolução de problemas através da combinação das decisões de diversos especialistas recebeu bastante atenção da comunidade científica: de fato, a teoria de sistemas com múltiplos classificadores (MCS – *Multiple Classifiers System*) já foi rigorosamente estudada em diversos trabalhos, como [48], [49], [50] e [51].

No âmbito da biometria, a consolidação de evidências apresentadas por múltiplas fontes é uma forma eficiente de melhorar o desempenho do sistema de reconhecimento. Alguns dos primeiros sistemas biométricos multi-modais descritos na literatura combinavam face e voz [52]

[53], de modo que, além do ganho de desempenho, também eram capazes de cobrir uma fatia maior da população. Um exemplo de sistema multibiométrico já em uso com sucesso é o IAFIS mantido pelo FBI, que integra a informação presente na impressão digital de diversos dedos para realizar a identificação dos suspeitos. Algumas das vantagens da multibiometria são listadas a seguir [42].

- Atenuam a questão da não-universalidade encontrada em sistemas uni-biométricos;
- Facilitam a filtragem ou indexação de bases de dados muito grandes;
- Torna-se mais difícil para um impostor burlar o sistema;
- Atenuam o problema do ruído;
- Ajudam no monitoramento contínuo ou rastreamento de um indivíduo em situações onde uma única característica não é suficiente;

Um sistema multibiométrico pode ser visto como um sistema tolerante a falhas, que pode continuar a operar mesmo quando algumas fontes biométricas tornam-se menos confiáveis, seja devido a uma falha de sensor, do algoritmo ou da manipulação deliberada do usuário. O conceito de tolerância a falhas é especialmente útil em sistemas de larga escala, envolvendo um grande número de usuários, como aplicações de controle de imigração em aeroportos.

O projeto de um sistema multibiométrico é guiado por diversos fatores, que incluem: a) a interface homem-máquina; b) o compromisso entre o custo adicional da introdução de novas fontes e o ganho obtido; c) as fontes de informação biométrica utilizadas; d) o nível de fusão (i.e., o tipo de informação que será fundida); e) o método de fusão adotado.

Baseado na natureza das fontes utilizadas para fusão, os sistemas multibiométricos podem ser classificados em seis categorias [42]: multisensor, multi-algoritmo, multi-instâncias, multi-amostras, multimodal e híbrido.

a) **Multisensor:** Emprega diversos sensores para capturar um único traço biométrico. Alguns exemplos encontrados na literatura são a utilização de várias câmeras para fotografar a face do indivíduo [54]; a combinação de sensores infra-vermelho com sensores de luz-visível para obter informação sobre a face [55], [56], [57]; o uso de câmeras multi-espectrais para coleta de imagens da íris, impressão digital ou face do usuário [58]. O uso de múltiplos sensores, em alguns casos, pode resultar na aquisição

de informações complementares capazes de melhorar a capacidade de reconhecimento do sistema.

b) **Multi-Algoritmo:** A utilização de múltiplos extratores de características e/ou algoritmos de classificação também podem melhorar o desempenho do sistema. Sistemas multi-algoritmo consolidam os dados provenientes de diversos extratores de características ou classificadores operando em um mesmo conjunto de amostras. O desenvolvimento deste tipo de sistema não requer o emprego de novos sensores e, portanto, tem custo inferior às outras formas de multibiometria. Por outro lado, a introdução de novos algoritmos pode aumentar demasiadamente a complexidade computacional, tornando o sistema inviável. Ross et al. [59] descrevem um sistema de reconhecimento de impressões digitais que utiliza tanto as minúcias quanto a informação de textura para representar e classificar as imagens.

c) **Multi-Instância:** Estes sistemas empregam múltiplas instâncias do mesmo traço biométrico, sendo também conhecidos como sistemas multi-unidade. O sistema IAFIS, citado anteriormente, é um exemplo desta modalidade de fusão.

d) **Mutli-Amostras:** Um único sensor pode ser usado para coletar diversas amostras do mesmo traço biométrico com o intuito de compensar variações que possam ocorrer na característica medida, ou mesmo para obter uma representação mais completa da mesma. Uma questão crucial nesta categoria multibiométrica é a determinação do número de amostras que devem ser coletadas. É importante que estas amostras representem tanto a variabilidade como as tipicalidades do traço biométrico do indivíduo.

e) **Multimodal:** São baseados em múltiplos traços biométricos. Características fisicamente descorrelacionados, em geral, resultam em maiores ganhos que características correlacionadas. O projeto destes sistemas também requer novos sensores, além do desenvolvimento de uma interface adequada para integrar todos os modos que se deseja utilizar. Um grande problema deste tipo de sistema é a "maldição da dimensionalidade": a inclusão de novas modalidades biométricas aumenta a dimensão do problema, enquanto o número de amostras permanece constante, o que dificulta um treinamento adequado dos classificadores [26].

f) Híbridos: Chang et al. [60] utilizam este termo para descrever sistemas que integram um subconjunto dos cinco casos anteriores. Brunelli et al. [53], por exemplo, discutem um arranjo no qual dois algoritmos de reconhecimento de locutor são combinados com três algoritmos de reconhecimento de face.

É possível também executar a fusão de dados biométricos com dados não-biométricos para aumentar a segurança do sistema. Em [61] é discutido um autenticador que combina números pseudo-aleatórios, impressos em um cartão pessoal, com um conjunto de características faciais, produzindo um sub-conjunto destas características que é específico para cada usuário, chamado de BioCode. Quando há suspeita de que a informação biométrica de um indivíduo foi comprometida um novo cartão é emitido, revogando o autenticador anterior. O uso de autenticadores biométricos e não biométricos em conjunto é uma forma poderosa de aumentar a segurança, no entanto, algumas das inconveniências associadas aos métodos tradicionais voltam à tona.

Uma categoria à parte de sistemas multibiométricos é formada por soluções que combinam identificadores biométricos primários (como face e impressão digital) com atributos biométricos secundários, conhecidos como *soft biometrics*, tais como sexo, altura, peso ou cor dos olhos. Os atributos secundários não são suficientes para realizar a identificação de um usuário, pois o mesmo atributo é compartilhado por diversas pessoas. No entanto, quando utilizados em conjunção com os atributos primários, o desempenho do sistema de autenticação como um todo pode ser melhorado significativamente [62].

Atributos secundários também podem ser utilizados para filtar grandes bases biométricas, limitando o número de comparações necessárias em um sistema de identificação. Se é sábido que o sujeito a ser identificado é um “Homem Latino”, por exemplo, o sistema pode restringir sua busca apenas às entradas do banco de dados que possuam estes rótulos. De forma semelhante, traços biométricos secundários podem ser utilizados em aplicações de vigilância, para decidir se é necessário coletar informação biométrica primária de um certo indivíduo. O desenvolvimento de técnicas automatizadas para a estimação de características biométricas secundárias é um dos ramos de pesquisa na biometria.

2.6.1 Níveis de Fusão

Em sistemas de reconhecimento de padrões, como os biométricos, a quantidade de informação disponível para a tomada de decisões é reduzida conforme prosseguimos ao longo dos vários módulos do sistema. Com base no tipo de informação disponível em um certo módulo, diferentes níveis de fusão podem ser definidos. Sanderson e Paliwal [63] dividem os diversos níveis de fusão em duas grandes categorias: pré-classificação e pós-classificação. Isto se dá porque a quantidade de informação é reduzida drasticamente após a aplicação do classificador. Esquemas de fusão pré-classificação geralmente necessitam que novas técnicas de classificação sejam desenvolvidas, já que os classificadores utilizados para as fontes tratadas individualmente podem não mais ser adequados ao problema, o que introduz desafios adicionais a esta tipo de fusão. A categoria de pré-classificação inclui técnicas de fusão no sensor e nos dados, enquanto a pós-classificação engloba a fusão de pontuações, colocação e decisão.

Para classificadores binários, que é o caso dos sistemas biométricos em um esquema de verificação, é fácil mostrar que o último nível de fusão, o das decisões, é particularmente pouco útil. Sejam dois sistemas biométricos hipotéticos S_1 e S_2, independentes entre si. Sejam $FAR_1(\tau_1)$, $FRR_1(\tau_1)$, $FAR_2(\tau_2)$ e $FRR_2(\tau_2)$ as taxas de falsa aceitação e rejeição para os sistemas S_1 e S_2, respectivamente. Analisaremos a seguir duas das formas de combinação possíveis para um classificador binário: união e interseção.

a) União: Neste caso o usuário é dito genuíno caso passe em qualquer um dos dois testes. Portanto, a única forma do usuário ser rejeitado é falhando nos dois testes simultaneamente, daí temos:

$$FRR_{\cup}(\tau_1, \tau_2) = FRR_1(\tau_1) FRR_2(\tau_2) \tag{2.4}$$

e a falsa aceitação pode ser dada pelo complemento da probabilidade de que nenhum dos testes causem falsas aceitações:

$$\begin{aligned} FAR_{\cup}(\tau_1, \tau_2) &= 1 - [1 - FAR_1(\tau_1)][1 - FAR_2(\tau_2)] \\ &= FAR_1(\tau_1) + FAR_2(\tau_2) - FAR_1(\tau_1) FAR_2(\tau_2) \end{aligned} \tag{2.5}$$

b) Interseção: O usuário será autenticado somente se tiver sucesso nos dois testes, simultaneamente. Aplicando raciocínio análogo ao caso anterior obtemos:

$$FAR_{\cap}(\tau_1, \tau_2) = FAR_1(\tau_1)FAR_2(\tau_2) \tag{2.6}$$

$$FRR_{\cap}(\tau_1, \tau_2) = FRR_1(\tau_1) + FRR_2(\tau_2) - FRR_1(\tau_1)FRR_2(\tau_2) \tag{2.7}$$

Em ambos os casos, um ganho em um tipo de erro é acompanhando de uma perda no outro. Caso ambos os sistemas operem com τ_i referente ao ponto de EER, o erro total do sistema resultante (i.e., $FAR + FRR$) será maior ou igual ao erro total do melhor sistema e menor ou igual ao erro total do pior sistema, com igualdade apenas se os dois sistemas possuírem o mesmo desempenho. No entanto, estas não são as únicas possibilidades de fusão de decisões, uma vasta teoria acerca deste assunto está disponível na literatura.

Vale notar também que a utilização das regras de união ou interseção geram sistemas mais liberais (baixa rejeição) ou mais restritos (baixa aceitação), respectivamente. Dependendo do escopo da aplicação, tais soluções de compromisso entre os dois tipos de erro podem ser toleradas.

2.7 Reconhecimento de Assinaturas

A forma como uma pessoa assina seu nome é sabidamente uma característica marcante do indivíduo [64] [65]. A assinatura ocupa um lugar muito especial dentre as modalidades biométricas [66] [67] [68] [69], principalmente por ser a forma de identificação mais difundida ao longo do tempo. Este método é comumente utilizado por instituições governamentais, legais e financeiras [70] [30] para autenticação. Além disso, a verificação de assinaturas não requer medidas invasivas, apenas o contato com o instrumento de escrita. A soma destes fatores faz com que as assinaturas tenham uma boa aceitação por parte dos usuários [71].

No entanto, a assinatura é o resultado de um processo psico-físico complexo, que depende tanto do estado psicológico do assinante quanto das condições sob a qual o processo ocorre. A assinatura de algumas pessoas pode variar significativamente, mesmo em coletas sucessivas. Portanto, apesar de algumas teorias terem sido propostas para modelar o mecanismo da escrita [72] [73] [74] [75] e o processo de deposição da tinta sobre o papel [76] [77] [78] [79], a verificação de assinaturas ainda é um desafio aberto.

Além de possuírem uma variância intra-classe elevada, falsificadores habilidosos são capazes de reproduzir assinaturas com precisão suficiente para enganar os sistemas de verificação [38]. Não bastassem tais desafios, a verificação é realizada, geralmente, com base em poucas amostras de referência [80], o que torna ainda mais difícil uma modelagem correta do processo subjacente.

A verificação de assinaturas envolve aspectos de disciplinas desde a anatomia até a engenharia, passando pela neurociência e computação [81]. Por este motivo, estudos acerca deste tema atraem pesquisadores das mais variadas áreas de atuação, trabalhando tanto para universidades quanto empresas, seja por interesse nos desafios científicos ou nas valiosas aplicações que o campo oferece [82]. O crescimento da Internet e do comércio eletrônico, em conjunto com a proliferação de PDAs e *tablet pcs*, tem feito com que a verificação automática de assinaturas seja vista com novo interesse, pois é a escolha mais óbvia em tais dispositivos. A criação de leis específicas, que já foram aprovadas em diversos países [83], [30], e a atenção dispensada por diversas associações e institutos à padronização de formatos de troca de dados referentes às assinaturas [84], [85], [86], são evidências disto. O objetivo destes esforços é facilitar a integração das tecnologias de verificação nos mais diversos equipamentos, com o intuito de gerar soluções completas para uma ampla gama de aplicações comerciais, como bancos, corretoras de seguros, clínicas e hospitais, segurança, comércio eletrônico e cartórios.

O reconhecimento de assinaturas pode ser dividido em duas modalidades distintas, estático e dinâmico. O caso estático refere-se às assinaturas "tradicionais", feitas com algum instrumento de escrita em uma superfície (e.g., cheques, recibos, documentos, pinturas), enquanto o caso dinâmico utiliza dispositivos eletrônicos adequados, como PDAs ou mesas digitalizadoras, onde os movimentos do instrumento de escrita são coletados na forma de uma série temporal.

Uma pessoa habilidosa, com acesso a uma assinatura original e um pouco de treino, é capaz de copiar uma assinatura com detalhes suficientes para enganar até mesmo especialistas no assunto. No entanto, imitar as velocidades e acelerações na execução dos traços, pressão exercida na escrita ou a forma de segurar a caneta, são tarefas bem mais difíceis. Estas novas informações fazem que os sistemas baseados em assinaturas dinâmicas possuam uma capacidade de discriminação maior que aqueles baseados em assinaturas estáticas, desde que tais dados sejam utilizados de forma correta, naturalmente. Além das informações dinâmicas, assinaturas *on-line*, como também são chamadas, possuem a vantagem de necessitar de um menor espaço de

armazenamento que assinaturas estáticas, pois apenas uma vetor de pontos é guardado, ao invés da imagem completa.

Outro ponto importante é que as assinaturas tradicionais, da forma como são utilizadas atualmente, não fornecem nenhuma segurança pró-ativa. Em instituições bancárias, por exemplo, assinaturas são conferidas apenas quando há alguma queixa ou quando transações suspeitas são detectadas (e.g., quantias excepcionalmente altas), pois a verificação de todos os casos tonaria o sistema financeiro extremamente lento. Este é um outro tipo de problema que pode ser sanado, ou ao menos reduzido, através da utilização de assinaturas dinâmicas, pois a verificação pode ser realizada no momento da escrita, ao invés de ser coletada em um meio físico (e.g., papel) para ser posteriormente digitalizada e processada.

Até agora consideramos apenas a aplicação que objetivamos neste trabalho: a biometria, ou, mais especificamente, a verificação de assinaturas. No capítulo seguinte, apresentamos a teoria da análise multifractal, que será a ferramenta utilizada na análise dos sinais de assinaturas dinâmicas com os quais trabalharemos.

Capítulo 3

Fractais

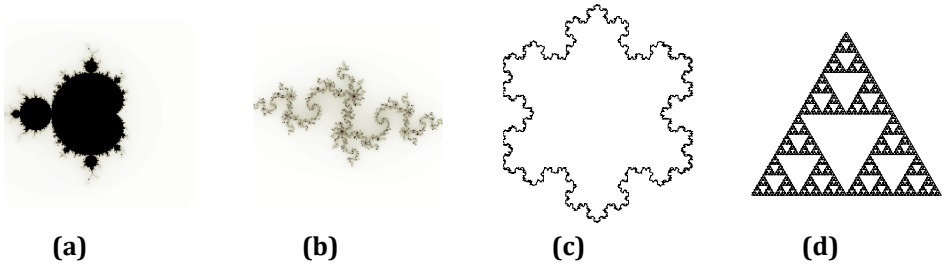
3.1 Introdução

Com origem no latim (*fractus* – irregular, quebrado), a palavra fractal foi cunhada por Benoît Mandelbrot [5] para descrever entidades geométricas caracterizadas por irregularidades que governam sua forma e complexidade, possuindo uma estrutura com detalhes em todos os níveis de resolução [5] [87] [88], ou, em suas próprias palavras, "um fractal é uma forma geométrica irregular ou fragmentada que pode ser quebrada em partes, cada uma das quais é (ao menos aproximadamente) uma cópia em tamanho reduzido do todo".

A definição formal de fractal é não-trivial e depende, primeiramente, da definição de dimensão de conjunto adotada. Todo conjunto matemático possui um número associado, chamado de dimesão topológica, que pode ser entendido como o número mínimo de parâmetros necessários para descrever todos os pontos a ele pertencentes. A geometria euclidiana, por exemplo, opera em conjuntos de dimensões topológicas inteiras (e.g., $\mathbb{R}^n$), onde são necessários n parâmetros para descrever seus elementos (e.g., x e y para um ponto bi-dimensional). As propriedades fundamentais que um fractal deve apresentar são: auto-similaridade, espectro segundo uma lei de potência e dimensão de Hausdorff maior que sua dimensão topológica.

Os conjuntos de Mandelbrot e Julia e as curvas de Koch e Sierpinski, ilustrados na Figura 3.1, são exemplos típicos e bem conhecidos de fractais, em geral contruídos a partir de funções iterativas [89] [5]. A curva de Von Koch , também conhecida como floco de neve de Von Koch, por exemplo, possui um comprimento infinito mas está limitada a uma região finita do espaço, este fenômeno aparentemente paradoxal é resultado de sua irregularidade. A geometria euclidiana é, portanto, mal adaptada para caracterizar conjuntos com tais comportamentos [87].

Além destas abstrações matemáticas citadas no parágrafo anterior, diversos estudos apontam a presença de comportamento aproximadamente fractal em fenômenos naturais, como os litorais dos continentes [90], séries temporais referentes aos batimentos cardíacos [9] e a distribuição de veias e artérias no corpo humano [10].



(a) (b) (c) (d)

Figura 3.1: Fractais
(a) conjunto de Mandelbrot; (b) conjunto de Julia; (c) floco de neve de Koch; (d) triângulo de Sierpinski

A dimensão de Hausdorff, definida em 1919, baseia-se no tamanho das variações dos conjuntos quando medidos ao longo de diferentes escalas, e foi utilizada por Mandelbrot no desenvolvimento da teoria fractal. Esta dimensão mede o grau das irregularidades e possui, geralmente, valores fracionários, sendo este um dos principais motivos para o nome "fractal". Além da dimensão de Hausdorff, a dimensão de capacidade, a dimensão de correlação e a dimensão de informação também são comumente utilizadas no estudo dos fractais.

3.1.1 Dimensão de Hausdorff e Dimensão de Capacidade

Seja S um subconjunto não-vazio do espaço $\mathbb{R}^n$. O diâmetro de S é dado por

$$|S| = sup\{|x - y|: x, y \in S\} \tag{3.1}$$

Seja $\{S_i\}$ uma coleção finita de conjuntos com diâmetros de valor máximo d, que cobre um subconjunto T em $\mathbb{R}^n$, tal que $T \subset \bigcup_{i=1}^{\infty} S_i$. Neste caso, chamamos $\{S_i\}$ de coleção de conjuntos de cobertura d de T, a Figura 3.2 ajuda a visualizar a definição.

De posse destes conceitos, podemos definir a medida de Hausdorff:

Definição 3.1.1 *Define-se* $M_d^k(T)$ *para algum* $d > 0$ *como*

$$M_d^k(T) = \inf\left\{\sum_{i=1}^{\infty} |S_i|^k : \{S_i\} \text{ é coleção de conjuntos de cobertura } d \text{ de } T\right\} \tag{3.2}$$

A medida de Hausdorff k-dimensional é dada por:

$$M^k(T) = \lim_{d \to 0} M_d^k(T) \tag{3.3}$$

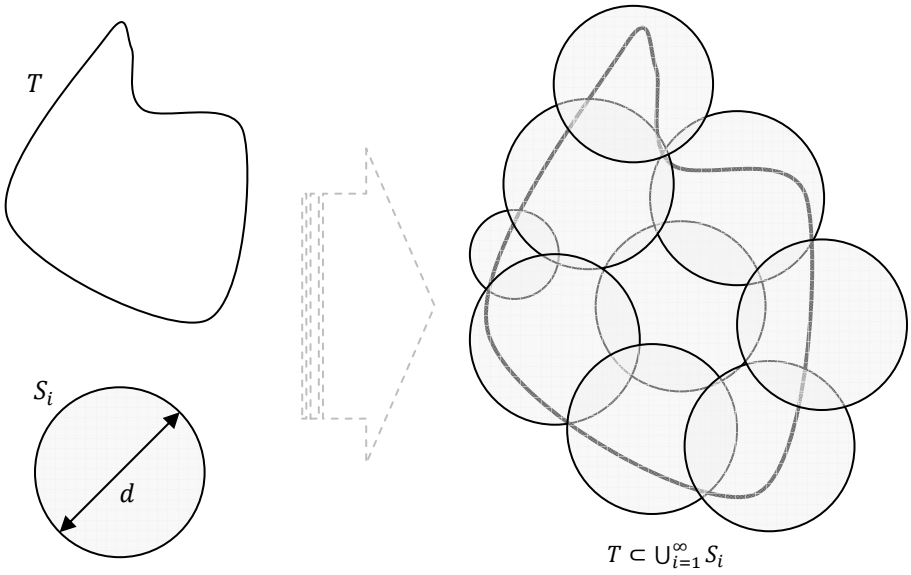


Figura 3.2: Coleção de conjuntos de cobertura

A partir da definição da medida, que existe para qualquer subconjunto T do $\mathbb{R}^n$ [5], define-se a dimensão de Hausdorff:

> ***Definição 3.1.2*** *O valor de k para o qual a medida* $M_d^k(T)$ *muda instantaneamente de* ∞ *para* 0 *é definido como dimensão de Hausdorff de* T*, e é denotada por* $D_H(T)$.

Estas definições aparentemente complicadas podem ser entendidas de forma bastante intuitiva. Suponha que você deseja cobrir uma mesa circular com um mosaico de ladrilhos quadrados, cada ladrilho possui lado de comprimento $l$cm. O número N de ladrilhos necessários para cobrir a parede é dado pela área da mesa dividido pela área do ladrilho (i.e., A/l^2). No entanto, nem sempre será possível cobrir completamente a mesa: provavelmente sobrará uma brecha onde o ladrilho não caberá. Ora, vamos então escolher ladrilhos menores e tentar novamente. Se dermos continuidade a este processo, em algum momento será possível preencher toda a mesa.

Matematicamente, podemos escrever este procedimento como: $A = \lim_{l \to 0} N \times l^2$. Ou, de um modo mais geral, $A = \lim_{l \to 0} N \times l^D$, onde D é a dimensão do conjunto que está sendo analisado, para os casos onde $D = 1$ e $D = 2$, ilustramos o conceito na Figura 3.3. Considerando que o objeto que desejamos cobrir cabe em um hipercubo de lado 1 (i.e., $A = 1$), com um pouco de manipulação algébrica podemos obter o valor de D através da equação:

$$D = \lim_{l \to 0} \frac{\log(N)}{\log\left(\frac{1}{l}\right)} \tag{3.4}$$

Esta é outra forma de definir a dimensão de Hausdorff, que é idêntica à definição da dimensão de Capacidade [5] [91], introduzida por Kolmogorov. A idéia de cobertura por elementos infinitesimais é a mesma empregada no cálculo integral.

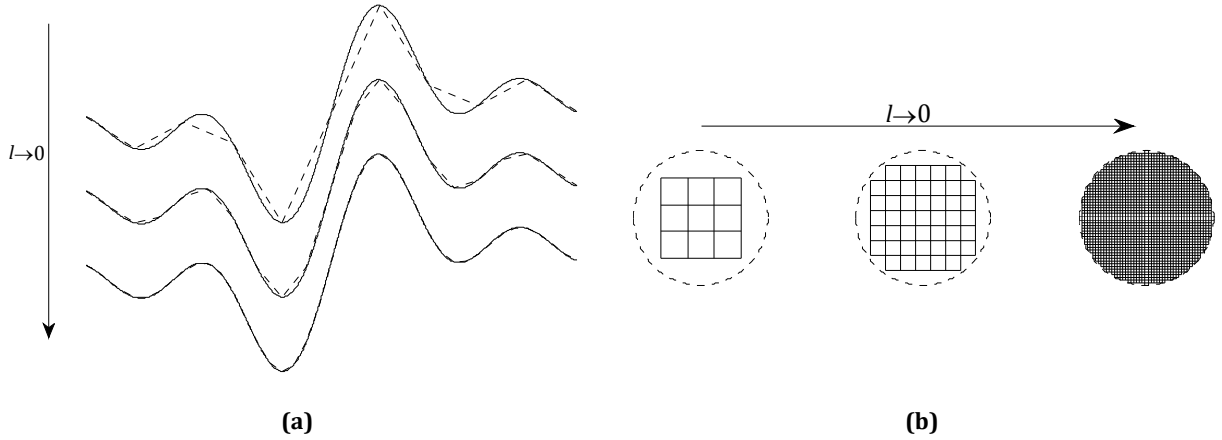


Figura 3.3: Aproximação por elementos infinitesimais
(a) Curva aproximada por segmentos de reta de comprimento l
(b) Círculo aproximado por quadrados de lado l

3.1.2 Dimensão de Correlação

Definição 3.1.3 *Seja um conjunto S de cardinalidade N, $S = \{X_1, X_2, \dots, X_N\}$. Seja $d(i,j) = |X_i - X_j|$. A dimensão de correlação é dada por:*

$$D_R = \lim_{\tau \to 0} \frac{\log \rho(\tau)}{\log \tau} \tag{3.5}$$

onde

$$\rho(\tau) = \lim_{N \to \infty} \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \delta(\tau - d(i,j)) \tag{3.6}$$

e

$$\delta(x) = \begin{cases} 1, \text{se } x \geq 0 \\ 0, \text{se } x < 0 \end{cases} \tag{3.7}$$

Note que $\rho(\tau)$ também pode ser escrita como uma lei de potência $\rho(\tau) \approx c\tau^{D}$, onde c é uma constante, e obtém-se um resultado semelhante àquele alcançado para D_H e D_C [5] [91].

Esta correlação em forma de uma lei de potência é bastante conhecida na termodinâmica e na mecânica estatística, sendo responsável pelo fenômeno de invariância em escala nos processos de mudança de fase.

3.1.3 Dimensão de Informação

Esta dimensão é definida no contexto de sistemas dinâmicos e Teoria do Caos [91]. Com o intuito de explanar seu significado, apresentamos no parágrafo seguinte, de forma bastante condensada alguns elementos básicos acerca deste tópico. Ao leitor que não está familiarizado com o assunto, ou deseja uma explanação mais detalhada, aconselhamos a leitura de materiais específicos sobre o tema, como [92].

Um sistema dinâmico pode ser modelado por um conjunto de entradas, saídas e variáveis de estados relacionados através de equações diferenciais. O espaço de estados, S, é um espaço cujos eixos são as variáveis de estado, e cada ponto deste espaço representa um estado particular do sistema dinâmico em questão. Um atrator é um sub-conjunto do espaço de estados em direção ao qual o sistema evolui ao longo do tempo. Para determinar a dimensão de informação, divide-se a região do atrator em células, c, de lado l, e define-se:

Definição 3.1.4 *Seja p_i a probabilidade de uma célula c_i de lado l possuir algum ponto em seu interior. A dimensão de informação é dada por:*

$$D_I = \lim_{l \to 0} -\frac{I(l)}{\log l} = \lim_{l \to 0} \frac{\sum_{(i=1)}^{N} p_i \log p_i}{\log l} \tag{3.8}$$

Atratores que possuem uma dimensão de informação diferente da dimensão topológica são conhecidos como Atratores Estranhos, e são objeto de estudo da teoria do caos.

Quando estas quatro dimensões aqui definidas coincidem, o objeto em questão é dito *monofractal*, quando elas são distintas, diz-se que ele é *multifractal* [87] [91].

3.2 Monofractais

Muito antes do termo fractal ser proposto, Kolmogorov, no início dos anos 40, introduziu o conceito de auto-similaridade para designar processos escalonáveis sem alterações em suas propriedades estatísticas.

Um ponto marcante dos processos auto-similares é a dependência de longo prazo, que faz com que a sua função de autocorrelação possua uma distribuição de cauda longa (*heavy-tailed*), com decaimento hiperbólico [93]. Mais precisamente, um processo X é dito de cauda longa se $P(X > x) \approx x^{-\alpha},\ x \to \infty$, onde $0 < \alpha < 2$ é chamado de índice de cauda. Estes processos possuem variância infinita e, além disso, para $\alpha < 1$, sua média também é infinita. Apesar de esta ser uma propriedade marcante dos processos auto-similares, a dependência de longo prazo não é suficiente para definir um processo como tal [94]. Apresentamos a seguir as definições matemáticas de dependência de longo prazo e auto-similaridade.

Definição 3.2.1 (dependência de longo prazo) *Um processo estacionário $X(t)$, com média zero e variância finita, apresenta dependência de longo prazo (ou memória longa) se sua função de autocorrelação $\rho(\tau) = E\{X(t+\tau)X(t)\}$ ou sua densidade espectral de potência $S_X(f)$ (transformada de Fourier de $\rho(\tau)$) satisfazem*

$$\rho(\tau) \approx c\tau^{-\beta}, \qquad \tau \to \infty \tag{3.9}$$

$$S_X(f) \approx k|f|^{-\gamma}, \qquad f \to \infty \tag{3.10}$$

onde $0 < \beta < 1$, $0 < \gamma < 1$, $\gamma = 1 - \beta$ e c e k são constantes não-negativas.

Apesar de dificultar a estimação de parâmetros estatísticos, a presença de memória de longo prazo, quando explorada adequadamente, pode auxiliar na predição de valores futuros do processo [94].

Definição 3.2.2 (auto-similaridade) *Seja $X_t = \{X(t), t \in \mathbb{R}^+\}$ um processo no tempo contínuo. X_t é dito auto-similar com parâmetro $H \in [0,1)$ se*

$$X(t) \sim c^{-H}X(ct), \qquad (\forall t \in T, \forall c > 0) \tag{3.11}$$

onde $\sim$ representa igualdade em distribuição.

Estes dois conceitos estão intimamente ligados, pois $H = 1 - \gamma/2$. O coeficiente H, que determina o grau de auto-similaridade do processo, é chamado de parâmetro de Hurst, em referência o hidrólogo Harold Hurst que, em 1951, publicou um artigo acerca da estrutura de correlação e dependência de longo prazo dos níveis das águas do rio Nilo [5] [94] [95].

Os processos *monofractais* são, em geral, resultantes de um processo aditivo, e governados pelo conteúdo de baixa frequência, sendo suficientemente caracterizados pelo seu parâmetro de Hurst [94] [96] [97] [98]. Pode-se ainda classificar o processo monofractal segundo o valor do seu parâmetro H:

a) $\mathbf{0,5 < H < 1}$: Neste caso, a integral da função de auto-correlação tende ao infinito, o que evidencia a longa memória do processo;
b) $\boldsymbol{H = 0,5}$**:** O processo é descorrelacionado. Um exemplo clássico de processo monofractal deste tipo é o Movimento Browniano;
c) $\mathbf{0 < H < 0,5}$**:** Neste caso a integral da função de auto-correlação é nula. O que indica dependência de curto prazo ou anti-persistente [94].

De um modo mais intuitivo, podemos dizer que o parâmetro de Hurst avalia a regularidade do função[2] analisada. Valores elevados indicam processos suaves, com poucas variações, valores baixos representam processos mais irregulares. Na Figura 3.4, apresentamos realizações de processos *monofractais* com diferentes valores do parâmetro H.

O parâmetro de Hurst está também intimamente relacionado à dimensão fractal, já que se pode mostrar que $D_H = 2 - H$, de modo que, quanto menor o parâmetro H, maior é a dimensão fractal, o que caracteriza sinais mais irregulares, corroborando a análise feita anteriormente.

Outra forma de classificar estes processos é segundo sua variância, que pode ser finita ou infinita. Entre os de variância finita temos os gaussianos, exponenciais e lognormais; no tocante aos de variância infinita, sua representação mais adequada é através dos processos denominados α-estáveis [99].

Diversos métodos de estimação do parâmetro H estão disponíveis na literatura [97] [99]. Além disso, vários processos monofractais foram amplamente estudados, sendo os mais comuns [94] o movimento Browniano fracionário (fBm – *fractional Brownian motion*), o ruído gaussiano

[2] Ao longo deste texto os termos "processo", "função" e "sinal" serão utilizados de forma indistinta.

fracionário (fGn – *fractional Gaussian noise*), MMPP (*Markov Modulated Poisson Process*) e F-ARIMA (*Fractional Autoregressive Integrated Moving Average*). Estes dois últimos podem convergir para o modelo fBm, conforme mostrado em [100]. Uma lista de sugestões de leitura acerca destes tópicos pode ser encontrada ao fim deste capítulo.

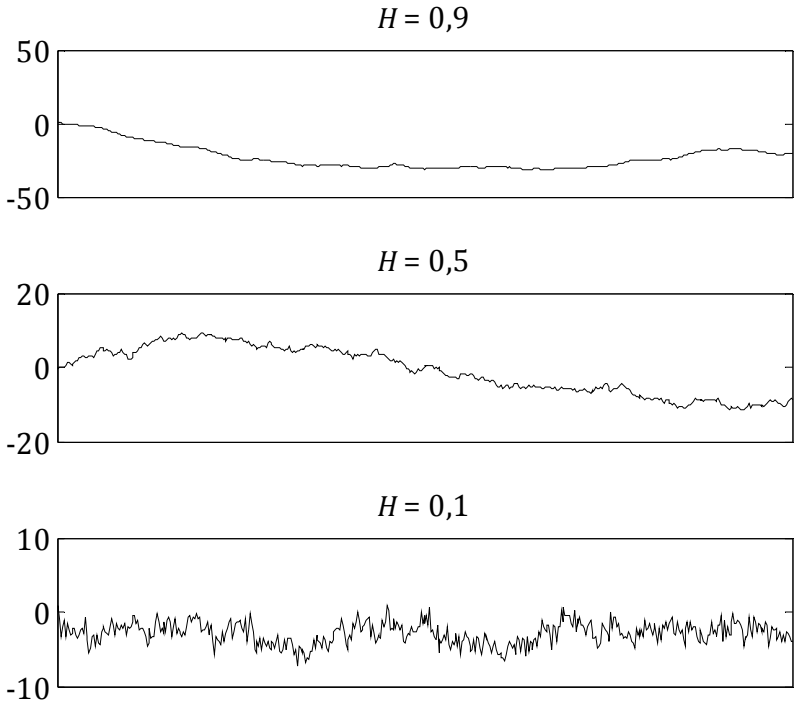


Figura 3.4: Processos monofractais com diferentes valores do parâmetro de Hurst

3.3 Multifractais

Do outro lado do espectro, encontram-se os processos multifractais, governados por conteúdo de alta frequência do sinal [101] [102]. Este fenômeno trata da estrutura de singularidades do sinal (i.e., pontos onde o mesmo torna-se degenerado), sendo geralmente formado a partir de processos multiplicativos. Esta diferença dá origem a vários comportamentos distintos em escalas diferentes (*multiple scaling*), de modo que o parâmetro de Hurst, que é uma constante, não é mais adequado à sua caracterização [103].

A dimensão, ou mesmo o parâmetro de Hurst, de um processo multifractal ainda podem ser utilizados para se obter uma noção global da complexidade do mesmo [88]. Contudo, o

comportamento local varia com o tempo, e, por isso, processos multifractais são definidos por leis de escala e momentos estatísticos em intervalos de tempo finitos.

Definição 3.3.1 *Um processo estocástico $X(t)$ é multifractal se satisfaz a equação*

$$E\{|X(t)|^q\} = c(q)t^{\tau(q)+1} \tag{3.12}$$

onde $t \in T$ e $q \in Q$. T e Q são intervalos na reta real. Normalmente assume-se que T e Q têm comprimentos positivos e que $0 \in T$ e $[0,1] \subseteq Q$.

Esta definição descreve o comportamento multifractal em termos de momentos estatísticos, onde $\tau(q)$ é chamado de função de escala, ou de partição, e $c(q)$ é denominado fator de momento do processo. Na realidade, esta definição pode ser vista como uma generalização do caso monofractal, uma vez que, se $\tau(q)$ é linear, voltamos ao caso monofractal. Com efeito, para o caso monofractal, pode-se mostrar que $\tau(q) = qH - 1$ e $c(q) = E\{|X(1)|^q\}$.

Sob uma ótica mais prática, os pontos singulares dos processos carregam em si uma grande quantidade de informação. Tanto em sinais referentes a eletrocardiogramas, fala, ou tráfego em redes de comunicação, a regularidade local é de grande utilidade para a análise [104] [105]. Portanto, precisamos buscar ferramentas para quantificar a regularidade de um sinal de modo localizado. Com este intuito, apresentamos a seguir uma série de definições que convergem para a medida de regularidade local desejada.

Definição 3.3.2 *Seja $\Omega \subset \mathbb{R}^n, k > 0$ e a função $f: \Omega \to \mathbb{R}$. $C^k(\Omega)$ define o conjunto de funções f que são k vezes diferenciáveis, com derivadas contínuas.*

Definição 3.3.3 (espaço métrico) *Um espaço métrico é um par ordenado (X, d) em que X é um conjunto não-vazio com uma função $d: X \times X \to [0, \infty)$, satisfazendo:*

1. *(não-degeneração) $d(x, y) = 0 \Leftrightarrow x = y$*
2. *(simetria) $d(x, y) = d(y, x), \forall\, x, y \in X$*
3. *(desigualdade triangular) $d(x, z) \leq d(x, y) + d(y, z), \forall x, y, z \in X$*

A função d é dita uma métrica em X e o número $d(x, y)$ é chamado de distância de x a y.

Definição 3.3.4 *Sejam* (X, d) *um espaço métrico,* r *um número positivo e* $x, x_0 \in X$*. O conjunto*

$$B(x_o, r) = \{x: d(x_o, x) < r\} \tag{3.13}$$

é denominado de bola aberta com centro em x_0 *e raio* r*.*

Definição 3.3.5 (condição de Hölder) *Seja* α *um número real positivo,* $x_0 \in \mathbb{R}$*, e uma função* $f: \mathbb{R} \to \mathbb{R}$*. Caso* $m < \alpha < m + 1, m \in \mathbb{N}$*, podemos dizer que* $f \in C^{\alpha}\big(B(x_0, r)\big)$ *caso exista uma constante* k *tal que, para todo* x, y *em* $B(x_0, r)$*,*

$$|\partial^m f(x) - \partial^m f(y)| \leq k|x - y|^{\alpha - m} \tag{3.14}$$

O conjunto $C^{\alpha}\big(B(x_0, r)\big)$ é o conjunto das funções contínuas segundo Hölder, também chamadas em alguns trabalhos de contínuas segundo Lipschitz, no entanto, a definição de Lipschitz não inclui o expoente α, sendo um caso particular da definição de Hölder.

Uma forma de interpretar esta condição é que existe uma área em torno do ponto y de modo que todos os pontos de f na bola $B(x_0, r)$ pertencem a esta região, exceto y, como ilustrado na Figura 3.5. Note que, para valores de α menores que 1, os vizinhos imediatos de $f(y)$ podem sofrer variações maiores e ainda assim satisfazer a condição.

A partir da definição de continuidade de Hölder são derivadas as ferramentas de quantificação da regularidade local de uma função [5], adequadamente chamadas de expoentes de Hölder, em homenagem ao matemático alemão Ludwig Hölder, autor da Definição 3.3.5.

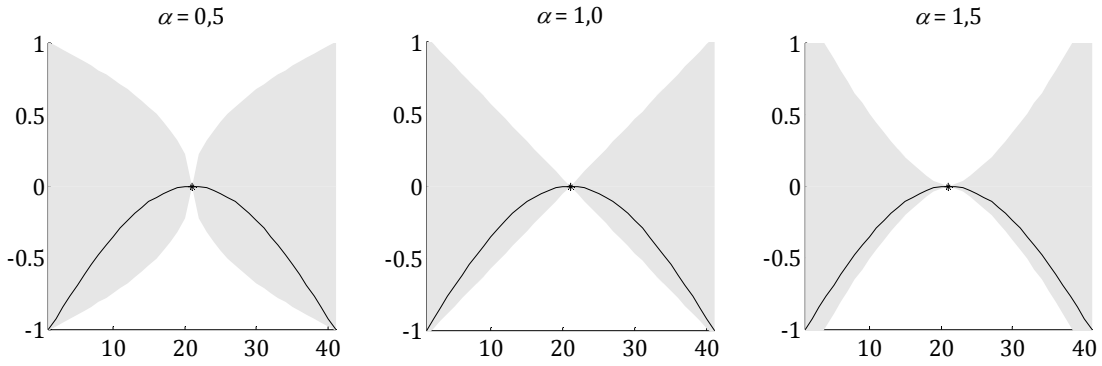


Figura 3.5: Condição de Hölder para diferentes valores de α

Definição 3.3.6 (expoente local de Hölder) *Seja f uma função contínua segundo Hölder e,*

$$\alpha_L\big(x_0, B(x_0, r)\big) = \sup\{\alpha: f \in C^{\alpha}\big(B(x_0, r)\big)\} \tag{3.15}$$

O expoente de Hölder local, α_L de f em x_0 é definido como

$$\alpha_L(x_0) = \lim_{r \to 0} \alpha_L\big(x_0, B(x_0, r)\big) \tag{3.16}$$

Definição 3.3.7 (expoente pontual de Hölder) *Seja α um número real, k uma constante e $x_0 \in \mathbb{R}$. A função $f: \mathbb{R} \to \mathbb{R} \in C^{\alpha}(x_0)$ se existe um polinômio P_n de grau $n < \alpha$, tal que*

$$|f(x) - P_n(x - x_0)\,| \leq k|x - x_0|^{\alpha} \tag{3.17}$$

O expoente de Hölder pontual α_P da função f em x_0 é definido como

$$\alpha_P(x_0) = \sup\{\alpha | f \in C^{\alpha}\,(x_0)\} \tag{3.18}$$

Estes expoentes são as medidas de regularidade mais utilizadas no campo da análise fractal. É fácil ver, seguindo a seqüência de definições, que uma operação de integração aumentará o expoente de Hölder de uma unidade, enquanto uma operação de diferenciação reduzir-lo-á de um, o que faz sentido também de um ponto de vista intuitivo, já que operações de integração causam uma suavização do sinal.

Outro modo de entender esta variação através das operações de integração e derivação é que o expoente de Hölder é uma medida de diferenciabilidade da função que, no caso de fractais, não é um número inteiro. Portanto, como o leitor já deve ter percebido, quanto maior o expoente de Hölder mais suave (regular) será f naquele ponto, para o caso α_P, ou região, no caso α_L.

Funções com $\alpha \geq 1$ são continuas e diferenciáveis $\alpha - 1$ vezes; para $0 < \alpha < 1$ são contínuas segundo Hölder e não diferenciáveis; para $-1 < \alpha \leq 0$ temos funções descontínuas e não-diferenciáveis e, finalmente, para $\alpha = -1$ a função não é mais localmente integrável.

A escolha de uso do expoente local ou pontual depende da aplicação em questão, o expoente local é estável sob a ação de operadores diferenciais ou integradores, mas possui a desvantagem de não ser capaz de observar um ponto regular em um ambiente irregular [106],

apenas as singularidades dominantes são observadas. De um modo geral, os expoentes pontuais são mais versáteis.

Como vimos na Definição 3.3.1, um processo multifractal é descrito pelas funções de partição e pelo fator de momento. Considere os dados X_i com suporte no intervalo $[0, T]$, em uma escala q_i. Define-se a soma-partição como

$$S_m^X(q_i) = \sum_{k=1}^{N/m} \left(\bar{X}_k^{(m)}\right)^{q_i} \tag{3.19}$$

onde

$$\bar{X}_k^{(m)} = \sum_{j=1}^{m} X_{(k-1)m+j} \tag{3.20}$$

Para q_i fixo, variam-se os valores de m, obtendo-se um conjunto de pontos que são analisados no plano $\log m \times \log S_m^X(q_i)$. Estima-se a função de partição $\tau(q_i)$ e o logaritmo do fator de monento $c(q_i)$ através do coeficiente angular e linear da reta obtida no passo anterior. Portanto, temos que

$$S_m^X(q) \cong \tau(q_i)\, log\, m + log\, c(q_i). \tag{3.21}$$

Repetindo-se o processo para diversos valores de q_i, obtêm-se as funções $\tau(q)$ e $c(q)$.

Outra função importante na análise multi-fractal é a função de partição baseada na análise *wavelets* [107] [108]. Dada uma função $f(t)$ em uma escala q, a função de partição *wavelet* é definida como

$$S(q, a) = \sum_t \left|T_\psi(t, a)\right|^q \tag{3.22}$$

onde

$$T_\psi(t, a) = \int_{-\infty}^{\infty} f(t)\, \psi_{t,a}(\gamma) d\gamma \tag{3.23}$$

onde a *wavelet* de análise ψ é uma função de suporte limitado, centrada em zero, cuja família de vetores *wavelet* é obtida por transalação t e dilação a [109].

Dado que $n_\psi > \alpha(t_0)$, onde n_ψ é o número de momentos evanescentes (*vanishing moments*) e $\alpha(t_0)$ é o expoente de Hölder no ponto t_0, a transformada *wavelet* em t_0, $T_\psi(t_0, a)$, apresenta o seguinte comportamento:

$$T_\psi(t_0, a) \approx a^{\alpha(t_0)}, \qquad (a \to 0^+) \tag{3.24}$$

Portanto, pode-se extrair o expoente pontual $\alpha(t_0)$, como a inclinação em um gráfico log-log da amplitude da transformada em função do fator de escala, a, como mostrado na Figura 3.7. Caso $n_\psi < \alpha(x_0)$, o comportamento ainda segue uma lei de potência mas, neste caso, o expoente será n_ψ.

Mallat e Hwang [110] notaram que o expoente local pode ser igualmente estimado analisando-se apenas os valores do módulo da transformada ao longo de uma linha de máximos que converge para o ponto t_0, o que reduz significativamente o custo computacional da estimação. Esta variante da transformada *wavelet* é conhecida como máximos em módulo da transformada wavelet (WTMM – *Wavelet Transform Modulus Maxima*). Em cada escala a, os máximos em módulo são calculados de modo que

$$\frac{\partial |T_\psi(t, a)|}{\partial t} = 0 \tag{3.25}$$

e as linhas de máximo são curvas conectadas no espaço-escala formada por pontos que são máximos em módulo e estão imersos num cone invertido definido por

$$|2^r k - t_0| \leq k 2^r \tag{3.26}$$

onde k é uma constante e r é o raio, ilustramos este procedimento na Figura 3.6. Dentro deste cone, a linha de máximos converge para a singularidade presente no ponto t_0 [111].

Após identificar a linha de máximos para o ponto desejado, é contruído um gráfico de $\log|T_\psi(t_0, a)| \times \log a$ e realizada uma aproximação linear, conforme explicado anteriormente. O coeficiente angular desta aproximação é o expoente de Hölder no ponto t_o. Note que, como o WTMM utiliza apenas máximos locais, ele tende a manter-se e igual a singularidade dominante em uma vizinhança de t_0. Na Figura 3.7 realizamos este procedimento com e sem o uso do WTMM, para a transformada *wavelet* da Figura 3.6.

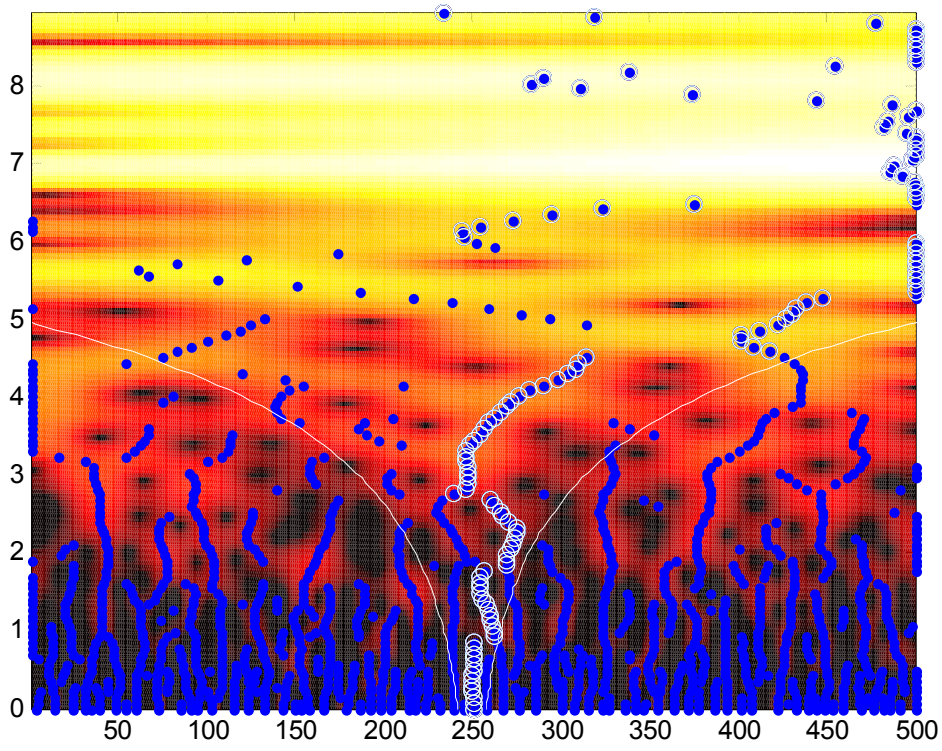
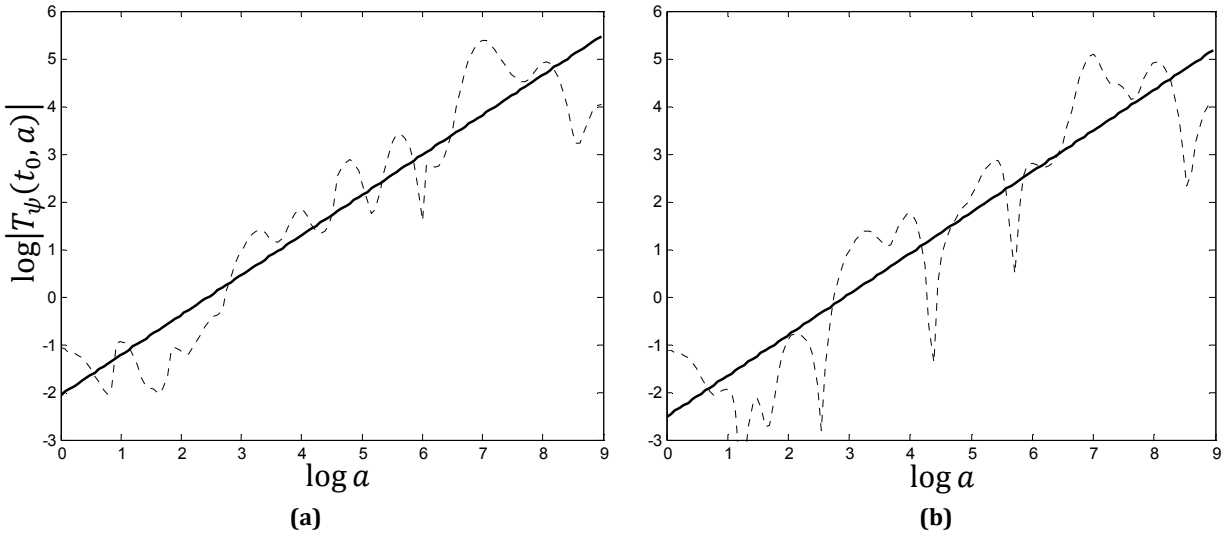


Figura 3.6: WTMM de um movimento Browniano no instante $t_0 = 250$
Pontos azuis são máximos locais, pontos circulados em branco formam a linha de máximos



(a) **(b)**

Figura 3.7: Estimação do expoente de Hölder pontual utilizando *wavelets*
(a) utilizando WTMM; (b) sem utilizar WTMM.

3.4 Espectro Multifractal

O espectro multifractal é a curva que descreve a distribuição geométrica ou estatística dos expoentes de Hölder de um determinado processo [87]. Existem três espectros multifractais de grande importância, o espectro de Hausdorff, o espectro de Grandes Desvios e o espectro de Legendre [112] [88]. Geralmente essa curva possui um perfil parabólico côncavo.

3.4.1 Espectro de Hausdorff

Dentre as dimensões fractais que descrevemos, a dimensão de Hausdorff é, matematicamente, a mais precisa delas. Além disso, ela possui a vantagem de ser definida para qualquer conjunto. Por este motivo, o espectro de Hausdorff é o mais desejado dentre os espectros multifractais. No entanto, este é também o espectro mais difícil de ser obtido [89].

Um conjunto de pontos com um mesmo grau de regularidade α constitui um conjunto fractal, cuja descrição geométrica é dada por sua dimensão de Hausdorff. Desse modo, define-se o espectro de Hausdorff como:

Definição 3.4.1 *Seja P_α o conjunto dos pontos de $X(t)$ com regularidade α. O espectro de Hausdorff de $X(t)$ é dado por*

$$f_H(\alpha) = D_H(P_\alpha) \tag{3.27}$$

3.4.2 Espectro de Grandes Desvios

Para uma realização fixa de um processo $X(t)$, sus variações em torno de t são descritas por

$$|X(t+\Delta t) - X(t)| \approx C_t(\Delta t)^{\alpha(t)} \tag{3.28}$$

onde C_t é chamado de pré-fator [5]. Percebe-se através desta relação que $\alpha(t)$ é um fator de escalonamento local em t. Portanto, o expoente de Hölder pode ser dado por

$$\alpha(t) = sup\left\{\alpha: \lim_{\Delta t \to 0} \frac{log|Z(t+\Delta t) - Z(t)|}{log\,\Delta t}\right\} \tag{3.29}$$

No entanto, na prática, o limite $\Delta t \to 0$ não pode ser obtido. Nesses casos, recorre-se ao conceito de expoentes aproximados de Hölder (*coarse Hölder exponents*).

Definição 3.4.2 *Seja um processo $X(t)$ com suporte no intervalo $[0, T]$. Subdivida iterativamente este intervalo em b^k partes iguais. O expoente aproximado de Hölder é dado por*

$$\alpha_k(t) \equiv \frac{log\left|Z\left(t + b^{-k}T\right) - Z(t)\right|}{log\, b^{-k}\, T} \tag{3.30}$$

Pode-se então estimar a probabilidade de que um ponto aleatoriamente escolhido tenha um dado expoente de Hölder. Se $N^{(k)}(\alpha)$ é o número de expoentes aproximados de Hölder iguais a α contidos em cada intervalo, segundo o princípio dos grandes desvios (LDP – *Large Deviation Principle*), quando $k \to \infty$, a razão $N^{(k)}\left(\alpha_j\right)/b^k$ converge para a probabilidade de um ponto t possuir expoente α [113]. De forma mais rigorosa temos:

Definição 3.4.3 *Seja $N^{(k)}(\alpha)$ o número de expoentes aproximados de Hölder iguais a α que ocorrem ao subdividir o processo $X(t)$ em b^k intervalos de mesmo tamanho. Então o espectro multifractal dos Grandes Desvios, ou de granularidade grosseira (coarse grain), é dado por:*

$$f_{GD}(\alpha) \equiv \lim_{k\to\infty}\left\{\frac{log\, N^{(k)}(\alpha)}{log\, b^k}\right\} \tag{3.31}$$

Mantendo a relação com o LDP, podemos dizer que o espectro $f_{GD}(\alpha)$ indica quão rápido a probabilidade de observar um α atípico decresce: ele pode ser considerado como uma função taxa sob este prisma. Caso este limite exista e seja definido para um único ponto, a função é monofractal, caso contrário, temos um processo multifractal.

Vale ainda notar que este espectro, apesar de indicar a probabilidade de α, não é a função densidade de probabilidade de α propriamente dita, e sim corresponde a uma dupla normalização logarítmica da mesma. A estimação deste espectro requer, em geral, a aplicação de ferramentas de estimação de densidade de probabilidade.

3.4.3 Espectro de Legendre

O espectro de Legendre é uma aproximação côncava do espectro de grandes desvios. Como este é o espectro mais fácil de ser computado, é também o mais utilizado na literatura. Define-se o espectro de Legendre como:

Definição 3.4.4 *Seja* $\tau(q)$ *a função de partição de um processo* $X(t)$*. O espectro de Legendre de* $X(t)$ *é dado por*

$$f_L(\alpha) \stackrel{\text{def}}{=} \tau^*(\alpha) \tag{3.32}$$

onde $\tau^*(\alpha)$ *é a transformada de Legendre de* $\tau(\alpha)$*, dada por* $\tau^*(\alpha) = inf_q\big(q\alpha - \tau(q)\big)$

Esta definição pode ser intuída a partir do comportamento em lei de potência da função de partição *wavelet*

$$S(q,a) \approx a^{\tau(q)}, \qquad (a \to 0^+) \tag{3.33}$$

e a forte analogia que liga o formalismo multifractal à termodinâmica [114] [115] [116]. As variáveis q e $\tau(q)$ são análogas à temperatura inversa e à energia livre de Gibbs na termodinâmica, sendo, por este motivo, também chamada de função de partição de Gibbs em alguns trabalhos. Ao invés de obter a energia e a entropia, que são as variáveis conjugadas a q e $\tau(q)$, a transformada de Legendre no caso multifractal resulta no expoente de Hölder α e no espectro multifractal $f_L(\alpha)$.

De acordo com Riedi [88], têm-se, com probabilidade 1, que $f_H \leq f_{GD} \leq f_L$. De fato, como f_L é uma aproximação côncava de f_G, é fácil perceber a última desigualdade.

3.5 Considerações Finais

Neste trabalho, implementados dois métodos de estimação de expoentes pontuais de Hölder: a) um método baseado em oscilações, que pode ser visto como uma variante do método de estimação dos expoentes aproximados de Hölder, proposto por Tricot [117]; e b) o método baseado em *wavelets* que descrevemos anteriormente, que pode ser executado com ou sem a

busca de linhas de máximo. Também foram implementadas a estimação de regularidade global e do espectro de Legendre a partir deste último método.

Por fim, para os leitores que desejaram obter maiores informações sobre processos multifractais, técnicas de estimação e aplicações da análise multifractal, recomendamos as seguintes publicações: [5] [87] [88] [101] [112] [114] [118] [119] [120] [121] [122] [123].

Capítulo 4

Análise Multifractal de Assinaturas

4.1 Introdução

A informação fractal já foi utilizada, com sucesso, em estudos nas mais diversas áreas, como a arte, a mecânica estatística, a análise do tráfego em redes, o processamento de sinais, a biologia e a economia. No entanto, apesar da imensa quantidade de artigos abordando aplicações biométricas, quase nenhuma atenção foi dada para as características fractais neste âmbito.

De fato, mesmo após a extensa revisão bibliográfica realizada durante este trabalho, encontramos menos de dez publicações que aplicam, de alguma forma, a informação fractal à biometria. Em [124], é proposta uma justificativa, baseada em um argumento fractal, para o processo de normalização em assinaturas online; [125] utiliza a dimensão fractal para caracterizar a íris humana; [126] aplica uma medida de dimensão fractal aos "tremores" nas bordas de assinaturas estáticas como uma forma de identificação de falsificações, tendo por base a idéia de que fatores como velocidade, pressão e confiança ao assinar alteram o processo de deposição da tinta no papel. Em [127], um modelo utilizando movimento browniano fracionário é utilizado para o reconhecimento de locutor, e os resultados obtidos foram ligeiramente superiores aos obtidos por métodos tradicionais; [128] utiliza características multifractais para realizar a autenticação baseada em imagens faciais; [129] apresenta um método de classificação de escrita cursiva em termos de medidas fractais, que, embora não esteja diretamente ligado à autenticação biométrica, foi utilizado posteriormente por [130] com relativo sucesso. [131] realiza uma modelagem multifractal das veias da palma da mão com o intuito de reconhecimento biométrico. Seu método baseia-se na largura e no valor máximo do espectro multi-fractal e os resultados, assim como nos demais trabalhos, apontam a viabilidade do uso dos fractais. Finalmente, [11] apresenta uma abrangente análise multifractal dos sinais relativos à caminhada (*gait*), mas nenhum experimento de verificação de identidade é realizado.

Esta escassez de trabalhos acerca do assunto foi um dos motivos que culminaram na realização desta dissertação. Até o momento, nenhum estudo foi realizado sobre a eficiência da análise multifractal na verificação biométrica de assinaturas.

Como já foi dito anteriormente, forçaremos a nossa análise em assinaturas dinâmicas, coletadas através de dispostivos adequados, como mesas digitalizadoras (*tablets*). Os sinais fornecidos por tais dispositivos são séries temporais referentes ao instrumento de escrita, como posição no espaço e pressão exercida sobre a superfície.

Os experimentos foram realizados sobre duas bases de assinaturas dinâmicas públicas: a base de assinaturas da primeira competição de verificação de assinaturas (SVC2004) [132] e a base MCYT-100 [133], coletada em conjunto por quatro universidades espanholas.

4.2 Descrição das Bases de Dados

A base SVC2004 é composta por 40 amostras provenientes de 40 usuários distintos, composta por assinaturas chinesas e americanas. Para cada usuário, 20 amostras são autênticas e 20 são falsificações treinadas provenientes de, pelo menos, 5 falsificadores diferentes.

Cada falsificador assistiu a um vídeo da amostra sendo escrita, de modo que fosse possível treinar, além da forma, a dinâmica do assinante. Este vídeo esteve disponível para treino até o copista sentir-se satisfeito com o resultado da cópia.

As amostras genuínas foram coletadas em duas sessões, com no mínimo uma semana de intervalo entre elas, de modo a captar as variações intra-classe dos indivíduos.

É importante salientar que, por questões de privacidade, foi pedido aos voluntários que inventassem uma assinatura diferente daquela que utilizam normalmente.

A base MCYT-100, por sua vez, possui 50 amostras relativas a cada um dos 100 usuários diferentes, sendo 25 destas amostras falsificações treinadas e 25 assinaturas genuínas.

O protocolo de coleta adotado segue o seguinte algoritmo:

1. Voluntário n apresenta-se para a coleta.
2. Para i de 1 até 5
   2.1. Fornece sua própria assinatura 5 vezes;
   2.2. Treina, no mínimo 10 vezes, a assinatura do voluntário $n - i$, observando uma imagem estática da mesma;
   2.3. Fornece 5 cópias da assinatura do voluntário $n - i$;
3. Fim do "Para"

Como o usuário desconcentra-se da tarefa de realizar sua própria assinatura para treinar e forjar outra, consideraremos cada conjunto de 5 amostras genuínas como uma sessão de aquisição independente. Nenhum comentário é feito acerca da originalidade das assinaturas.

Ambas as bases foram coletadas utilizando mesas digitalizadoras da mesma marca, embora não seja especificado o modelo adotado na base SVC2004, e adotam uma freqüência de amostragem de 100Hz. Como a duração típica de uma assinatura é em média 3 segundos, teremos, em geral, cerca de 300 pontos disponíveis por assinatura. Para as bases de dados em questão, a média foi de 310 pontos, com a assinatura mais curta durando 1 segundo e a assinatura mais longa durando em torno de 9 segundos.

As seguintes séries temporais estão disponíveis nas duas bases de assinaturas:

a) Posição no eixo x, X_t
b) Posição no eixo y, Y_t
c) Pressão aplicada pela caneta, P_t
d) Ângulo de azimute da caneta em relação à superfície, Θ_t
e) Ângulo de altitude da caneta em relação à superfície, Φ_t

Apesar de alguns trabalhos apontarem benefícios na utilização das informações de altitude e azimute, neste trabalho, utilizaremos apenas as variáveis X_t, Y_t e P_t.

Escolhemos estas três variáveis devido aos seguintes motivos: a) estudos preliminares indicam que a maior parte da informação discriminativa está nestas variáveis; b) vários dispostivos de coleta não são capazes de fornecer informações referentes aos ângulos Θ_t e Φ_t, o descarte dos mesmos torna mais fácil a comparação dos resultados realizados por estudos posteriores; c) os ângulos são bastante suscetíveis à posição do assinante durante a escrita, e, em alguns casos, pode piorar o resultado obtido devido a um aumento excessivo da variância intra-classe.

Nas Figura 4.1 e Figura 4.2 apresentamos algumas amostras de cada uma das bases de dados, juntamente com as séries temporais X_t, Y_t e P_t. As assinaturas na parte superior da figura são genuínas e as da parte inferior são falsificações. Note que as assinaturas falsas da base SVC2004 são mais parecidas com as assinaturas genuínas que na base MCYT-100.

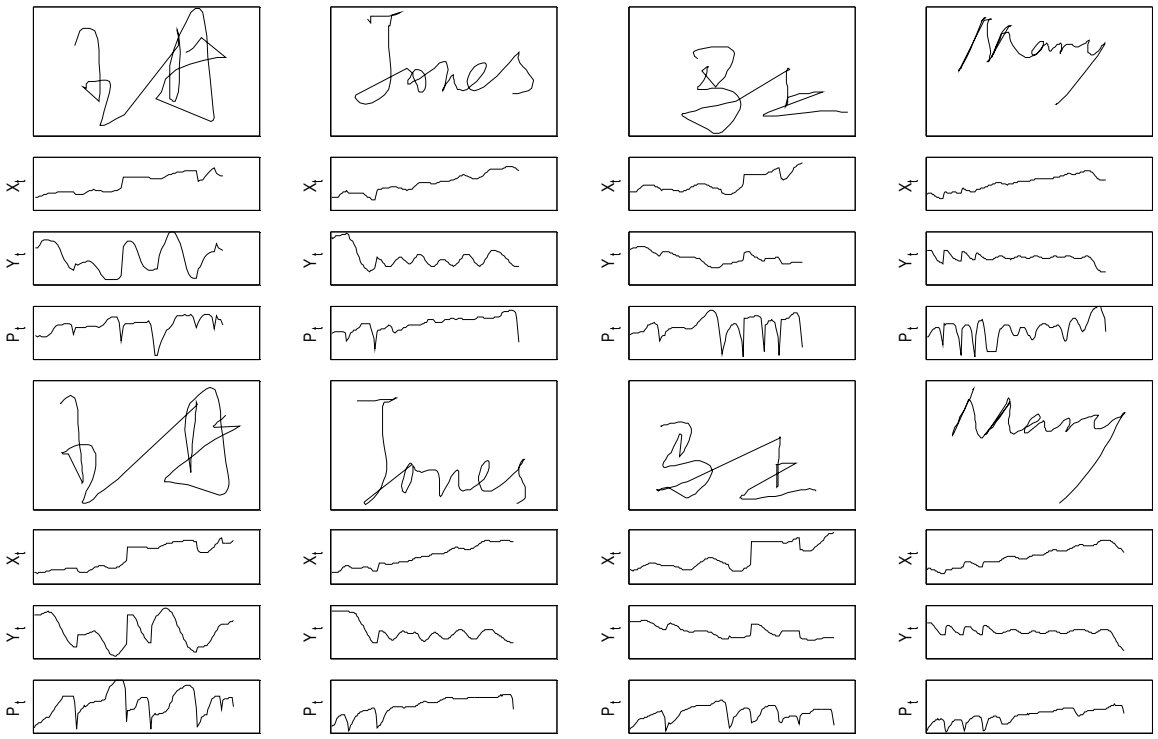


Figura 4.1: Exemplos de assinaturas da base SVC2004.
Assinaturas genuínas na parte superior e falsificações na parte inferior

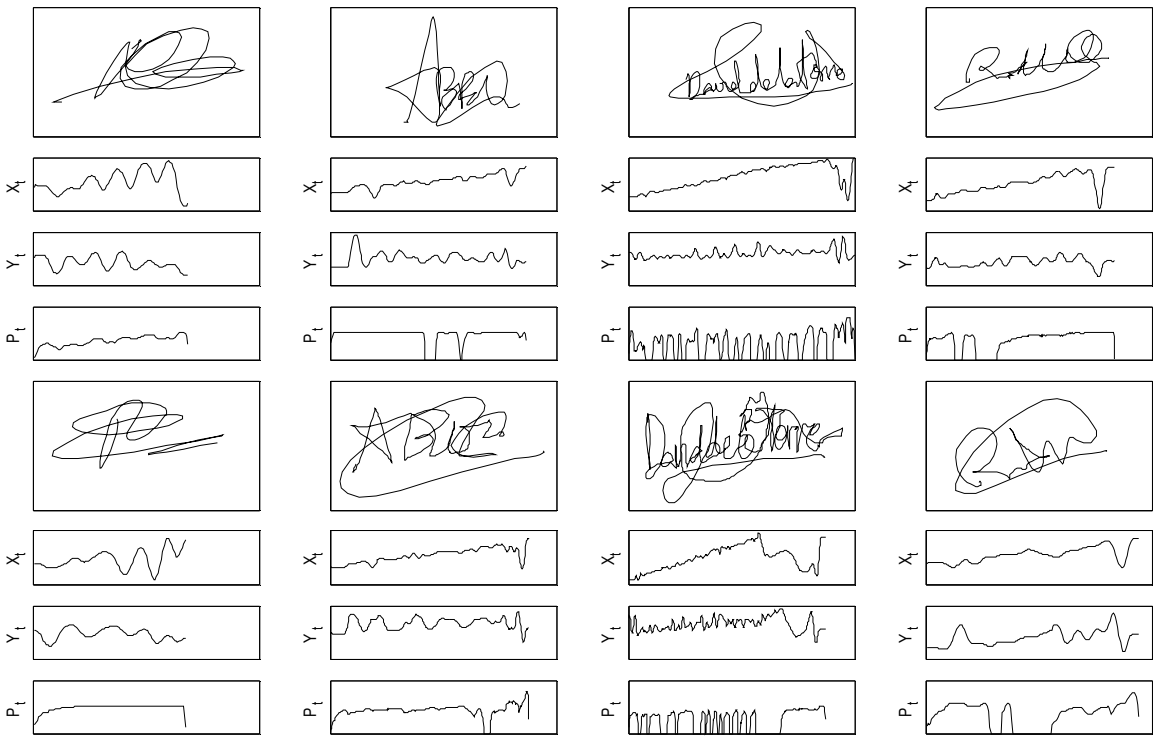


Figura 4.2: Exemplos de assinaturas da base MCYT-100
Assinaturas genuínas na parte superior e falsificações na parte inferior

Como trata-se de uma biométrica comportamental, que pode ser bastante influenciada por fatores impassíveis de controle por parte da equipe científica, vale a pena chamar atenção para algumas pontos importantes acerca das bases utilizadas:

a) Quão adaptados os voluntários estavam ao dispositivo de aquisição antes de iniciar a aquisição? Em experiências de coleta deste tipo, alguns voluntários sentem-se desconfortáveis com o atrito diferente em relação à escrita comum, o que altera significativamente os dados adquiridos, quer seja pela mudança nas forças atuantes no processo físico ou pela "sensação" de incômodo causada pelas mesmas;

b) O longo protocolo de coleta da base MCYT-100 (um total de, pelo menos, 100 assinaturas é gerado em sequência por cada usuário), apesar de fornecer mais tempo para o usuário adaptar-se ao equipamento, pode tornar-se entediante. Qual é seu efeito na qualidade das assinaturas coletadas?

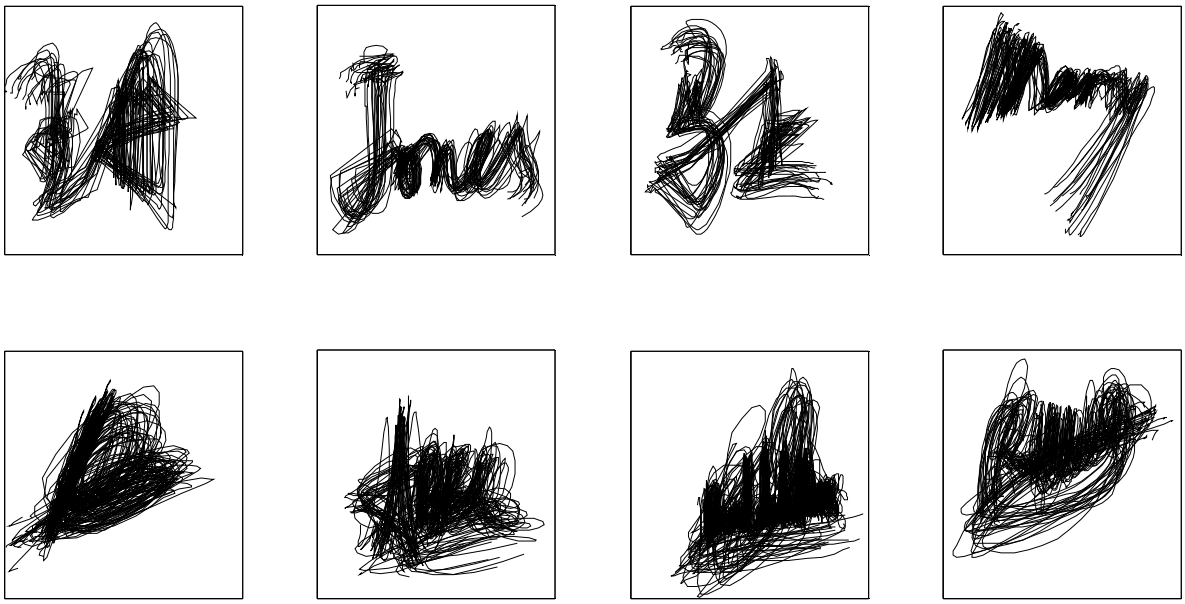
c) A utilização de vídeos para treinamento das falsificações, na base SVC2004, melhorou a qualidade das cópias?

d) A utilização de assinaturas artificiais, na base SVC2004, altera os resultados obtidos?

Na Figura 4.3 fizemos uma sobreposição de todas as assinaturas verdadeiras de quatro usuários da base SVC2004 (parte superior) e da base MCYT-100 (parte inferior). Em seguida, na Figura 4.4, acrescentamos as falsificações correspondentes a cada usuário, em vermelho. Perceba que a variação na base SVC2004 é bem menor. Note também que a qualidade das cópias não é consequência do uso de um vídeo, pois estamos fazendo uma comparação apenas da forma.

Estas são apenas algumas das perguntas que podem ser levantadas diante da utilização de bases biométricas, especialmente no caso de traços comportamentais. Apesar de ser um tópico interessante, e que pode ser discutido longamente, dando espaço para argumentações variadas dependendo do enfoque adotado (e.g., fisiológico, psicológico, estatístico), o aprofundamento destas questões foge ao escopo deste trabalho, e tais pontos são apresentados apenas para instigar o leitor a imaginar a quantidade de variáveis que, em geral, não são levadas em consideração.

SVC2004



MCYT-100

Figura 4.3: Sobreposição de assinaturas genuínas

SVC2004

MCYT-100

Figura 4.4: Sobreposição de assinaturas falsas e genuínas

4.3 Estimação da Característica Fractal

Como pode ser intuído a partir da Definição 3.3.1, e é mostrado em [134] e [135], os estimadores de característica multifractal estão intimamente ligados a estimação de momentos estatísticos de ordem superior a dois, o que representa um grande problema para conjuntos de dados com poucas amostas. Espectros que, teoricamente, deveriam ser reduzidos a um único ponto, o caso de processos monofractais, apresentam-se similar ao multifractal pelo simples fato dos sinais analisados não possuírem comprimento infinito.

Deste modo, para considerar a multifractalidade das assinaturas, comparamos os espectros multifractais das mesmas com aqueles provenientes de processos sabidamente monofractais de mesmo comprimento, conforme sugerido em [11]. A comparação da largura destes espectros determina a presença ou não de características multifractais no processo em teste, em nosso caso, as assinaturas. Neste primeiro teste, independentemente do valor adotado para o parâmetro de Hurst na geração do fBm, os sinais de assinaturas apresentaram um espectro com largura pelo menos 1,5 vezes as dos processos monofractais. Na Figura 4.5 apresentamos o resultado referente aos valores médios obtidos.

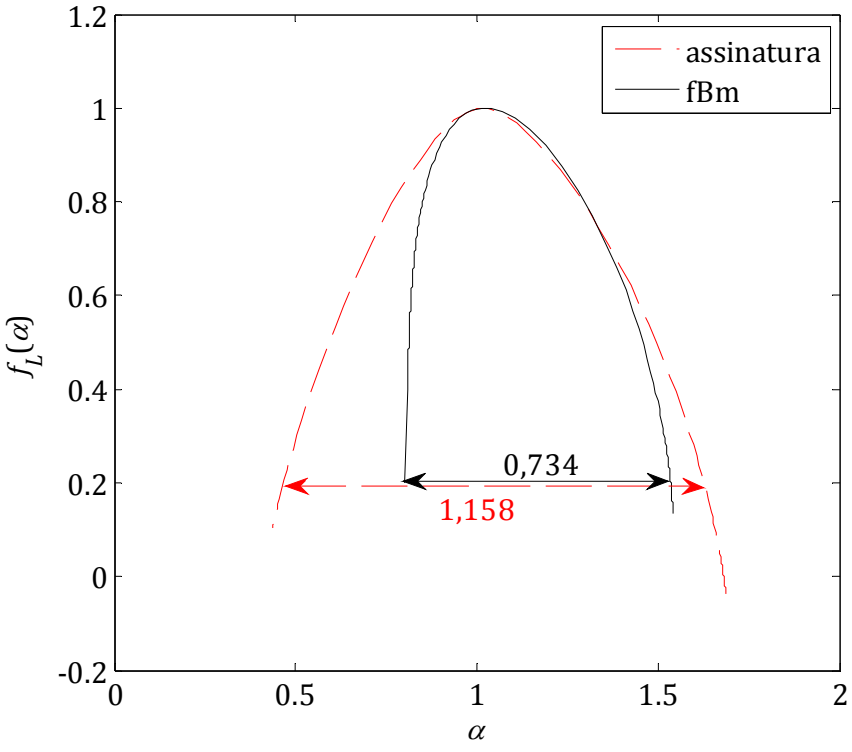


Figura 4.5: Comparação do espectro multifractal. Assinatura x fBm

No entanto, a comparação dos espectros entre assinaturas não mostrou-se suficiente para diferenciar umas das outras. Todas possuem o mesmo formato apresentado na Figura 4.5, com pequenas variações na largura, e abscissa referente ao ponto de máximo oscilando em torno de 1. Na Figura 4.6 apresentamos o espectro de todas as assinaturas genuínas de diversos usuários, onde cada cor representa um usuário distinto. Há sempre a superposição do espectro com pelo menos um outro usuário da base, e uma variação intra-classe da mesma ordem da variação inter-classes.

Por outro lado, esta semelhança dos espectros pode ser um indício de que realmente existe um mecanismo psico-físico comum a todas as pessoas, como sugerem alguns pesquisadores [72].

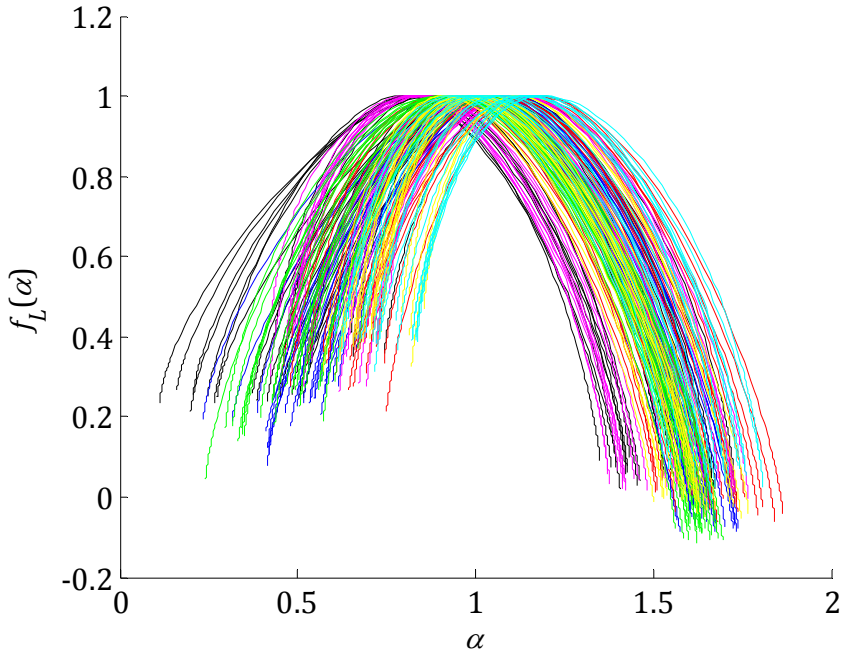


Figura 4.6: Espectro multifractal de diversas assinaturas

Apesar da comparação de espectros, que é o caso mais comum na análise multifractal, não ter dado resultados, esta não é a única solução possível. Resta-nos ainda analisar a informação de regularidade local. Diversos estudos já apontam a utilidade deste tipo de informação em sinais relativamente regulares, como as assinaturas, principalmente na segmentação e filtragem dos mesmos [136] [137] [138] [139].

Portanto, utilizando os métodos de estimação implementados, obtivemos as funções Hölder (i.e., expoente pontual de Hölder para cada instante de tempo) de cada uma das séries temporais referentes às assinaturas. O método *wavelet* foi utilizado sem a detecção de máximos

locais proposta no WTMM, pois tal algoritmo foca nas descontinuidades, sendo inadequado à análise de sinais mais regulares, como as assinaturas [134]. Apenas como ilustração deste fato, apresentamos na Figura 4.7 os resultados da aplicação de cada um dos métodos de estimação a um sinal senoidal. Na Figura 4.8, apresentamos o máximos locais da transformada *wavelet*: note que há poucas opções de linhas de máximo, o que explica a convergência para apenas três valores, que pode ser observada na Figura 4.7.

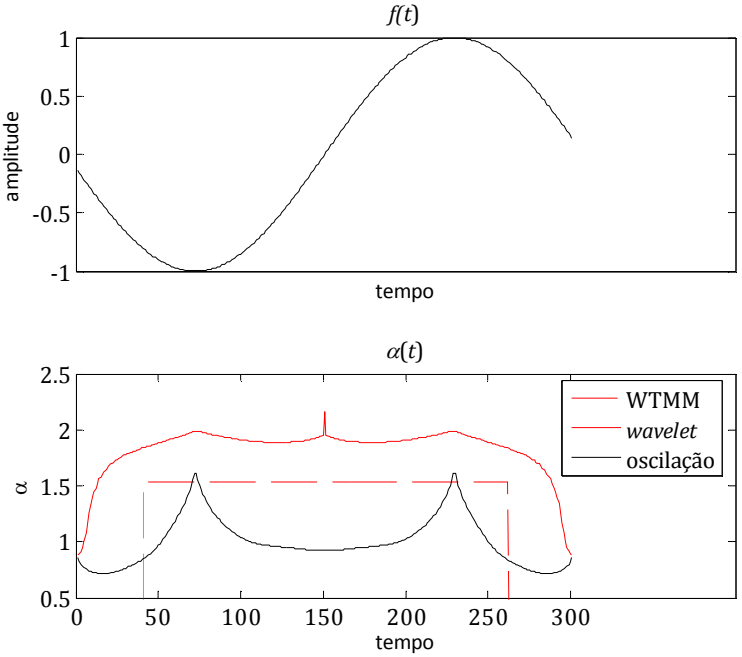


Figura 4.7: Estimação da função Hölder pelos três métodos propostos

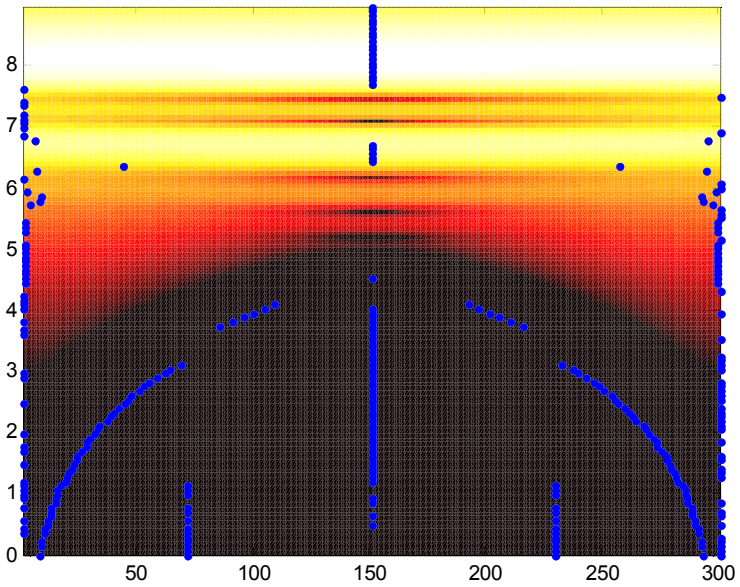


Figura 4.8: Máximos locais da transformada *wavelet* do sinal senoidal da Figura 4.7

A proposta deste trabalho é utilizar a informação de regularidade local, obtida através dos expoentes de Hölder, como uma nova série temporal para a verificação de assinaturas. Três perguntas fundamentais precisam ser respondidas:

1. *É possível distinguir assinaturas através da informação de regularidade local?*
2. *A informação contida nos expoentes de Hölder é redundante com aquela dos sinais originais? Em outras palavras, num contexto de fusão de informação, há um ganho de desempenho do sistema?*
3. *Caso não seja redundante, quanta informação é adicionada pela análise multifractal?*

4.4 Verificação Biométrica de Assinaturas

Como já citamos anteriormente no Capítulo 2 e neste mesmo capítulo, a assinatura é um traço biométrico comportamental, e, como tal, apresenta grande variabilidade para uma mesma pessoa. Este fato faz com que as estratégias de verificação mais eficientes sejam aquelas capazes de tolerar tais mudanças, como as distâncias adaptativas, (e.g., *Dynamic Time Warping* (DTW), distância de Levenshtein) e a modelagem estatística (e.g, Modelos Ocultos de Markov (HMM – *Hidden Markov Models*) e GMM). Outro ponto crucial é que, durante o cadastro do usuário, tipicamente, coletam-se poucas amostras. Neste trabalho utilizaremos apenas 5 assinaturas como referência nas tarefas de classificação.

Este trabalho visa estudar a viabilidade e eficiência da utilização da informação fractal, sem a pretensão de alcançar ou superar os desmpenhos apresentados na literatura. Deste modo, optamos por implementar dois métodos relativamente simples de verificação, que não necessitam de ajustes de parâmetros por parte do usuário, e, ainda assim, fornecem resultados competitivos. O primeiro método, que denominaremos DIST, é baseado na DTW, e o segundo, que chamaremos de STAT, é baseado em modelagem estatística, através do GMM. Nas seções seguintes descrevemos com mais detalhes as duas técnicas adotadas

4.4.1 Método DIST

As séries temporais referentes às assinaturas são, em geral, de comprimentos diferentes. Portanto, para realizar uma comparação, é necessário sincronizar as duas séries temporais. O

método mais simples para tal é fazer uma reamostragem, utilizando algum método de interpolação, para que as duas sequências apresentem o mesmo tamanho. Porém, o resultado de tal processo geralmente resulta em uma sincronização ruim.

Para sanar este problema, adotamos um método de programação dinâmica, amplamente utilizado em processamento de sinais de voz, conhecido como DTW [27]. Mostramos na Figura 4.9 as comparações feitas pela reamostragem e pelo DTW de dois sinais de comprimentos e formas distintos. Na Figura 4.10 mostramos a superposição dos sinais após a reamostragem e após o DTW para facilitar a visualização da sincronização realizada por cada método.

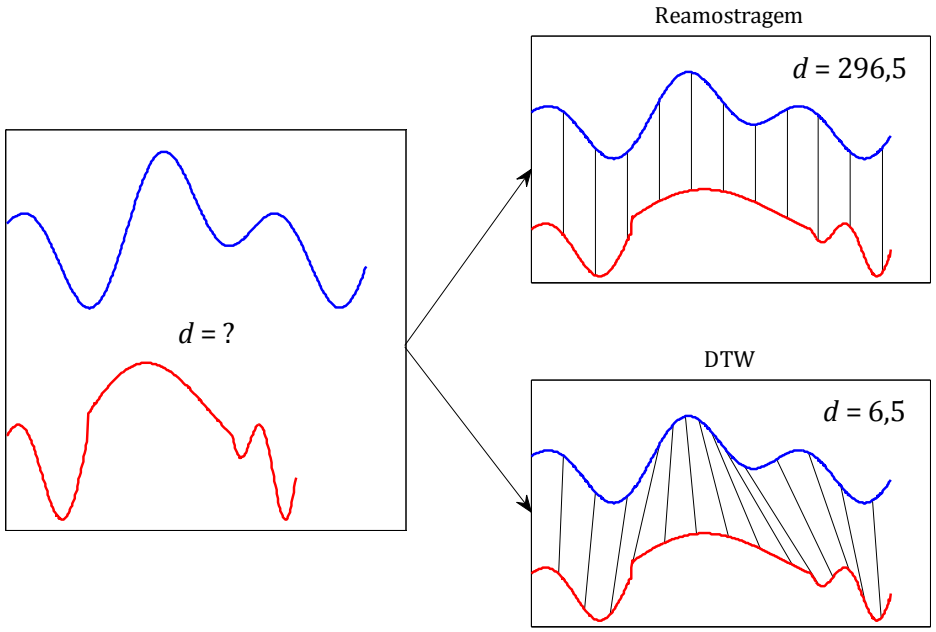


Figura 4.9: Comparações de dois sinais de comprimentos diferentes realizadas pelo método da reamostragem e DTW

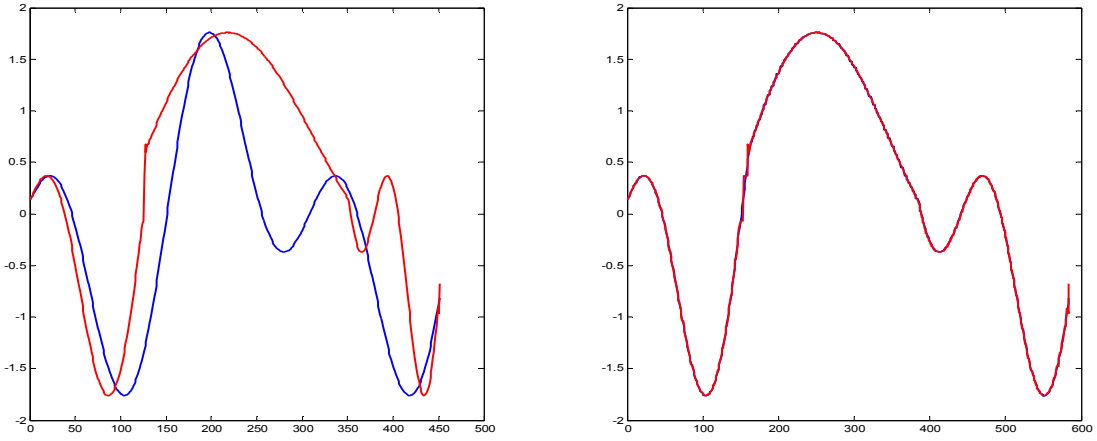


Figura 4.10: Superposição das curvas da Figura 4.9 após a sincronização através da reamostragem e do DTW

Descrevemos o problema em questão e a solução DTW em seguida. Sejam $R(i)$, $i = 1,2,\dots,I$, e $T(j)$, $j = 1,2,\dots,J$, as séries de referência e teste respectivamente, onde, possivelmente, $I \neq J$. Nosso objetivo é desenvolver uma medida de distância adequada entre estas sequências. Para tal, montaremos uma matriz M de tamanho $I \times J$, onde cada elemento $m_{i,j}$ representa um custo, dado pela função $d(i,j)$. Um caminho C ao longo desta matriz que vai de $m_{0,0}$ até um ponto final $m_{f,f}$ é um conjunto ordenado de pontos da forma

$$C = c(1), c(2), \dots, c(f) \tag{4.1}$$

onde $c(k) = (i_k, j_k)$. Associado a este caminho, definimos o custo total do caminho como:

$$D(C) = \sum_{k=1}^{f} d(i_k, j_k) \tag{4.2}$$

A distância entre R e T é definida como o menor D dentre todos os possíveis caminhos existentes. Obviamente, calcular todos os caminhos possíveis é um procedimento de custo computacional elevado e, muitas vezes, impraticável. Para reduzir este custo, utilizamos o princípio de programação dinâmica de Bellman [140].

Seja o caminho ótimo entre um ponto inicial (i_1, j_1) e um ponto final (i_f, j_f) denotado por

$$(i_1, j_1) \xrightarrow{o} (i_f, j_f) \tag{4.3}$$

Se (i_k, j_k) é um ponto intermediário entre o ínicio e o fim, vamos denotar o caminho ótimo que passa por (i_k, j_k) como

$$(i_1, j_1) \xrightarrow{o(i_k, j_k)} (i_f, j_f) \tag{4.4}$$

O princípio de Bellman diz que

$$(i_1, j_1) \xrightarrow{o(i_k, j_k)} (i_f, j_f) = (i_1, j_1) \xrightarrow{o} (i_k, j_k) \oplus (i_k, j_k) \xrightarrow{o} (i_f, j_f) \tag{4.5}$$

onde $\oplus$ é a concatenação dos caminhos. Em resumo, o caminho ótimo que passa por um dado ponto intermediário, é a combinação do caminho ótimo até o ponto intermediário com o caminho ótimo deste ponto até o final. Ora, estendendo este raciocínio pode-se facilmente concluir que,

para obter o caminho ótimo, basta que cada passo seja ótimo. No entanto, com o intuito de reduzir o espaço de busca, normalmente adotam-se restrições (heurísticas) acerca de quais passos são permitidos a partir de um dado ponto, portanto, o resultado obtido pode ser sub-ótimo. Em nossa implementação, duas restrições são aplicadas:

a) O caminho deve ser *completo*, ou seja incia em (i_1, j_1) e termina em (i_I, j_J)
b) Estando no ponto (i_k, j_k), pode-se deslocar apenas para os pontos (i_{k+1}, j_k), (i_k, j_{k+1}) ou (i_{k+1}, j_{k+1}) , como mostramos na Figura 4.11.

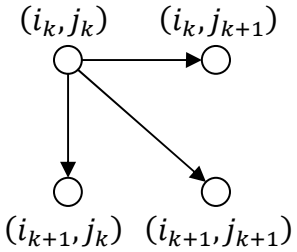


Figura 4.11: Restrições de busca local para o DTW implementado

Além disso, a função custo adotada foi a distância euclidiana, ou seja

$$d(i_k, j_k) = \|R(i_k) - T(j_k)\| \tag{4.6}$$

Existem diversas variações e melhoramentos deste método, a versão aqui apresentada pode ser vista como a clássica, ou básica, do mesmo. Vale notar ainda que a distância entre R e T obtida por este método não é uma métrica, no sentido matemático, como visto na Definição 3.3.3, pois não possui a propriedade de desigualdade triangular. Como uma ilustração do método, apresentamos na Figura 4.12 a matriz de distância acumulada e o caminho ótimo encontrado para as séries da Figura 4.9.

Como utilizamos cinco assinaturas de referência, a assinatura em teste é comparada a cada uma das referências e adotamos como pontuação a média das distâncias obtidas. Além disso, sabemos que as falsificações, em geral, são mais lentas que assinaturas genuínas, pois o falsificador precisa tomar mais cuidado ao tentar reproduzir sua forma. De fato, utilizando apenas o tempo total para classificação das assinaturas, é possível obter uma EER de cerca de 22%, em

ambas as bases utilizadas neste trabalho. Portanto, as distâncias obtidas pelo algoritmo DTW são multiplicadas por um fator k_d dador por

$$k_d = \frac{|t_{teste} - t_{médio}|}{t_{médio}} + 1 \tag{4.7}$$

onde t_{teste} é o tempo total da assinatura sendo comparada e $t_{médio}$ é o tempo total médio das assinaturas usadas como referência, deste modo estamos penalizando as assinaturas que divergem do tempo médio do modelo do usuário.

As distâncias entre assinaturas genuínas de uma mesma pessoa, como definido no Capítulo 2, serão chamadas de Pontuações Genuínas e as demais comparações serão chamadas de Pontuações Impostoras.

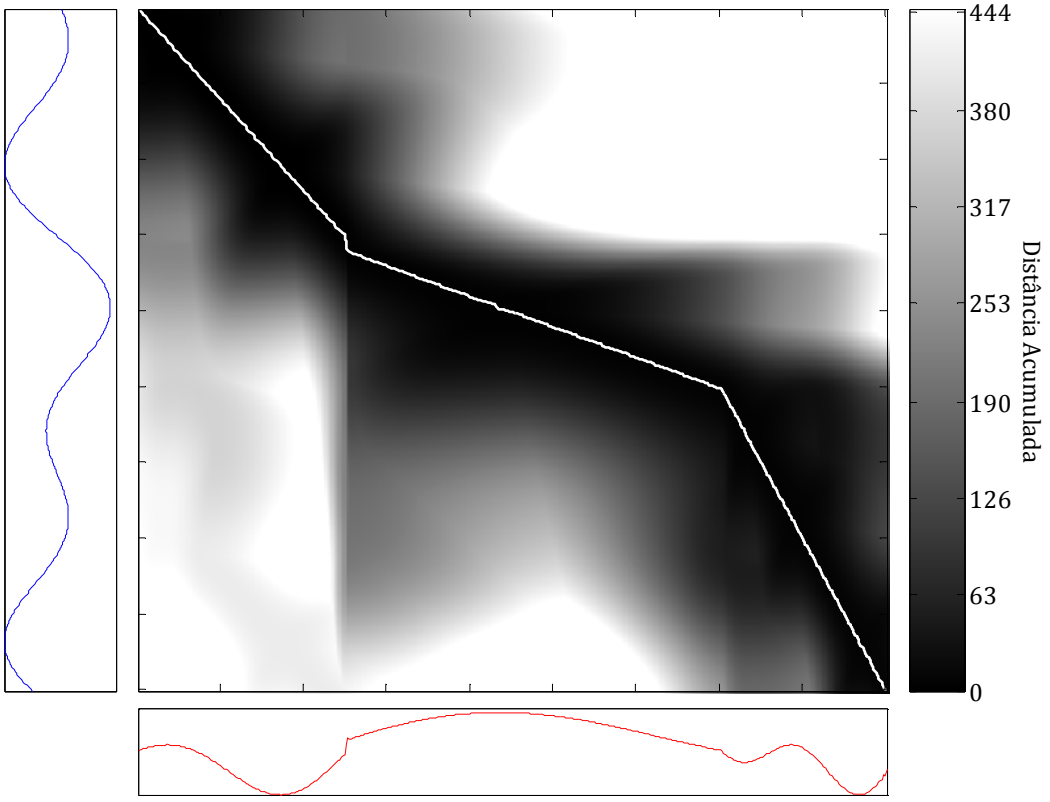


Figura 4.12: Matriz de distância acumulada e caminho ótimo para as curvas da Figura 4.9

4.4.2 Método STAT

A idéia da modelagem estatística para classificação de padrões consiste em, a partir de um conjunto de dados, estimar a função densidade de probabilidade dos mesmos e, no momento da comparação, medir a verossimilhança de o padrão em teste pertencer àquela distribuição. Dentro deste contexto, o modelo de mistura de Gaussianas (GMM) é bastante utilizado [26] [27].

O problema pode ser enunciado como a seguir: Seja $\boldsymbol{X} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_N\}$ um conjunto de amostras de dimensão D provenientes de uma pdf desconhecida $p(\boldsymbol{x})$. Assumiremos que existe uma aproximação adequada para $p(\boldsymbol{x})$ dada por uma mistura de distribuições gaussianas:

$$p(\boldsymbol{x}) \approx \hat{p}(\boldsymbol{x}|\boldsymbol{\Theta}) = \sum_{k=1}^{M} \alpha_k G(x|\boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k) \tag{4.8}$$

onde $\boldsymbol{\Theta}$ são os parâmetros da mistura (i.e., $\mathrm{A} = [\alpha_1, \dots, \alpha_M]$, $\mathrm{M} = [\boldsymbol{\mu}_1, \dots \boldsymbol{\mu}_M]$ e $C = [\boldsymbol{\Lambda}_1, \dots, \boldsymbol{\Lambda}_M]$) e $G(x|\boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k)$ é o k-ésimo núcleo gaussiano da mistura, com vetor de médias $\boldsymbol{\mu}_k$ e matriz de covariância $\boldsymbol{\Lambda}_k$:

$$G(\boldsymbol{x}|\boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k) = \frac{1}{(2\pi)^{\frac{D}{2}} |\boldsymbol{\Lambda}_k|^{\frac{1}{2}}} exp\left(-\frac{(\boldsymbol{x} - \boldsymbol{\mu}_k)^T \boldsymbol{\Lambda}_k^{-1} (\boldsymbol{x} - \boldsymbol{\mu}_k)}{2}\right) \tag{4.9}$$

Denotando a verossimilhança de $\boldsymbol{X}$ por $L(\boldsymbol{\Theta}) = p(\boldsymbol{X}|\boldsymbol{\Theta})$, o problema em questão resume-se a encontrar o conjunto de parâmetros ótimo, $\boldsymbol{\Theta}_0$, que maximiza o logaritmo da verossimilhança:

$$\boldsymbol{\Theta}_0 = \underset{\Theta}{\operatorname{argmax}}(log\, L(\boldsymbol{\Theta})) \tag{4.10}$$

Se as N amostras são independentes, podemos escrever

$$log\, L(\boldsymbol{\Theta}) = log\left(\prod_{k=1}^{N} p(\boldsymbol{x}_k|\boldsymbol{\Theta})\right) = \sum_{k=1}^{N} log\, p(\boldsymbol{x}_i|\boldsymbol{\Theta}) \tag{4.11}$$

Um método bem conhecido, e amplamente difundido, para a resolução deste problema é o algoritmo EM (*Expectation-Maximization*) [141]. No entanto, além de não ser o método de estimação mais rápido, este algoritmo apresenta alguns outros problemas como [142]: a)

convergência potencialmente ruim, a depender da distribuição dos dados e da escolha inicial para seus parâmetros; b) seu critério baseado em verossimilhança apresenta diversos máximos que resultam em modelos ruins, o que é especialmente verdade para conjuntos de dados pequenos, onde esta convergência inadequada ocorre com maior frequência.

De fato, sabe-se que a verossimilhança é pouco representativa em problemas de dimensão elevada, e até mesmo em alguns casos de baixa dimensão [143]. Uma solução para tais quesitos é a utilização de métodos de regularização, que impõem restrições ao problema de otimização para aumentar a capacidade de generalização do mesmo [144].

Tendo em vista tais limitações, e como já vimos que o estamos diante de um problema com poucos dados disponíveis, optamos por utilizar o método de Parzen, conforme proposto em [145]. O método de Parzen, originalmente não-paramétrico, é considerado neste trabalho segundo uma perspectiva paramétrica, como uma GMM regularizada.

Inicialmente, dividimos as amostras $\boldsymbol{X}$ disponíveis em dois grupos, $\boldsymbol{P} = \{\boldsymbol{p}_1, \boldsymbol{p}_2, \dots, \boldsymbol{p}_M\}$ para geração dos protótipos e $\boldsymbol{V} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_{N-M}\}$ para a validação. As restrições impostas são as seguintes:

a) $\alpha_k = 1/M$, ou seja, todas os núcleos gaussianos do Parzen possuem o mesmo peso;
b) $\boldsymbol{\mu}_k = \boldsymbol{x}_k$, as amostras são os centros das gaussianas;
c) $\boldsymbol{\Lambda}_k = \sigma^2\mathbf{I}$, onde $\mathbf{I}$ é a matriz identidadade $D \times D$, de modo que os núcleos gaussianos são todos idênticos e isotrópicos.

Claramente, dadas tais restrições, o único parâmetro livre do sistema é σ, que otimizaremos conforme o seguinte algoritmo:

1. Para cada amostra p_k de P, encontram-se os dois vizinhos mais próximos, p_i e p_j, e realiza-se uma estimação grosseira da variância.

$$s_k^2 = \left(\|\boldsymbol{p}_k - \boldsymbol{p}_i\|^2 + \left\|\boldsymbol{p}_k - \boldsymbol{p}_j\right\|^2\right)/2 \tag{4.12}$$

2. Toma-se a mediana s_m^2 das variâncias estimadas no passo 1, e define-se:

$$s_{min} = \frac{s_m}{10}, \qquad s_{max} = 10s_m, \qquad \Delta s = s_m/20 \tag{4.13}$$

3. Obtém-se um modelo de mistura de Gaussianas conforme a Equação 4.8 e as restrições mencionadas.

4. Com os elementos do grupo de validação, calcula-se o logaritmo da verossimilhança para cada valor de s no conjunto $\{s_{min}, s_{min} + \Delta s, \dots, s_{max}\}$ e adota-se o valor σ_o que maximiza $\log L(\Theta)$

$$\sigma_o = \underset{s}{\operatorname{argmax}} \left(\sum_{k=1}^{N-M} log\, \hat{p}(\boldsymbol{v}_k|s) \right), \qquad (s \in \{s_{min}, s_{min} + \Delta s, \dots, s_{max}\}) \tag{4.14}$$

Utilizamos apenas cinco assinaturas para gerar o modelo do usuário, como no caso anterior. Este procedimento é então repetido cinco vezes, cada uma das quais utilizando uma assinatura diferente para validação. O σ_0 adotado para o modelo é a média dos σ obtidos em cada validação. Neste caso, as pontuações genuínas e impostoras são formadas pelo negativo dos valores do logaritmo da verossimilhança. Mais detalhes acerca deste método de estimação de pdf, com exemplos e comparações ao algoritmo EM, podem ser encontrados em [145].

Para a aplicação em assinaturas, especificamente, Montalvão et al. [146] propõem o acréscimo de uma dimensão aos dados, referente ao tempo. Esta dimensão é uma forma simples de fazer com que as coordenadas de pontos de uma assinatura que foram coletadas em instantes afastados também sejam representados separadamente no espaço dos sinais, permitindo uma modelagem adequada das características dinâmicas do sinal. A Figura 4.13 ajuda a entender esta idéia, perceba que não é possível distinguir na visão frontal, que desconsidera o tempo, os dois pontos selecionados.

Em cada ponto da assinatura, o método Parzen, conforme proposto, posiciona um núcleo gaussiano onde a probabilidade de ocorrência de um ponto reduz-se com o desvio padrão. Como utilizamos cinco assinaturas para criar o modelo, estes núcleos combinam-se e criam uma região em torno da assinatura onde os pontos devem estar localizados com maior probabilidade. Para facilitar a visualização, apresentamos na Figura 4.14 essa região apenas para a variável x no tempo. Note que as áreas com poucas variações entre as amostras de referência apresentam uma maior densidade de probabilidade.

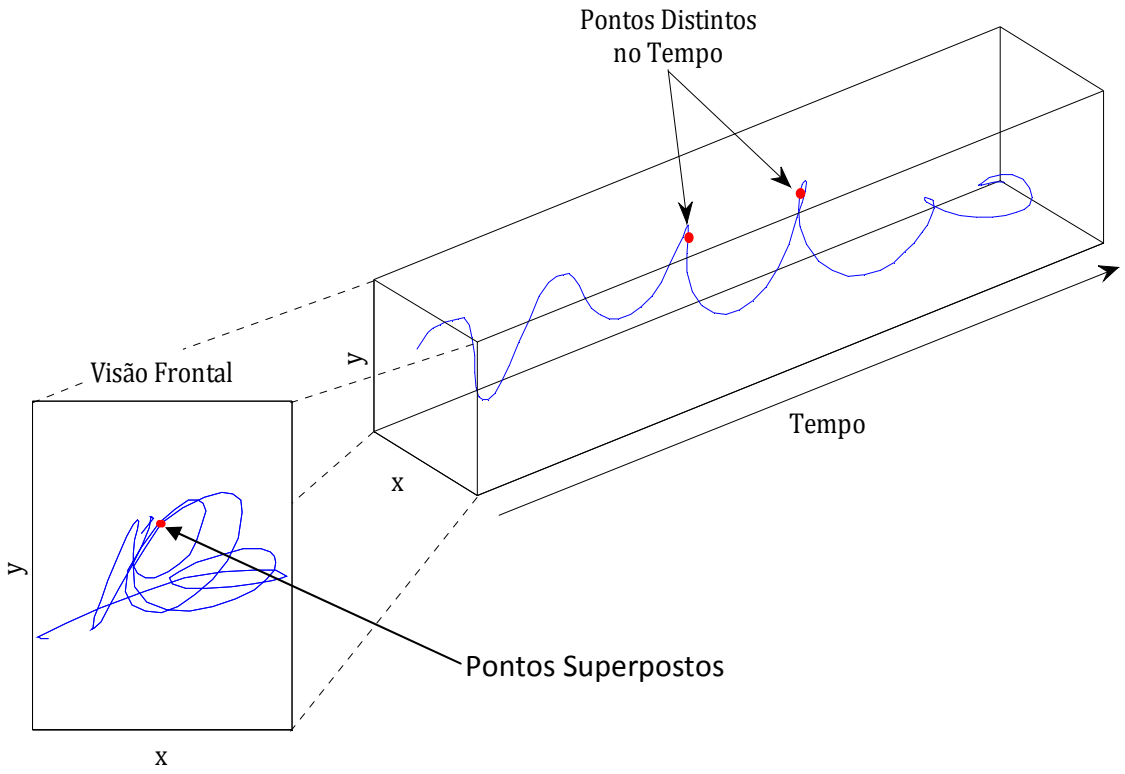


Figura 4.13: Efeito da inserção do tempo no espaço dos sinais

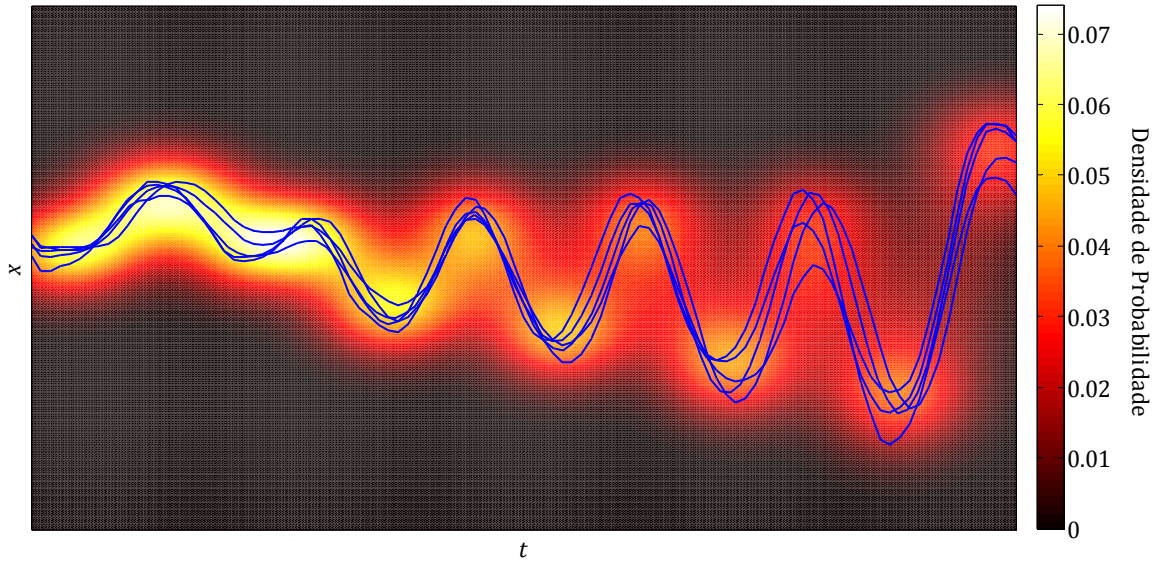


Figura 4.14: Estimação da densidade de probabilidade utilizando o método STAT. Cada curva representa uma assinatura genuína utilizada como modelo.

Semelhante ao que foi feito para o método DIST, o algoritmo proposto em [146] foi modificado para incluir uma penalização às assinaturas que divergem do tempo médio das assinaturas de referência. Neste caso a dimensão de tempo introduzida é esticada, ou comprimida, por um fator k_S dado por

$$k_S = \frac{t_{teste}}{t_{médio}} \tag{4.15}$$

Note que este método é, de certa forma, semelhante ao método DIST, no qual utilizamos cinco amostras e consideramos a distância média. Porém, neste caso, as variações locais são melhor modeladas. Pode-se ainda enxergá-lo como uma versão probabilística do DTW, no entanto, uma análise detalhada de tais relações que propomos aqui apenas intuitivamente é assunto suficiente para uma outra dissertação.

4.5 Resultados Experimentais

Na verificação de assinaturas, os impostores podem ser divididos em dois tipos: aleatórios e treinados. Falsificações aleatórias são aquelas em que o falsificador não conhece a assinatura real e, portanto, utiliza qualquer coisa, até mesmo a sua própria assinatura, para tentar burlar o sistema. Impostores treinados, por outro lado, representam os reais desafios para os sistemas de verificação, pois, nestes casos, as assinaturas podem ser bastante semelhantes. De um modo geral, dado um sistema de reconhecimento de assinaturas, o desempenho na presença de falsificações aleatórias é significativamente superior àquele na presença de falsificações treinadas. Por este motivo, as análises que faremos neste trabalho estarão focadas neste último tipo de falsificação.

Propomos a utilização da análise multifractal como uma ferramenta de pré-processamento e extração de características. As funções Hölder podem ser vistas como uma transformação dos dados originais que, possivelmente, fornecerão aos sistemas de verificação informações que antes estavam implícitas no sinal bruto.

Primeiramente, apresentamos os resultados dos dois métodos implementados, utilizando os dados brutos, X_t, Y_t e P_t, e suas respectivas funções Hölder, que serão indicadas deste ponto em diante pelos expoentes $\sim$ e ψ (i.e., $X_t^{\sim}$, Y_t^{ψ}), representando os métodos de estimação por oscilação e *wavelet*, respectivamente. O expoente H indicará função Hölder, sem especificar o

método (i.e., $X_t^H \asymp Y_t^{\psi}$ ou $Y_t^{\sim}$). Na Tabela 4.1 constam os resultados referentes à base SVC2004. A Tabela 4.2, mostra os resultados para a base MCYT-100.

Tabela 4.1 - Resultados de Verificação Utilizando 1 Variável (Base SCV2004)

	EER_{DIST} (%)	EER_{STAT} (%)
X_t	25,3±1,9	15,3±0,3
Y_t	17,1±0,8	13,7±0,3
P_t	26,9±1,2	15,8±0,6
$X_t^{\sim}$	23,1±0,6	16,2±0,8
$Y_t^{\sim}$	25,2±0,7	17,2±0,4
$P_t^{\sim}$	29,7±1,1	17,4±0,8
X_t^{ψ}	31,4±1,2	17,8±0,7
Y_t^{ψ}	30,6±1,0	19,8±0,6
P_t^{ψ}	31,9±0,6	17,9±0,8

Tabela 4.2 - Resultados de Verificação Utilizando 1 Variável (Base MCYT-100)

	EER_{DIST} (%)	EER_{STAT} (%)
X_t	16,4±0,3	8,5±0,1
Y_t	6,9±0,3	7,2±0,1
P_t	12,0±0,4	8,5±0,2
$X_t^{\sim}$	13,9 ±0,5	9,4±0,4
$Y_t^{\sim}$	15,1±0,2	9,8±0,5
$P_t^{\sim}$	16,6±0,2	11,3±0,2
X_t^{ψ}	13,0±0,5	10,1±0,3
Y_t^{ψ}	14,1±0,2	10,6±0,3
P_t^{ψ}	9,6±0,1	10,9±0,7

Os desvios padrões foram calculados a partir de 10 execuções independentes, escolhendo-se, a cada uma delas, um grupo de 5 assinaturas diferente como referência.

Algumas informações importantes podem ser inferidas a partir deste primeiro conjunto de resultados. Inicialmente, quanto aos métodos escolhidos, percebe-se que o método estatístico é capaz de modelar melhor os dados: sua superioridade pode ser notada, em especial na base SVC2004. Podemos ainda perceber que o desempenho na base MCYT-100 é quase duas vezes melhor que na base SVC2004. Os motivos para esta discrepância observada ainda estão sendo investigados, no entanto, crê-se que os seguintes fatores são suas principais causas:

1. A utilização de assinaturas artificiais pode ter removido uma característica marcante das assinaturas: o movimento balístico. Quando utilizamos a assinatura a que estamos acostumados não há um mecanismo de realimentação [72] no movimento de escrita, são movimentos semelhantes aos reflexos;

2. O treinamento com informação temporal, na base SVC2004, pode ter reduzido significativamente as variações inter-classes. Como vimos na Seção 4.2, isto é verdade quanto à forma das assinaturas.
3. A presença de assinaturas chinesas na base SVC2004, que são, em geral, compostas por segmentos de curva curtos e rápidos ao invés de curvas mais suaves e contínuas, como as assinaturas ocidentais;
4. Quanto às funções Hölder, que possuem desempenho muito inferior na base SVC2004, percebemos ser resultado da forma de coleta empregada. Na base MCYT-100 os pontos são registrados de forma ininterrupta, enquanto na base SVC2004 a aquisição é pausada quando não há contato da caneta com a superfície (pressão nula), o que provoca uma série de descontinuidades artificiais nas séries temporais.

Uma análise detalhada destes fatores, no entanto, foge ao escopo desta tese. Outro ponto, este importante para os objetivos deste trabalho, é que, apesar dos resultados de classificação obtidos com base nas funções Hölder serem piores que aqueles obtidos pelos dados brutos, a primeira questão proposta já pode ser respondida:

1. ***É possível utilizar informação de regularidade local para classificar assinaturas?***
   R: *Sim, embora os resultados não sejam tão bons quanto aqueles obtidos através dos dados originais.*

Sendo a resposta é afirmativa, pode-se continuar a análise e partir em busca de uma resposta para a segunda questão. A etapa seguinte consiste na realização de experimentos de fusão de informação. Neste caso, têm-se três possibilidades:

1. **Fusão de Dados:** Concatenar as variáveis e apresentá-las aos algoritmos como um único vetor multi-dimensional;
2. **Fusão de Pontuações:** Combinam-se as pontuações, através de uma operação adequada, e só então a classificação é realizada;
3. **Fusão de Decisões:** A classificação é realizada independentemente em cada variável, como apresentado há pouco nas Tabela 4.1 e Tabela 4.2, e as decisões são combinadas de modo adequado (e.g., através de uma votação).

Perceba que, para os tipos 2 e 3, é possível realizar também a fusão dos resultados dos dois algoritmos. No entanto, pelos motivos citados na Seção 2.6.1, não implementaremos a fusão no nível de decisão. As sessões seguintes detalham as abordagem adoatadas para os casos 1 e 2, e os resultados obtidos com os mesmos.

4.5.1 Fusão no Nível dos Dados

Se considerarmos todas as combinações tomadas 2 a 2, 3 a 3, e assim por diante, das 9 variáveis disponíveis, temos um total de 511 combinações possíveis. Como a realização de múltiplos experimentos, para estimação dos intervalos de confiança para todas estas possibilidades tomaria um tempo inadmissível (cerca de 60 dias). Optou-se por utilizar apenas algumas das combinações possíveis, listadas a seguir:

a) $C(X,Y)$ e $C(X^H,Y^H)$: Estas combinações representam o caso onde apenas os dados referentes à posição estão disponíveis, como acontece em alguns dispositivos de coleta. As combinações dos dados brutos e das funções Hölder são realizadas independentemente para comparar o desempenho de classificação, de modo semelhante ao que foi apresntado das Tabela 4.1 e Tabela 4.2.
b) $C(X,Y,P)$ e $C(X^H,Y^H,P^H)$: Neste caso, todas as variáveis disponíveis são utilizadas, mas os dados brutos e suas funções Hölder ainda são considerados separadamente.
c) $C(X,X^H)$, $C(Y,Y^H)$ e $C(P,P^H)$: Cada variável é combinada com sua respectiva função Hölder, os resultados obitdos neste caso podem ser comparados àqueles das Tabela 4.1 e Tabela 4.2 para analisar se há um ganho de desempenho resultante da fusão de informação.
d) $C(X,Y,X^H,Y^H)$: Análogo ao que foi feito no item anterior, mas compara-se aos resultados das combinações no item (a)
e) $C(X,Y,P,X^H,Y^H,P^H)$: Também análogo ao item (c). Neste caso, compara-se aos resultados obtidos no item (b).

onde $C(V_1,\dots V_n)$ representa a concatenação das variáveis V_i em um vetor n-dimensional (e.g., $C(X,Y)=[X_t \quad Y_t]$, $C(X,X^H)=[X_t \quad X_t^H]$). Para reduzir a notação utilizaremos $C(S)$ para representar $C(X,Y,P)$ e de forma análoga $C(S^H)$ representa $C(X^H,Y^H,P^H)$. Em todos os casos,

as variáveis são normalizadas para média nula e variância unitária antes de serem concatenadas e apresentadas aos algoritmos. Além disso, para os casos onde X e Y são considerados conjuntamente, as assinaturas são giradas de modo que seu eixo principal esteja na direção horizontal, reduzindo as diferenças relativas à orientação da escrita.

4.5.2 Fusão no Nível das Pontuações

Neste caso há uma quantidade ainda maior de combinações possíveis, pois podemos considerar, além das pontuações obtidas utilizando as variáveis de forma independente, as pontuações relativas à fusão no nível dos dados, o que nos daria um total de 520 variáveis e algo da ordem de $3{,}4 \times 10^{156}$ combinações possíveis. Para manter a consistência das análises, realizaremos as mesmas combinações propostas para a fusão no nível dos dados.

Existem diversos estudos relacionados à fusão de informação nas mais variadas áreas da ciência. Neste trabalho aplicamos uma técnica bastante simples, que já foi utilizada com sucesso na fusão de dados biométricos [147] [148]. Este método, conhecido apropriadamente como regra da soma (*sum rule*), consiste simplesmente em somar as pontuações obtidas e escolher um novo limiar adequado.

Em [149], realizamos uma transformação não-linear das pontuações, com base em sua distribuição de probabilidade conforme sugerido em [150] e [151], o que melhorou significativamente o desempenho do sistema. No entanto, como este procedimento requer a utilização de informações concernentes às falsificações, que numa aplicação real não estariam disponíveis, optamos por não realizá-lo aqui.

4.5.3 Resultados dos Experimentos de Fusão

Inicialmente, continuaremos a comparação de desempenho dos algoritmos sem realizar a fusão entre a informação fractal e os dados brutos, consideraremos apenas as combinações de um mesmo tipo de informação. Na Tabela 4.3 apresentamos tais resultados para a base SVC2004. Para conveniência do leitor, reproduziremos quando necessário os resultados das Tabela 4.1 e Tabela 4.2.

Tabela 4.3: Combinações de dados de um mesmo tipo para a base SVC2004

SVC2004	DIST		STAT	
	EER_{DADOS} (%)	EER_{PONTOS} (%)	EER_{DADOS} (%)	EER_{PONTOS} (%)
X	25,3±1,9	–	15,3±0,3	–
Y	17,1±0,8	–	13,7±0,3	–
P	26,9±1,2	–	15,8±0,6	–
$C(X,Y)$	16,2±0,8	17,0±2,1	10,8±0,5	18,5±0,4
$C(S)$	14,9±0,4	18,9±2,4	9,4±0,6	16,4±0,7
$X^{\sim}$	23,1±0,6	–	16,2±0,8	–
$Y^{\sim}$	25,2±0,7	–	17,2±0,4	–
$P^{\sim}$	29,7±1,1	–	17,4±0,8	–
$C(X^{\sim},Y^{\sim})$	16,1±0,6	21,9±0,9	15,8±1,2	16,9±0,9
$C(S^{\sim})$	15,4±1,3	21,7±1,4	15,0±1,1	16,1±1,2
X^{ψ}	31,4±1,2	–	17,8±0,7	–
Y^{ψ}	30,6±1,0	–	19,8±0,6	–
P^{ψ}	31,9±0,6	–	17,9±0,8	–
$C(X^{\psi},Y^{\psi})$	25,7±0,8	30,4±1,6	18,2±1,1	22,2±0,9
$C(S^{\psi})$	15,5±1,0	30,4±1,7	20,0±0,7	20,4±1,2

O conjunto de resultados para a base SVC2004 também apresenta desempenhos atípicos na fusão dos dados. Para o método DIST, como se é esperado, o desempenho melhora conforme acrescentamos mais informação, no entanto, é curioso notar que em todos os três casos $C(S^*)$ possui o mesmo desempenho, com EER em torno de 15%.

Para o método STAT, os resultados também melhoram com a inclusão de novas informações, com exceção do caso onde as funções Hölder são estimadas através do método *wavelet*. Isto se dá porque, como explicamos na Seção 4.3, não utilizamos o método WTMM, pois tratavam-se de sinais bastante regulares. No entanto, para a base SVC2004, especificamente, as descontinuidades na aquisição geram diversas singularidades nas séries temporais, onde provavelmente seria mais adequado a utilização da WTMM.

Os erros introduzidos por esta estimação inconsistente são modelados pelo método estatístico, provocando uma degradação do desempenho. Na Tabela 4.4, realizamos a mesma análise para a base MCYT-100.

Tabela 4.4: Combinações de dados de um mesmo tipo para a base MCYT-100

MCYT-100	DIST		STAT	
	EER_{DADOS} (%)	EER_{PONTOS} (%)	EER_{DADOS} (%)	EER_{PONTOS} (%)
$\boldsymbol{X}$	16,4±0,3	–	8,5±0,1	–
$\boldsymbol{Y}$	6,9±0,3	–	7,2±0,1	–
$\boldsymbol{P}$	12,0±0,4	–	8,5±0,2	–
$\boldsymbol{C(X,Y)}$	6,3±0,6	7,0±0,4	5,7±0,3	7,8±0,1
$\boldsymbol{C(S)}$	4,3±0,2	5,4±0,6	4,2±0,3	6,8±0,2
$\boldsymbol{X^{\sim}}$	13,9 ±0,5	–	9,4±0,4	–
$\boldsymbol{Y^{\sim}}$	15,1±0,2	–	9,8±0,5	–
$\boldsymbol{P^{\sim}}$	16,6±0,2	–	11,3±0,2	–
$\boldsymbol{C(X^{\sim},Y^{\sim})}$	9,3±0,6	11,6±0,5	7,8±0,1	9,2±0,6
$\boldsymbol{C(S^{\sim})}$	7,1±0,4	9,5±0,6	6,3±0,3	8,6±0,7
$\boldsymbol{X^{\psi}}$	13,0±0,5	–	10,1±0,3	–
$\boldsymbol{Y^{\psi}}$	14,1±0,2	–	10,6±0,3	–
$\boldsymbol{P^{\psi}}$	9,6±0,1	–	10,9±0,7	–
$\boldsymbol{C(X^{\psi},Y^{\psi})}$	9,6±0,5	10,7±0,5	9,1±0,4	9,6±0,4
$\boldsymbol{C(S^{\psi})}$	6,4±0,1	7,2±0,5	6,5±0,4	8,9±0,8

Para a base MCYT-100, os resultados seguem conforme o esperado: a inclusão de novas informações melhora o desempenho dos sistemas. É interessante notar que, apesar de possuir um desmpenho superior quando consideramos cada variável isoladamente, os resultados obtidos pelo método STAT são muito semelhantes àqueles obtidos pelo método DIST sobre o conjunto de todos os dados.

Podemos perceber também que em todos os casos a fusão no espaço dos dados é superior à fusão das pontuações. Este também é um resultado esperado, pois o mapeamento não linear

gerado pelos extratores de características tem uma capacidade de representação maior que a regra da soma utilizada na fusão das pontuações.

Uma contribuição secundária desta tese também pode ser extraída através da comparação destes resultados com o estado da arte nesta área. Na Tabela 4.5 apresentamos os melhores resultados na detecção de falsificações treinadas para a base MCYT, juntamente com os resultados dos métodos que utilizamos aqui e uma breve descrição de cada um dos métodos listados.

Tabela 4.5: Estado da arte para a base MCYT (unibiométrico)

Base	Método	EER (%)
MCYT-280	d_{edit} [1]	15,9
MCYT-100	FFT [2]	12,1
MCYT-280	VQ [3]	11,8
MCYT-100	DTW [4]	9,8
MCYT-280	DTW [4]	8,9
MCYT-280	HMM [5]	8,4
MCYT-100	STFT [6]	7,0
MCYT-280	HMM [7]	5,7
MCYT-100	DIST	4,3
MCYT-100	STAT	4,2

1. Este identifica uma série de "eventos" no sinal da assinatura (voltas, pontas, arcos, mudança de velocidade) e transforma a assinatura em uma série de caracteres. Em seguida, aplica a Distância de Levenshtein, também conhecida como distância de edição, que é um tipo de distância elástica e pode ser calculada de forma semelhante à DTW aqui descrita [152].
2. Utiliza descritores de Fourier para representar características globais do sinal. Todas as características (x, y, pressão, azimute e altitude) são utizadas. De fato, os descritores de Fourier podem ser vistos como uma *wavelet* de suporte infinito, portanto só são capazes de representar características globais [153].
3. Utiliza método de quantizaçõ vetorial (VQ – *Vector Quantization*) e vizinhos mais próximos (NN – *Nearest Neighbours*) [154].

4. Este método, dos mesmos autores do método 2, foi o primeiro colocado da SVC2004, com um EER de 2,9% naquela base. Mais comentários acerca dos resultados nesta base serão tecidos nas páginas seguintes. Este mesmo algoritmo foi implementado pelos autores do método 3 e aplicado à base MCYT-280 [155] [154].
5. Utiliza um conjunto de 14 características para modelar as assinaturas. Este foi o segundo colocado da SVC2004, com um EER de 5,0% [156].
6. Neste, é utilizada a transformada de Fourier localizada (STFT – *Short Time Fourier Transform*) que, ao contrário da FFT, é capaz de fornecer informação local da assinatura através do spectrograma. Neste trabalho apenas as variáveis x e y são utilizadas, além da duração total da assinatura. Infelizmente, é difícil comparar os resultados obtidos aqui, pois foram utilizadas 15 assinaturas para gerar os modelos [157].
7. Utiliza um conjunto de 25 características dinâmicas extraídas das assinaturas. Este algoritmo também foi utilizado na SVC2004, no entanto, 10 das assinaturas de testes não foram corretamente processadas e seus resultados não foram apresentados junto com os demais. Se considerarmos tais assinaturas como erros de classificação, este passaria a ser o segundo colocado, com EER de 4,5% [158].

Não obstante a relativa simplicidade dos algoritmos que adotamos, sem qualquer procedimento de segmentação ou pré-processamento sofisticado, o desempenho foi superior ao de todos os métodos descritos na literatura. Apesar de utilizarmos apenas a base MCYT-100 que é um subconjunto da base MCYT-280, o método 4, utilizado em ambas, nos leva a crer que os resultados não seriam muito diferentes se dispuséssemos de tal conjunto de dados para teste.

O leitor mais atento deve ter percebido uma discrepância entre os resultados do estado da arte e as descrições de suas colocações na SVC2004. Os métodos 7, 5 e 4 são os melhores, nesta ordem, na base MCYT, enquanto na competição de verificação de assinaturas de 2004, a ordem foi 4, 7, 5. Nas Tabela 4.6 e Tabela 4.7 são apresentados os resultados obtidos por estes três métodos na competição SVC2004, usando apenas as variáveis X_t e Y_t, e usando todos os dados disponíveis, respectivamente

Tabela 4.6: Resultados da SVC2004 utilizando apenas as informações de posição (X e Y)

	Base de Treino (40 Usuários)			Base de Testes (60 Usuários)		
Método	EER (%)	EER_{max} (%)	EER_{min} (%)	EER (%)	EER_{max} (%)	EER_{min} (%)
DTW [4]	5,5±7,7	30,0	0,0	2,9±5,6	30,0	0,0
HMM [5]	12,0±17,7	95,0	0,0	5,9±9,2	50,0	0,0
HMM [7]	–	–	–	–	–	–

Tabela 4.7: Resultados da SVC2004 utilizando todas as informações disponíveis

	Base de Treino (40 Usuários)			Base de Testes (60 Usuários)		
Método	EER (%)	EER_{max} (%)	EER_{min} (%)	EER (%)	EER_{max} (%)	EER_{min} (%)
DTW [4]	7,0±11,8	65,0	0,0	2,9±5,7	30,0	0,0
HMM [5]	6,9±09,5	50,0	0,0	5,0±9,1	50,0	0,0
HMM [7]	7,6±12,6	60,0	0,0	4,5±9,4	65,0	0,0

Outros pontos interessantes são o fato de os algoritmos terem desempenho melhor na base de teste que naquela para qual eles foram treinados, e os valores dos desvios padrão, que, se comparados aos que obtivemos até agora, são muito elevados.

Após uma investigação cuidadosa destes problemas, concluímos que o cálculo da EER para a competição foi realizado utilizando um limiar específico para cada usuário. Tal método de avaliação, no entanto, não condiz com o que se espera de um algoritmo de biometria, já que, em uma aplicação real, não temos à disposição informações acerca das Pontuações Impostoras para cada usuário. Apesar de discordarmos do método empregado, repetimos o mesmo procedimento adotado na competição para comparar os desempenhos obtidos pelos sistemas aqui utilizados, tais resultado são apresentados na Tabela 4.8. Curiosamente, neste caso, o algoritmo DIST passa a ser superior.

De volta ao tema principal da dissertação, comparamos agora o desempenho dos algoritmos após a fusão dos dados brutos com a informação de regularidade local. Para facilitar a leitura, apresentaremos apenas alguns dos resultados. A tabela completa pode ser encontrada no Apêndice A. Apresentamos primeiramente os resultados para a base SVC2004, na Tabela 4.9

Tabela 4.8: Resultados dos algoritmos DIST e STAT seguindo o procedimento da SVC2004 (base de treino apenas)

	$C(X,Y)$			$C(S)$		
Método	EER (%)	EER_{max} (%)	EER_{min} (%)	EER (%)	EER_{max} (%)	EER_{min} (%)
DTW [4]	5,5±7,7	30,0	0,0	7,0±11,8	30,0	0,0
HMM [5]	12,0±17,7	95,0	0,0	6,9±9,5	50,0	0,0
HMM [7]	–	–	–	7,6±12,6	65,0	0,0
DIST	5,6±6,8	40,0	0,0	3,0±6,1	40,0	0,0
STAT	8,4±9,5	45,8	0,0	5,5±7,4	40,0	0,0

Tabela 4.9: Fusão da informação multifractal com os dados originais (base SVC2004)

SVC2004	**DIST**		**STAT**	
	EER_{DADOS} (%)	EER_{PONTOS} (%)	EER_{DADOS} (%)	EER_{PONTOS} (%)
Y	17,1±0,8	–	13,7±0,3	–
$C(Y,Y^{\sim})$	19,1±0,4	21,5±1,1	14,6±1,0	16,9±0,5
$C(Y,Y^{\sim},Y^{\psi})$	20,4±1,0	21,2±1,5	15,2±0,8	17,6±0,8
$C(X,Y)$	16,2±0,8	17,0±2,1	10,8±0,5	18,5±0,4
$C(X,Y,X^{\sim},Y^{\sim})$	14,9±0,6	16,9±2,3	12,6±1,0	17,1±1,0
$C(X,Y,X^{\sim},Y^{\sim},X^{\psi},Y^{\psi})$	15,8±0,4	20,7±2,7	13,0±1,1	17,7±1,4
$C(S)$	14,9±0,4	18,9±2,4	9,4±0,6	16,4±0,7
$C(S,S^{\sim})$	12,8±0,8	16,9±2,8	10,5±0,8	15,6±1,4
$C(S,S^{\sim},S^{\psi})$	15,7±0,5	21,0±3,2	12,3±0,7	15,7±1,9

Novamente, os testes na base SVC2004 apresentam resultados inesperados. A inclusão da informação multifractal para o método DIST melhora o desempenho, desde que utilizemos apenas o método da oscilação. Como já vimos, o método *wavelet* aqui utilizado não se adequa bem à forma como os dados se apresentam nesta base.

Para o método STAT, por outro lado, a inclusão da informação multifractal, independente do método utilizado, degrada o resultado apesar de este ser o método que obteve o melhor resultado com a utilização das funções de Hölder de forma independente.

No entanto, as várias questões que já levantamos acerca desta base não nos permitem tirar conclusões ainda, em especial a questão referente à forma de aquisição, que afeta diretamente os estimadores dos expoentes de Hölder. Como os métodos foram otimizados tendo a base MCYT em foco, avaliaremos agora o que acontece neste banco de dados. Os resultados podem ser observados na Tabela 4.10.

Neste caso, os resultados são bem consistentes entre as duas bases. O método estatístico mostra-se superior, na presença de poucos dados, utilizando toda a informação disponível os métodos possuem desempenhos semelhantes. Além disso, ele é capaz de aproveitar melhor a informação multifractal. O ganho médio relativo, em termos de desempenho, é de 18% utilizando a informação de apenas um estimador e de 29% utilizando ambos. Para o algoritmo DIST estes ganhos são de 8% e 18% respectivamente.

Tabela 4.10: Fusão da informação multifractal com os dados originais (base MCYT-100)

MCYT-100	DIST		STAT	
	EER_{DADOS} (%)	EER_{PONTOS} (%)	EER_{DADOS} (%)	EER_{PONTOS} (%)
$\boldsymbol{Y}$	6,9±0,3	–	7,2±0,1	–
$\boldsymbol{C(Y,Y^{\sim})}$	6,6±0,2	7,5±0,4	6,5±0,3	7,8±0,5
$\boldsymbol{C(Y,Y^{\sim},Y^{\psi})}$	6,2±0,2	7,7±0,4	5,6±0,1	7,7±0,6
$\boldsymbol{C(X,Y)}$	6,3±0,6	7,0±0,4	5,7±0,3	7,8±0,1
$\boldsymbol{C(X,Y,X^{\sim},Y^{\sim})}$	5,6±0,1	7,1±0,7	4,6±0,1	7,6±0,7
$\boldsymbol{C(X,Y,X^{\sim},Y^{\sim},X^{\psi},Y^{\psi})}$	4,8±0,3	6,6±0,9	4,3±0,1	7,2±0,8
$\boldsymbol{C(S)}$	4,3±0,2	5,4±0,6	4,2±0,3	6,8±0,2
$\boldsymbol{C(S,S^{\sim})}$	4,0±0,2	5,6±0,8	3,5±0,1	6,6±0,7
$\boldsymbol{C(S,S^{\sim},S^{\psi})}$	3,9±0,3	5,2±1,0	3,3±0,1	6,8±1,1

Podemos então considerar os casos de fusão como um sistema multibiométrico, que compararemos novamente com o estado da arte na base MCYT, mas desta vez considerando os sistemas que fazem uso da fusão de algoritmos.

Tabela 4.11: Estado da arte para a base MCYT (multibiométrico)

Base	Método	EER (%)
MCYT-280	SVM [8]	17,0
MCYT-100	*Wavelet*-DCT [9]	9,8
MCYT-280	SVM [8] + DTW [4]	7,6
MCYT-100	FFT [2] + DTW [4]	7,2
MCYT-280	VQ [3] + DTW [4]	5,4
MCYT-100	*Wavelet*-DCT [9] + DTW + HMM + GMM	5,2
MCYT-280	HMM [5] + HMM [7]	3,4
MCYT-100	DIST	4,0
MCYT-100	STAT	3,5

Os algoritmos numerados de 1 a 7 são os mesmos descritos após a Tabela 4.5. A seguir, descrevemos brevemente os dois novos algoritmos que constam apenas na Tabela 4.11, SVM (8) e *Wavelet*-DCT (9):

8. Este método utiliza um conjunto de 100 características globais extraídas a partir das séries temporais das assinaturas [159].
9. Proposto pelo mesmo autor do método 8, as características são extraídas a partir das séries temporais através de uma transformada *wavelet* unidimensional e em seguida comprimidas para um tamanho desejado através da Transformada Discreta do Cosseno (DCT – *Discrete Cosine Transform*). No mesmo artigo o autor também realiza a fusão deste método com DTW, HMM e GMM, obtendo o melhor resultado na literatura para a base MCYT-100 [160].

Os métodos propostos conseguem novamente equiparar-se ao estado da arte. Note que, em nosso caso, não estamos fazendo a fusão de métodos distintos, apenas uma transformação dos dados de entrada e utilizando o algoritmo sem qualquer ajuste. Além de poder fundir os métodos aqui descritos com outros disponíveis na literatura, a informação multifractal pode também ser utilizada como entrada nos métodos já existentes, possivelmente beneficiando-os.

Terminada esta etapa, podemos responder então a pergunta número dois:

2. ***Num contexto de fusão de informação, há um ganho de desempenho do sistema?***
   R: *Sim. Apesar de os resultados para a base SVC2004 não representarem isto, acreditamos que isto se dá devido a uma série de fatores inerentes ao método de coleta adotado para esta base, que degradam o resultado, em especial o dos expoentes de Hölder. Na base MCYT-100, por outro lado, fica claro que há um ganho na utilização da informação multifractal.*

A partir deste ponto, passaremos a utilizar apenas a base MCYT-100 e o método estatístico. De fato, não foi por acaso que tal método foi escolhido, sua utilização facilitará a próxima etapa do trabalho, onde tentaremos quantificar a informação fornecida pelos expoentes de Hölder.

Como os resultados para os expoentes estimados pelos dois métodos são bastante semelhantes e o ganho resultante da introdução do segundo estimador não é tão significativo, acreditamos que seja melhor, em uma aplicação real, utilizar apenas os expoentes estimados através do método da oscilação. O tempo de execução, em ambos os casos, cresce linearmente com o núnero de amostras: no entanto, o tempo da estimação por *wavelets*, do modo como está implementado, cresce quase 1000 vezes mais rápido que o tempo da estimação por oscilação. Para estimar a função Hölder de um sinal com 100 amostras, por exemplo, o método da oscilação leva $10ms$ enquanto o método *wavelet* demora $4{,}9s$, em experimentos realizados utilizando o Matlab v.7.9.0 (R2009b) 64bits em um computador com dois processadores de 64bits com velocidade de clock de 2.10GHz , Cache de 2MB e 4GB de Memória RAM, utilizando o sistema operacional Windows Vista Home Premium 64bits. Na Figura 4.15 apresentamos o tempo em função do número de amostras para o método da oscilação, na Figura 4.16 o mesmo é feito para o método *wavelet*.

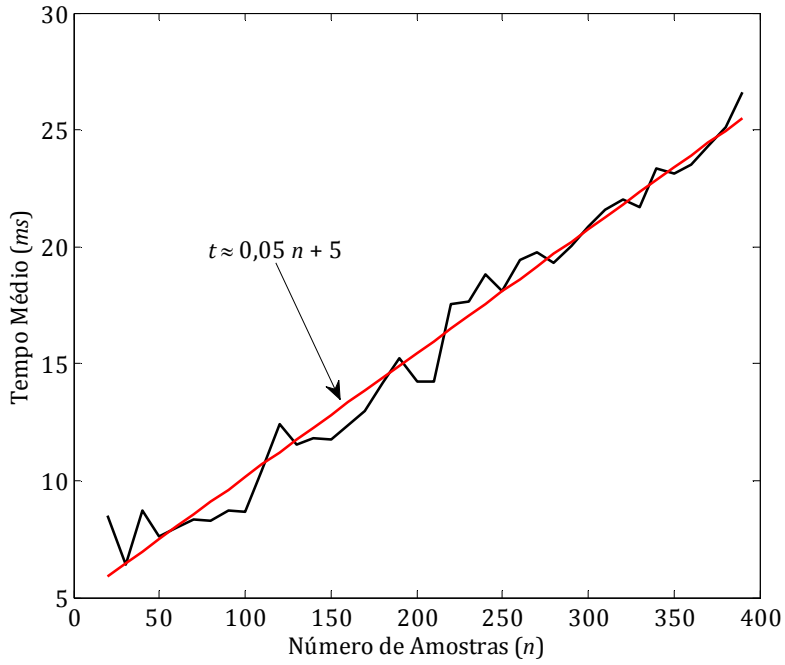


Figura 4.15: Tempo de execução em função do número de amostras para o método da oscilação

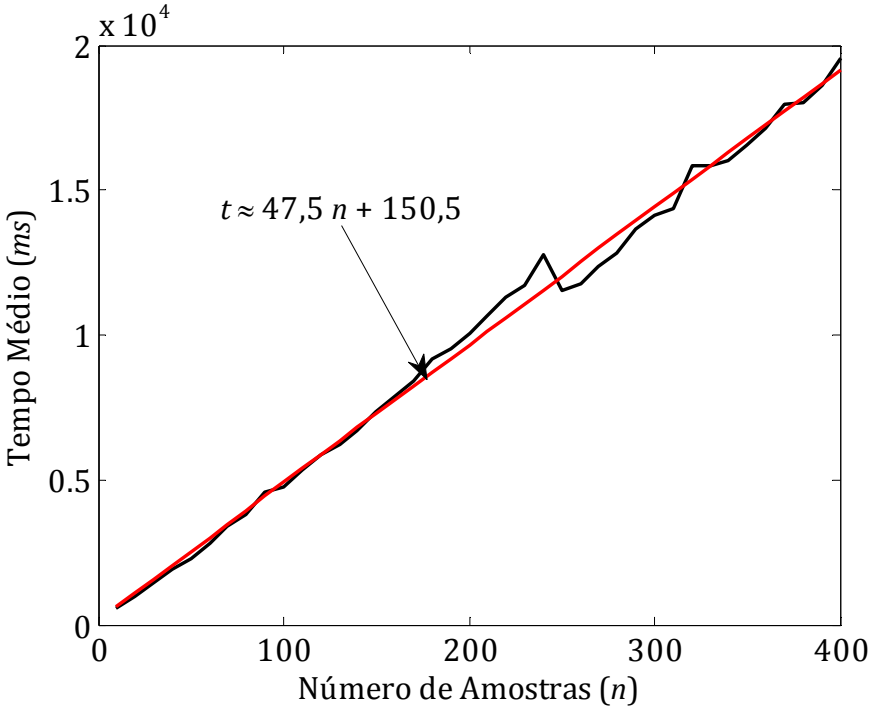


Figura 4.16: Tempo de execução em função do número de amostras para o método *wavelet*

4.5.4 Quantificação da Informação

Sob o ponto de vista da teoria da informação, o que desejamos em um problema de classificação, ao extrair características de um sinal, é obter o máximo de informação possível acerca da classe a que ele pertence. Para iniciar esta análise, partiremos da desigualdade de Fano [161].

Definição 4.2 (limite fraco de Fano) Dadas uma variável aleatória C representando as classes e uma variável aleatória X relacionada à C, a probabilidade de se realizar uma estimativa $\hat{C}$ da classe C incorretamente é limitada inferiormente por:

$$P\left(C \neq \hat{C}\right) \geq \frac{h(C|X) - 1}{\log(N-1)} = \frac{h(C) - I(C,X) - 1}{\log(N-1)} \tag{4.16}$$

onde N é o número de classes representadas por C.

Portanto, o limite inferior do erro é reduzido conforme maximizamos a Informação Mútua $I(C,Y)$ entre C e Y, que definiremos como

$$I(X,Y) = h(Y) - h(Y|X) \tag{4.17}$$

onde $h(*)$ é a medida de entropia diferencial [162] dada por

$$h(Y) = -\int_{-\infty}^{\infty} \log(f_Y(y))\, f_Y(y) dy. \tag{4.18}$$

A Figura 4.17 ajuda a entender este conceito. Supomos a existência de uma distribuição subjacente às classes, dos quais os dados brutos X são realizações com pdf condicionada à classe que os gera. Em seguida, um extrator de características $g(*,\Theta)$ é aplicado ao sinal, que nos dá as características Y a partir das quais realizaremos nossa estimação $\hat{C}$ da classe à qual X pertence. Portanto, nosso objetivo é encontrar os parâmetros Θ da função g que maximizam a informação mútua entre C e Y.

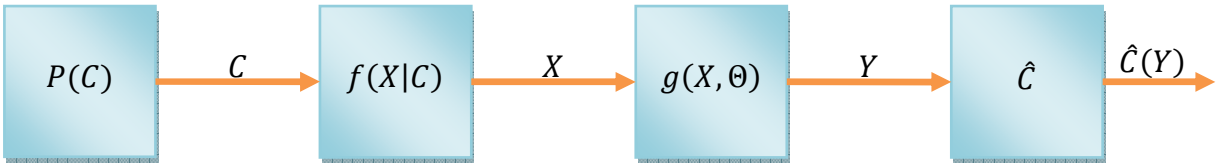


Figura 4.17: Problema da classificação em um contexto bayesiano

Perceba que esta descrição que acabamos de fazer é bem semelhante àquela que foi utilizada para a explanação do método da mistura de gaussianas. Esta discussão, até este ponto, pode ser visto como uma demonstração alternativa do princípio da maximização de informação de Linsker [163].

Há diversos métodos na literatura para alcançar este objetivo, sendo o InfoMax, de Bell e Sejnowski [164], provavelmente o mais conhecido. No entanto, para nosso caso especificamente, os métodos descritos por Viola et al. [165] e em uma série de artigos de Fisher e Príncipe [166] [167] [168] [169] são particularmente interessantes. Estas técnicas utilizam o método Parzen para realizar uma estimativa, não paramétrica, da pdf $f_Y(y)$ e comparam-na à distribuição uniforme, que tem entropia máxima se o domínio de y é um conjunto limitado do $\mathbb{R}^n$.

Em seguida, uma função de custo é formulada e uma regra de atualização para Θ é deduzida, que é utilizada pelo sistema através do algoritmo de retro-propagação do erro. Pode-se mostrar que a minimização deste erro é equivalente à maximização da verossimilhança [170] [171]. Se re-interpretarmos a Figura 4.17 como um problema de decodificação fica fácil perceber esta equivalência. Observe a Figura 4.18.

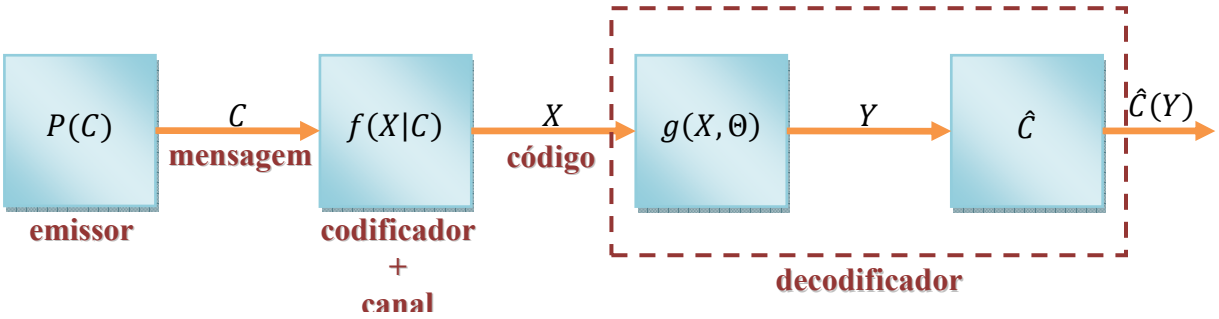


Figura 4.18: Problema da classificação como um problema de decodificação

Sabe-se da teoria da codificação, que o decodificador ótimo é o de máxima verossimilhança. O algoritmo Parzen, que modificamos para um contexto paramétrico, também pode ser visto desta forma. De fato, em [172], é mostrado que a estimativa do desvio padrão σ_0

dos núcleos gaussianos através da validação cruzada do tipo todos-menos-um (*leave-one-out*) é melhor que o valor σ obtido pelo método de descida do gradiente.

Se analizarmos a Figura 4.14 novamente, o método Parzen aqui aplicado realiza um mapeamento de $\mathbb{R}^n$ em $\mathbb{R}^+$, onde os pontos mais distantes daqueles utilizados no treinamento são comprimidos próximo ao zero e as regiões de alta concentração são expandidas, realizando intrinsecamente um processo semelhante a uma equalização, conforme desejado no algoritmo de Fisher e Principe, citado anteriormente.

Tendo explicado rapidamente como o método de Parzen pode ser visto como um bom estimador das estatísticas informativas, podemos voltar ao ponto inicial desta seção. A desigualdade de Fano, portanto, pode nos dar um limitante inferior para o erro ou para a informação mútua entre a classe e as pontuações, ou ainda, de forma inversa, um limite superior para a incerteza sobre a classe dada a pontuação. No entanto, o termo $\log(N-1)$ no denominador torna esta desigualdade impossível de ser utilizada para o caso binário, que é o que temos em mãos. Utilizaremos então o limite forte de Fano [161], que no caso binário reduz-se apenas a

$$h(C|X) \leq h(E) \tag{4.19}$$

onde E é uma variável aleatória binária que indica a ocorrência de erro de classificação.

Por outro lado, o limite de Hellman-Raviv [173], dado por

$$h(C|X) \geq 2P\left(C \neq \hat{C}\right), \tag{4.20}$$

nos dá um limite superior para o erro e informação mútua, e um limitante inferior para a entropia da classe dada a observação. Na Figura 4.19 apresentamos estes dois limites para o caso binário.

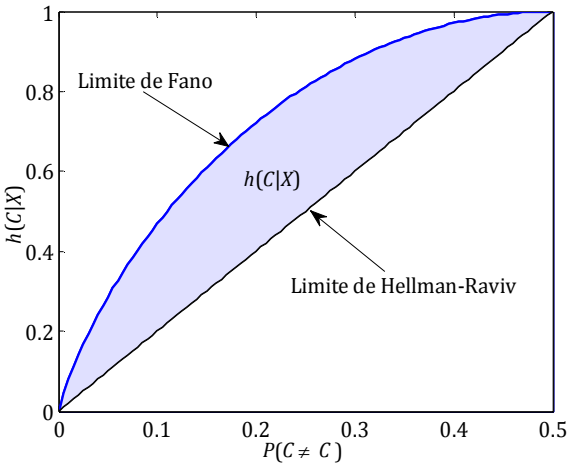


Figura 4.19: Limite de Fano e Limite de Hellman-Raviv

Conforme explicado nesta seção, o método estatístico adotado, empregando janelas de Parzen, aproxima-se do estimador ótimo, segundo [171]. De fato, comparando com os resultados obtidos na literatura, ele chega no limite do desempenho dos demais sistemas. Podemos então, a partir dos resultados obtidos, estimar quanta incerteza os expoentes de Hölder removem acerca da classificação.

Inicialmente, calcularemos os limites de Fano, $h(C|\rho)_{sup}$, e de Hellman-Raviv, $h(C|\rho)_{inf}$, para cada variável isoladamente, onde ρ aqui são as características que utilizamos para estimar a autenticidade do usuário, ou seja, as pontuações. Vamos ainda supor que, *a priori*, não possuímos nenhuma informação sobre as classes e que elas são equiprováveis, ou seja $h(C) = 1$ bit. Os resultados obtidos para variáveis isoladas encontram-se na Tabela 4.12, enquanto na Tabela 4.13 apresentamos as estimativas para combinações de variáveis.

A partir da comparação dos valores obtidos nas Tabela 4.12 e Tabela 4.13 podemos estimar quanta incerteza acerca da classe é removida por cada variável em relação a outra. Por exemplo, na combinação $C(X, X^{\sim})$, podemos estimar quanta incerteza X remove em relação ao uso de $X^{\sim}$ sozinho e vice-versa. Na Tabela 4.14 apresentamos o limite superior, calculado através do limite de Fano, desta remoção de incerteza (ou ganho de informação), promovido por cada variável. As linhas indicam quem está sendo introduzido e as colunas representam com quem ele está sendo combinado.

Tabela 4.12: Estimação da incerteza sobre a classe dada a pontuação. Variáveis consideradas isoladamente

	$h(C\|\rho)_{inf}$ (bits)	$h(C\|\rho)_{sup}$ (bits)
X	0,170	0,420
Y	0,144	0,373
P	0,170	0,420
$X^{\sim}$	0,188	0,450
$Y^{\sim}$	0,196	0,463
$P^{\sim}$	0,226	0,509

Tabela 4.13: Estimação da incerteza sobre a classe dada a pontuação. Combinações de variáveis.

	$h(C\|\rho)_{inf}$ (bits)	$h(C\|\rho)_{sup}$ (bits)
$C(X, X^{\sim})$	0,144	0,373
$C(Y, Y^{\sim})$	0,130	0,347
$C(P, P^{\sim})$	0,158	0,399
$C(X, Y)$	0,114	0,315
$C(X^{\sim}, Y^{\sim})$	0,146	0,377
$C(X, Y^{\sim})$	0,144	0,374
$C(X^{\sim}, Y)$	0,119	0,325
$C(X, P)$	0,116	0,318
$C(X^{\sim}, P^{\sim})$	0,143	0,371
$C(X, P^{\sim})$	0,156	0,396
$C(X^{\sim}, P)$	0,123	0,333
$C(Y, P)$	0,094	0,273
$C(Y^{\sim}, P^{\sim})$	0,153	0,390
$C(Y, P^{\sim})$	0,126	0,340
$C(Y^{\sim}, P)$	0,126	0,339

Tabela 4.14: Remoção de incerteza em bits promovida pela fusão de dados

	X	Y	P	$X^{\sim}$	$Y^{\sim}$	$P^{\sim}$
X	–	0,058	0,102	0,077	0,081	0,113
Y	0,105	–	0,147	0,125	0,116	0,169
P	0,102	0,055	–	0,117	0,124	0,110
$X^{\sim}$	0,047	0,048	0,087	–	0,086	0,138
$Y^{\sim}$	0,046	0,026	0,081	0,073	–	0,119
$P^{\sim}$	0,024	0,033	0,021	0,079	0,073	–

Observa-se então que a contribuição dos expoentes de Hölder é, em geral, significativamente inferior à dos dados brutos. Se compararmos a contribuição de Y e $Y^{\sim}$ com relação a X, por exemplo, percebemos que utilizando o dado bruto, Y, há uma redução de

incerteza duas vezes maior que utilizando $Y^{\sim}$. Este já era um fenômeno esperado, uma vez que as funções Hölder são apenas transformações dos dados brutos.

De modo resumido, temos que a redução média de incerteza causada pelas variáveis originais em relação às funções Hölder é de 0,115bits enquanto a redução provocada pelas funções Hölder em relação às primeiras é de apenas 0,046bits, sendo $P^{\sim}$ a menos significativa das informações multifractais e Y a mais significativa das informações brutas.

Vale relembrar que todas as estimativas foram realizadas a partir das pontuações fornecidas pelo algoritmo STAT sobre a base MCYT-100. Supondo que todos os estimadores utilizados sejam ótimos, podemos dizer que a informação e a redução de incerteza fornecidas pela análise multifractal são limitadas aos valores que obtivemos aqui através do limite de Fano.

Finalmente, podemos responder a terceira questão proposta neste trabalho:

3. ***Quanta informação é acrescentada pela análise multifractal?***

   R: *Supondo que os estimadores utilizados sejam ótimos e que as ocorrências das classes "genuíno" e "impostor" sejam equiprováveis, a informação multifractal é capaz de proporcioanar uma redução de incerteza de no máximo 0,046bits.*

Apesar do ganho proporcionado ser baixo se o compararmos à utilização dos dados brutos, o uso dos expoentes de Hölder é viável, pois, como já vimos, o método de estimação adotado é bastante rápido; e contribui para o desempenho do sistema como um todo. Portanto, acreditamos que não há motivos para descartar a utilização desta informação extra. Além disso, os métodos adotados aqui foram bastante simples e genéricos, sem ajuste de parâmetros específicos à aplicação, o que ainda pode dar margem a melhorias na forma como esta informação é utilizada e, consequentemente, aumentar a contruibuição da regularidade local para o desempenho do sistema.

Capítulo 5

Conclusões

Os resultados obtidos ao longo deste trabalho nos levam a crer que as informações multifractais, amplamente usada para a nálise de séries temporais em diversos ramos da ciência, podem contribuir também para o desempenho dos sitemas de verificação de assinaturas dinâmicas. De fato, apesar de não termos extendido o estudo a outros tipos de sinais biométricos, acreditamos que o mesmo princípio pode ser utilizado no contexto das outras biometrias.

A idéia central aqui foi utilizar a informação de regularidade local como característica complementar àquelas que já são utilizadas rotineiramente nesta área de pesquisa. Os resultados da Seção 4.5.4 mostram que há bastante redundância entre a informação multifractal e a informação dos dados brutos, o que já era esperado, pois se trata de uma transformação destes. Apesar disso, conseguimos um ganho relativo de até 30%, em termos de EER, em um sistema cuja performance já se encontra no estado da arte.

Como este foi o primeiro trabalho a utilizar a informação multifractal desta forma, não possuímos referências para comparar o desempenho dos expoentes de Hölder, especificamente. No entanto, comparando com trabalhos de outras áreas, percebemos que as conclusões obtidas são semelhantes: a informação (multi)fractal ajuda a classificar, mas não é boa o suficiente para ser usada no lugar de outras técnicas já conhecidas.

Além disso, optamos por utilizar métodos relativamente simples de comparação, que não necessitam qualquer ajuste de parâmetros por parte do usuários, de modo que ainda há a possibilidade de melhorias, seja no algoritmo como um todo, seja na forma como a informação fractal é explorada. Nas sessões seguintes indicamos algumas possibilidades alternativas de utilização dos expoentes de Hölder, que podem ser vistas como sugestões para trabalhos futuros, e explicitamos de forma objetiva as contribuições desta dissertação.

5.1 Aplicações Alternativas dos Expoentes de Hölder na Verificação de Assinaturas

Além da utilização como característica discriminativa, vislumbramos o uso da informação multi-fractal de duas outras formas, as quais descreveremos aqui brevemente.

5.1.1 Qualidade da Assinatura

Como explicamos ao longo deste trabalho, as assinaturas, apesar de possuírem algumas características multi-fractais, são sinais bastante regulares. A informação de regularidade global (i.e. o parâmetro de Hurst) é em geral muito alto para as assinaturas. Deste modo, acreditamos que seja possível utilizar o parâmetro de Hurst para descartar rabiscos aleatórios ou assinaturas de qualidade muito ruim.

Realizamos experimentos iniciais utilizando movimento Browniano multifracionário e o *mouse* para gerar os rabiscos e avaliamos a diferença entre o parâmetro de Hurst e a média dos expoentes de Hölder. Todos os sinais gerados automaticamente através do movimento Browniano foram corretamente descartados. Para os "rabiscos" gerados com o *mouse*, nem todas foram descartadas corretamente, mas deixaremos um estudo mais aprofundado desta aplicação para trabalhos futuros.

5.1.2 Amostragem Adaptativa

A informação de regularidade local está intimamente ligada aos modelos termodinâmicos. De forma intuitiva, podemos entender a regularidade local como a "temperatura" da função naquele ponto, ou seja, o grau de agitação (variação) dela. Deste modo, acreditamos que seja possível utilizar a informação multifractal para a seleção adaptativa de amostras, onde a taxa de amostragem deve ser inversamente proporcional ao valor do expoente naquela região.

Na forma como implementamos o modelo estatístico, por exemplo, selecionamos apenas um de cada dois pontos da amostra, para reduzir a carga computacional. Utilizando a informação dos expoentes de Hölder poderíamos focar a obtenção de amostras em regiões de maior "temperatura". No entanto, não realizamos nenhum experimento neste sentido ainda.

Uma idéia bastante semelhante, utilizando entropia, é utilizada na seleção de janelas de processamento para reconhecimento de fala em [174].

5.2 Resumo

Aqui resumimos, de forma objetiva, os resultados e contribuições deste trabalho.

- Análise multifractal dos sinais de assinatura;
- Proposta de um método, baseado em DTW, com resultados superiores àqueles encontrados na literatura;
- Modificação do método proposto por [146] para penalização explícita baseada no tempo total de assinatura, também melhorando seu desempenho;
- Análise da viabilidade e eficiência da informação multifractal como característica discriminativa, respondendo as três perguntas propostas:
  - ***É possível utilizar informação de regularidade local para classificar assinaturas?***

  R: *Sim, embora os resultados não sejam tão bons quanto aqueles obtidos através dos dados originais.*
  - ***Num contexto de fusão de informação, há um ganho de desempenho do sistema?***

  R: *Sim. Apesar dos resultados para a base SVC2004 não representarem isto, acreditamos que isto se dá devido a uma série de fatores inerentes ao método de coleta adotado para esta base, que degradam o resultado, em especial o dos expoentes de Hölder. Na base MCYT-100, por outro lado, fica claro que há um ganho na utilização da informação multifractal.*
  - ***Quanta informação é acrescentada pela análise multifractal?***

  R: *Supondo que os estimadores utilizados sejam ótimos e que as ocorrências das classes "genuíno" e "impostor" sejam equiprováveis, a informação multifractal é capaz de proporcioanar uma redução de incerteza de no máximo 0,046bits.*

Referências Bibliográficas

[1]. Jain, A. K., Flynn, P e Ross, A. *Handbook of Biometrics.* s.l. : Springer-Verlag, 2007.

[2]. International Biometric Group. *Biometrics Market and Industry Report 2009-2014.* 2008.

[3]. Ross, A. e Jain, A. K. Human Recognition Using Biometrics: an Overview. *Annals of Telecommunications.* Janeiro de 2007, Vol. 62, 1-2, pp. 11-35. 2007.

[4]. Impedovo, D. e Pirlo, G. Automatic Signature Verification: The State of the Art. *IEEE Transactions on Systems, Man and Cybernetics - Part C: Applications and Reviews.* Setembro de 2008, Vol. 38, 5, pp. 609-635. 2008.

[5]. Mandelbrot, B. B. *The Fractal Geometry of Nature.* New York : W. H. Freeman & Co., 1977.

[6]. Bianchi, G. R., Vieira, F. H. e Lee, L. L. *A novel network traffic predictor based on multifractal traffic characteristic*. IEEE Global Telecommunications Conference. Vol. 6, pp. 680-684. 2004.

[7]. Mandelbrot, B. B. e Hudson, R. L. *The (Mis)Behavior of Markets.* s.l. : Basic Books, 2004.

[8]. Nyikos, L., Balázs, L. e Schiller, R. Fractal Analysis of Artistic Images: from Cubism to Fractalism. *Fractals.* Vol. 2, 1, pp. 143-152. 1994.

[9]. Joshi, A., et al. Arrhytmia Classification Using Local Hölder Exponents and Support Vector Machine. *Lecture Notes in Computer Science: Pattern Recognition and Machine Intelligence.* Dezembro de 2005, Vol. 3776, pp. 242-247.

[10]. Grasman, J., et al. The Multifractal Structure of Arterial Trees. *Journal of Theoretical Biology.* Janeiro de 2003, Vol. 220, 1, pp. 75-82.

[11]. Scafetta, N., Griffin, L. e West, B. J. Hölder Exponent Spectra for Human Gait. *Physica A: Statistical Mechanics and its Applications.* 2003, Vol. 328, pp. 561-583.

[12]. Jain, A. K., Bolle, R. e Pankanti, S. *Biometrics: Personal Identification in Networked Society.* s.l. : Kluwer Acadamic Publishers, 1999.

[13]. Bolle, R., et al. *Guide to Biometrics.* s.l. : Springer, 2003.

[14]. Jain, A. K. e Ross, A. An Introduction to Biometric Recognition. *IEEE Transactions on Circuits and System for Video Technology, Special Issue on Image and Video Based Biometrics.* Janeiro de 2004, Vol. 14, 1, pp. 4-20.

[15]. Jain, A. K. Biometrics. *The World Book Encyclopedia.* 2007.

[16]. Jain, A. K., Hong, L. e Pankanti, S. Biometric Identification. *Communications of the ACM.* Fevereiro de 2000, Vol. 43, 2, pp. 91-98.

[17]. Vielhauer, C. e Dittmann, J. *Biometrics for User Authentication: Encyclopedia of Multimidia.* [ed.] B. Furth. s.l. : Springer, 2006.

[18]. *Biometrics research agenda: Report of the NSF workshop.* Rood, E. e Jain, A. K. Morgantown : s.n., 2003. Workshop for a Biometric Research Agenda.

[19]. Boyer, K. W., Govindaraju, V. e Ratha, N. K. Special Issue on Recent Advances in Biometric Systems. *IEEE Transactions on Systems, Man and Cybernetics.* Outubro de 2007, Vol. 37, 5, pp. 1091-1095.

[20]. Prabhakar, S., et al. Special Issue on Biometrics: Progress and Directions. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* Abril de 2007, Vol. 29, 4, pp. 513-516.

[21]. Zhang, K., et al. Special Issue on Biometric Systems. *IEEE Transactions on Systems, Man and Cybernetics.* Agosto de 2005, Vol. 35, 3, pp. 273-275.

[22]. Frischholz, R. W. e Dieckmann, U. BioID: A Multimodal Biometric Identification System. *IEEE Computer Science.* 2000, Vol. 33, 2.

[23]. Davies, S. Touching Big Brother: How Biometric Technology Will Fuse Flesh and Machine. *Information Technology and People.* 1994, Vol. 7, 4.

[24]. Kenny, S. e Borking, J. J. The Value of Privacy Engineering. *The Journal of Information, Law and Technology.* 2002, Vol. 7, 1.

[25]. Rejman-Greene, M. Privacy Issues in the Application of Biometrics: A European Perspective. [ed.] J. L. Wayman, et al. *Biometric Systems: Technology, Design and Performance Evaluation.* s.l. : Springer, 2005, pp. 335-339.

[26]. Duda, R. O., Hart, P. E. e Stork, D. G. *Pattern Classification.* s.l. : Willey-Interscience, 2000.

[27]. Theodoridis, S. e Koutroumbas, K. *Pattern Recognition.* s.l. : Academic Press, 2006.

[28]. *Biometrics: A Grand Challenge.* Jain, A. K., et al. Cambridge : s.n., 2004. Proceedings of International Conference on Pattern Recognition. Vol. 2, pp. 935-942.

[29]. Vielhauer, C. A Behavioural Biometrics. *Public Service Review: European Union.* 2005, Vol. 20, 9, pp. 113-115.

[30]. Wijesoma, W. S., et al. Online Handwritten Signature Verification for Electronic Commerce over the Internet. [ed.] N. Zhong, et al. *Web Intelligence: Research and Development, Lecture Notes in Artificial Intelligence 2198.* Berlin : Springer-Verlag, 2001, pp. 227-236.

[31]. Veeramacheneni, K., Osadciw, L. A. e Varshney, P. K. An Adaptive Multimodal Biometric Management Algorithm. *IEEE Transactions on Systems, Man and Cybernetics.* Agosto de 2005, Vol. 35, 3, pp. 344-356.

[32]. *Signature Verification Using Fractal Transformation.* Huang, K. e Yan, H. Barcelona : s.n., 2000. Proceedings of the 15th International Conference on Pattern Recognition. Vol. 2, pp. 851-854.

[33]. Egan, J. *Signal Detection Theory and ROC Analysis.* s.l. : Academic Press, 1975.

[34]. Neyman, J. e Pearson, E. S. On the Problem of the Most Efficient Test of Statistical Hypotheses. *Philosophy Transactions of the Royal Society of London, Series A.* 1933, Vol. 231, pp. 289-337.

[35]. *Sheeps, Goats, Lambs and Wolves: A Statistical Analysis of Speaker Performance in the NIST 1998 Speaker Reocgnition Evaluation.* Doddington, G., et al. Sydney : s.n., 1998. Processdings of the 5th International Conference on Spolenn Language Processings.

[36]. Pankanti, S., Prabhakar, S. e Jain, A. K. On the Individuality of Fingerprints. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 2002, Vol. 24, 8, pp. 1010-1025.

[37]. *How Flexible is the Human Voice? A Case Study of Mimicry.* Eriksson, A. e Wretling, P. Rhodes : s.n., 1997. Proceedings of the European Conference on Speech Technology. pp. 1043-1046.

[38]. Harrison, W. R. *Suspect Documents, their Scientific Examination.* s.l. : Nelson-Hall Publishers, 1981.

[39]. *Impact of Artificial Gummy Fingers on Fingerprint Systems.* Matsumoto, T., et al. San Jose : s.n., 2002. Optical Security and Counterfeit Deterrence Techniques IV, Proceedings of SPIE. Vol. 4677, pp. 275-289.

[40]. *Biometrical Fingerprint Recognition: Don't Get Your Fingers Burned.* Putte, T. e Keuning, J. 2000. Proceedings of IFIP TC8/WG8. 8 Fourth Working COnference on Smart Card Research and Advanced Applications. pp. 289-303.

[41]. *An Analysis of Minutiae Matching Strength.* Ratha, N. K., Connell, J. H. e Bolle, R. M. Halmstad : s.n., 2001. Proceedings of 3rd International Conference on Audio and Video Based Biometric Person Authentication. pp. 223-228.

[42]. Ross, A., Nandakumar, K. e Jain, A. K. *Handbook of Multibiometrics.* s.l. : Springer, 2006.

[43]. Jain, A. K. e Ross, A. Multibiometric Systems. *Communications of the ACM, Special Issue on Multimodal* Interfaces. Janeiro de 2004, Vol. 47, 1, pp. 34-40.

[44]. *Can Multibiometrics Improve Performance?* Hong, L., Jain, A. K. e Pankanti, S. New Jersey : s.n., 1999. Proceedings of IEEE Workshop on Automatic Identification Advanced Technologies. pp. 59-64.

[45]. Palmer, T. N. Predicting Uncertainty in Forecasts of Weather Climate. *Reports on Progress in Physics.* 2000, Vol. 63, pp. 71-116.

[46]. Blum, R. S. e Liu, Z., [ed.]. *Multi-Sensor Image Fusion and Its Applications.* s.l. : CRC Press, 2006.

[47]. Abidi, M. A. e Gonzalez, R. C. *Data Fusion in Robotics and Machine Intelligence.* s.l. : Academic Press, 1992.

[48]. Kuncheva, L. I. *Combining Pattern Classifiers - Methods and Algorithms.* s.l. : Wiley, 2004.

[49]. Kittler, J., et al. On Combining Classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* Março de 1998, Vol. 20, 3, pp. 226-239.

[50]. Xu, L., Krzyzak, A. e Suen, C. Y. Methods for Combining Multiple Classifiers and their Applications to Handwritting Recognition. *IEEE Transactions no Systems, Man and Cybernetics.* 1992, Vol. 22, 3, pp. 418-435.

[51]. *Muticlassifier Systems: Back to the Future.* Ghosh, J. Cagliari : s.n., 2002. Proceedings of Third International Workshop on Multiple Classifier Systems. pp. 1-15.

[52]. Chibelushi, C. C., Deravi, F. e Mason, J. S. Voice and Facial Image Integration for Speaker Recognition. [ed.] R. I. Damper, W. Hall e J. W. RIchards. *Multimedia Technologies and Future Applications.* Londres : Pentech Press, 1994, pp. 155-161.

[53]. Brunelli, R. e Falavigna, D. Person Identification Using Multiple Cues. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* Outubro de 1995, Vol. 17, 10, pp. 955-966.

[54]. *Finding Optimal Views for 3D Face Shape Modeling.* Lee, J., et al. Seul : s.n., 2004. Proceedings of the IEEE INternational Conference on Automatic Face and Gesture Recognition. pp. 31-36.

[55]. Kong, A., et al. Recent Advances in VIsual and Infrared Face Recognition - A Review. *Computer Vision and Image Understanding.* Janeiro de 2005, Vol. 97, 1, pp. 103-135.

[56]. Chen, X., Flynn, P. J. e Bowyer, K. W. IR and Visible Light Face Recognition. *Computer Vision and Image Understanding.* Setembro de 2005, Vol. 99, 3, pp. 332-358.

[57]. Socolinsky, D. A., Selinger, A. e Neuheisel, J. D. Face Recognition with Visible and Thermal Infrared Imagery. *Computer Vision and Image Understanding.* Julho-Agosto de 2003, Vol. 91, 1-2, pp. 72-114.

[58]. Marcialis, G. L. e Roli, F. Fingerprint Verification by Fusion of Optical and Capacitive Sensors. *Pattern Recognition Letters.* Agosto de 2004, Vol. 25, 11, pp. 1315-1322.

[59]. Ross, A., Nandakumar, K. e Jain, A. K. A Hybrid Fingerprint Matcher. *Pattern Recognition.* Julho de 2003, Vol. 36, 7, pp. 1661-1673.

[60]. Chang, K. I., Bowyer, K. W. e Flynn, P. J. An Evaluation of Multimodal 2D+3D Face Biometrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* Abril de 2005, Vol. 27, 4, pp. 619-624.

[61]. *An Integrated Dual Factro Authenticator Based on the Face Data and Tokenised Random Number.* Jin, A. T. B, Ling, D. N. C. e Goh, A. Hong Kong : s.n., 2004. 1st International Conference on Biometric Authentication. pp. 117-123.

[62]. *Integrating Faces, Fingerprints and Soft Biometric Traits for User Recognition.* Jain, A. K., et al. Praga : Springer, 2004. Proceedings of ECCV International Workshop on Biometric Authentication. Vol. LNCS 3087, pp. 259-269.

[63]. Sanderson, C. e Paliwal, K. K. *Information Fusion and Person Verification Using Speech and Face Information.* s.l. : IDIAP, 2002. Research Paper IDIAP-RR 02-33.

[64]. Nalwa, V. S. Automatic On-Line Signature Verification. *Proceedings of the IEEE.* Fevereiro de 1997, Vol. 85, 2, pp. 215-239.

[65]. Lee, L., Berger, T. e Aviczer, E. Reliable On-Line Human Signature Verification Systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* Junho de 1996, Vol. 18, 6, pp. 643-647.

[66]. Fairhurst, M. C. Signature Verification Revisited: Promoting Practical Exploitation of Biometric Technology. *IEE Electronics and Communication Engineering Journal.* Dezembro de 1997, Vol. 9, 6, pp. 273-280.

[67]. Fairhurst, M. C. e Kaplani, E. Perceptual Analysis of Handwritten Signatures for Biometric Authentication. *IEE Proceedings on Vision, Image and Signal Processing.* Dezembro de 2003, Vol. 150, 6, pp. 389-394.

[68]. Leclerc, F. e Plamondon, R. Automatic Signature Verification: The State of the Art (1989-1993). *International Journal on Pattern Recognition and Artificial Intelligence.* Junho de 1994, Vol. 8, 3, pp. 643-660.

[69]. Pirlo, G. Algorithms for Signature Verification. [ed.] S. Impedovo. *Proceedings of the NATO-ASI Series: Fundamentals in Handwritting Recognition.* 1994, pp. 433-454.

[70]. Nanavati, S., Thieme, M. e Nanavati, R. *Biometrics: Identity Verification in a Networked World.* Nova Iorque : Wiley, 2002.

[71]. Plamondon, R. e Srihari, S. N. On-Line and Off-Line Handwritting Recognition: A Comprehensive Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* Janeiro de 2000, Vol. 22, 1, pp. 63-84.

[72]. Plamondon, R. A Kinematic Theory of Rapid Human Movements: Part I: Movement Representation and Generation. *Biological Cybernetics.* 1995, Vol. 72, 4, pp. 295-307.

[73]. Plamondon, R. A Kinematic Theory of Rapid Human Movement: Part II: Movement Time and Control. *Biological Cybernetics.* 1995, Vol. 72, 4, pp. 309-320.

[74]. *The Origin of 2/3 Power Law.* Plamondon, R. 1997. Proceedings of the 8th International Conference of the Graphonomics Society. pp. 17-20.

[75]. Plamondon, R. A Kinematic Theory of Rapid Human Movement: Part III: Kinetics Outcome. *Biological Cybernetics.* Janeiro de 1997.

[76]. Doermann, D. S. e Rosenfeld, A. Recovery of Temporal Information from Static Images of Handwritting. *International Journal of Computer Vision.* 1995, Vol. 15, pp. 143-164.

[77]. *Ink Texture Analysis for Writer Identification.* Franke, K., Bünnemeyer, O. e Sy, T. 2002. Proceedings of the 8th International Workshop on Frontiers of Handwritting Recognition. pp. 268-273.

[78]. *Ink Deposition Analysis Using Temporal (on-line) Data.* Franke, K. e Rose, S. La Baule : s.n., 2006. Proceedings of the 10th Workshop on Frontiers of Handwritting Recognition. pp. 447-453.

[79]. *Ink Deposition Model, the Relation of Writing and Ink Deposition Processes.* Franke, K. e Rose, S. Kichijoji : s.n., 2004. Proceedings of the 9th International Workshop on Frontiers of Handwritting Recognition. pp. 173-178.

[80]. Plamondon, R., [ed.]. *Progress in Automatic Signature Verification.* s.l. : World Scientific, 1994.

[81]. Miller, B. Vital Signs of Identity. *IEEE Spectrum.* Fevereiro de 1994, Vol. 31, 2, pp. 22-30.

[82]. Newham, E. Survey: Signature Verification Technologies. *Signature Verification.* Amsterdam : Elsevier, 2000, pp. 8-10.

[83]. *Mouse Based Signature Verification for Internet Based Transactions.* Lei, H., Palla, S. e Govindaraju, V. San Jose : s.n., 2005. Proceedings of the SPIE Symposium on Electronic Imaging Science & Technology: Electronic Imaging Visualization. Vol. 5673, pp. 153-160.

[84]. ANSI. *Information Technology - Biometric Data Interchange Formats - Signatures/Sign Data.* s.l. : ANSI Standard ANSI INCITS 392-2005, 2005.

[85]. ISO. *Information Technology - Biometric Data Interchange Formats - Part 7: Signatures/Sign Time Series Data.* s.l. : ISO Standard ISO/IEC FCD 19794-7, 2006.

[86]. ISO. *Information Technology - Biometric Data Interchange Formats - Part 11: Signature/Sign Processed Dynamic Data.* s.l. : ISO Standard ISO/IEC WD 19794-11, 2007.

[87]. Peitgen, H. O., Jurgens, H. e Saupe, D. *Chaos and Fractals, New Frontiers of Science.* s.l. : Springer-Verlag, 1994.

[88]. Riedi, R. H. *Introduction to Multifractals.* Departamento de ECE, Universidade de Rice. Houston : s.n., 1999.

[89]. Falconer, K. *Fractal Geometry: Mathematical Foundations and Applications.* Nova York : John Wiley and Sons, 1990.

[90]. Mandelbrot, B. B. How Long is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension. *Science.* Maio de 1967, Vol. 156, 3775, pp. 636-638.

[91]. Thanki, S. G. *Classification of Galaxies using Fractal Dimensions.* Departamento de Física, Universidade de Nevada. 1999. Tese de Mestrado.

[92]. Hirsch, M. W., Smale, S. e Devaney, R. *Differential Equations, dynamical systems, and an introduction to chaos.* s.l. : Academic Press, 2003.

[93]. Park, K. e Willinger, W. *Self-Similar Network Traffic and Performance Evaluation.* New York : John Wiley and Sons, 2000.

[94]. Beran, J. *Statistics for Long-Memory Process.* New York : Chapman & Hall, 1994.

[95]. Willinger, W., Sherman, R. e Wilson, D. Self-Similarity Through High-Variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level. *IEEE/ACM Transactions on Networking.* Fevereiro de 1997, Vol. 5, 1.

[96]. Taqqu, M., Teverovsky, V. e Willinger, W. Is Network Traffic Self-Similar or Multifractal? *Fractals.* 1996.

[97]. Taqqu, M. S. e Teverovsky, V. Robustness of Whittle-type Estimator for Time Series with Long-Range Dependence. *Stochastic Models.* 1997.

[98]. Leland, W. E., Willinger, W. e Wilson, D. V. On the Self-Similar Nature of Ethernet Traffic (Extended Version). *IEEE.* 1994, pp. 1-15.

[99]. Taqqu, M. S. e Teverovsky, V. On Estimating the Intensity of Long-Range Dependence in Finite and Infinite Variance Time Series. *A Practical Guide to Heavy Tails: Statistical Techniques and Applications.* s.l. : Birkhauser, 1998, pp. 177-217.

[100]. *On the Convergence of MMPP and Fractional ARIMA Process with Long-Range Dependence to Fractional Brownian Motion.* Sikdar, B. e Vastola, K. S. New Jersey : s.n., 2000. Proceedings of the 34th CISS.

[101]. Véhel, J. L., Lutton, E. e Tricot, C. *Fractals in Engineering.* Londres : INRIA, Springer-Verlag, 1997.

[102]. Véhel, J. L. e Riedi, R. *Fractional Brownian Motion and Data Traffic Modeling: The Other End of the Spectrum.* Rocquencourt : INRIA, 1997.

[103]. Billingsley, P. *Convergence of Probability Measures.* New York : John Wiley & Sons, 1979.

[104]. *Speech Modeling Based on Regularity Analysis.* Daoudi, K. e Véhel, J. L. Las Vegas : s.n. Proceedings of the IASTED/IEEE International Conference on Signal and Image Processing.

[105]. West, B. J., et al. Influence of Progressive Central Hypovolemia on Hölder Exponent Distribution of Cardiac Interbeat Intervals. *Annals of Biomedical Engineering.* 2004, Vol. 32, 8, pp. 1077-1087.

[106]. Daoudi, K. e Véhel, J. L., Meyer, Y. Construction of Continuous Functions with Prescribed Local Regularity. *Journal of Constructive Approximation.* 1998, Vol. 14, 3, pp. 349-385.

[107]. Arneodo, A., Grasseau, G. e Holschneider, M. Wavelet Transform of Multifractals. *Physical Review Letters.* Vol. 61, pp. 2281-2284.

[108]. Wornell, G. W. *Signal Processing with Fractals: A Wavelet-Based Approach.* s.l. : Prentice-Hall, 1996.

[109]. Daubechies, I. *Ten Lectures on Wavelets.* New York : SIAM, 1992.

[110]. Mallat, S. e Hwang, W. Singularity Detection and Processing with Wavelets. *IEEE Transactions on Information Theory.* 1992, Vol. 38, pp. 617-643.

[111]. Seuret, S. e Gilbert, A. C. Pointwise Hölder Exponent Estimation in Data Network Traffic. *ITC Specialist Seminar.*

[112]. Canus, C. *Robust Large Deviation Multifractal Spectrum Estimation.* Rocquencourt : INRIA, 1998.

[113]. Mandelbrot, B. B., Fisher, A. e Calvet, L. *A Multifractal Model of Asset Return.* Universidade de Yale. 1997. Relatório Técnico.

[114]. Peltier, R. F. e Véhel, J. L. *"Multifractal" Brownian Motion: Definition and Preliminary Results.* Rocquecourt : INRIA, 1995.

[115]. Bohr, T. e Tèl, T. The Thermodynamics of Fractals. *Direction in Chaos.* 1988, Vol. 2, pp. 194-237.

[116]. Arneodo, A., Bacry, E. e Muzy, J. F. The Thermodynamics of Fractals Revisited with Wavelets. *Physica A.* 1995, Vol. 213, pp. 232-275.

[117]. Tricot, C. *Curves and Fractal Dimension.* s.l. : Springer-Verlag, 1994.

[118]. Ayache, A. e Véhel, J. L. *The Generalized Multifractal Brownian Motion.* Rocquencourt : INRIA, 1991.

[119]. Canus, C., Véhel, J. L. e Tricot, C. *Continuous Large Deviation Multifractal Spectrum: Definition and Estimation.* Rocquencourt : INRIA, 1998.

[120]. Dekking, M., et al. *Fractals: Theory and Applications in Engineering.* s.l. : Springer-Verlag, 1999.

[121]. Nowak, R. W. *Discrete Cascade Universal Multifractal Simulation and Analysis.* Universidade McGill. Montreal : s.n., 1999. Tese de Mestrado.

[122]. Vojak, R. e Véhel, J. L. *Higher Order Multifractal Analysis.* Rocquencourt : INRIA.

[123]. Riedi, R. H. e Scheuring, I. Conditional and Relative Multifractal Spectra. *Fractals.* 1997, Vol. 5, 1.

[124]. *A Fractal Justification of the Normalization Step for Online Handwritting Recognition.* Vincent, N. e Dorizzi, B. 2000. Proceedings of the 7th International Workshop on Frontiers in Handwritten Recognition. pp. 535-540.

[125]. *A Novel Personal Biometric Authentication Technique Using Human Iris Based on Fractal Dimension Features.* Chen, W. S. e Yuan, S. Y. 2003. ICASSP 2003. Vol. 3, pp. 201-204.

[126]. *The Detection of Forged Handwritting Using a Fractal Number Estimate of Wrinkliness.* Chen, H. C., et al. 2003. Proceedings of the 11th International Graphonomics Society Conference. pp. 312-315.

[127]. Sant'Anna, R., Coelho, R. e Alcaim, A. Text-independent Speaker Recognition Based on the Hurst Parameter and the Multidimensional Fractional Brownian Motion Model. *IEEE Transactions on Audio, Speech & Language Processing.* 2006, Vol. 14, 3, pp. 931-940.

[128]. *Facial Biometry by Stimulating Salient Singularity Masks.* Lefevbre, G. e Garcia, C. 2007. IEEE Conference on Advanced Video and Signal Based Surveillance. pp. 511-516.

[129]. Boulétreau, V. *Vers un Classement de l'Écrit par des Méthodes Fractales.* Universidade Claude Bernarde Lyon I. 1997. Tese de Doutorado.

[130]. Wirotius, M. *Authentification par Signature Manuscrite sur Support Nomade.* Universidade François-Rabelais. 2005. Tese de Doutorado.

[131]. Shang, P. e Li, T. Multifractal Characteristics of Palmprint and Its Extracted Algorithm. *Applied Mathematical Modeling.* Dezembro de 2009, Vol. 33, 12, pp. 4378-4387.

[132]. *SVC2004: First International Signature Verification Competition.* Yeung, D. Y., et al. Hong Kong : Springer, 2004. Proceedings of the International Conference on Biometric Authentication. pp. 16-22.

[133]. Ortega-Garcia, J., et al. MCYT Baseline Corpus: a Bimodal Biometric Database. *IEE Proceedings on Vision, Image and Signal Processing.* Dezembro de 2003, Vol. 150, 6, pp. 395-401.

[134]. *Existence Test of Moments: Application to Multifractal Analysis.* Gonçalvès, P. Acapulco : s.n., 2000. Proceedings of the International Conference on Telecommunications.

[135]. Venugopal, V., et al. Revisiting Multifractality of High-Resolution Temporal Rainfall Using a Wavelet-Based Formalism. *Water Resources Research.* 2006, Vol. 42.

[136]. Stojic, T., Reljin, I. e Reljin, B. Adaptation of Multifractal Analysis to Segmentation of Microcalcification in Digital Mammograms. *Physica A.* 2006, Vol. 367, pp. 494-508.

[137]. Daoudi, K. e Véhel, J. L. Signal Representation and Segmentation Based on Multifractal Stationarity. *Signal Processing with Heavy-Tailed Models.* Dezembro de 2002, Vol. 82, 12, pp. 2015-2024.

[138]. Turiel, A. e Parga, N. Multifractal Wavelet Filter of Natural Images. *Physical Review Letters.* Outubro de 2000, Vol. 85, 15, pp. 3325-3328.

[139]. Zhong, J. e Ning, R. Image Denoising Based on Wavelets and Multifractals for Singularity Detection. *IEEE Transactions on Image Processing.* Outubro de 2005, Vol. 14, 10, pp. 1435-1447.

[140]. Bellman, R. *Dynamic Programming.* s.l. : Princeton University Press, 1957.

[141]. Dempster, A., Laird, N. e Rubin, D. Maximum Likelihood Estimation from Incomplete Data Using the EM Algorithm. *Journal of the Royal Statistics Society.* 1977, Vol. 39, pp. 1-38.

[142]. Webb, A. *Statistical Pattern Recognition.* s.l. : Wiley, 2002.

[143]. MacKay, D. J. C. *Information Theory, Inference, and Learning Algorithms.* Cambridge : Cambridge University Press, 2003.

[144]. Larsen, J. *Design of Neural Network Filters.* Instituto de Eletrônica, Technical University of Denmark. 1996. Tese de Doutorado.

[145]. *Gaussian Mixture Regularization Through Parzen Method with Prunning - An Acceptance Region Based Approach.* Montalvão, J. R., Canuto, J. C. Juiz de Fora : SBA, 2008. Anais do Congresso Brasileiro de Automática.

[146]. *Comparing GMM and Parzen in Automatic Signature Recognition - a Step Backward or Forward?* Montalvão, J. R., Houmani, N. e Dorizzi, B. s.l. : SBA, 2010. Anais do Congresso Brasileiro de Automática.

[147]. Bigun, E. S., et al. Expert Conciliation for Multi Modal Person Authentication Systems by Bayesian Statistics. *Lecture Notes in Computer Science 1206: Proceedings of the AVBPA.* s.l. : Springer, 1997, pp. 291-300.

[148]. Kittler, J., et al. On Combining Classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 1998, Vol. 20, pp. 226-239.

[149]. *Effectiveness of Hölder Function for Online Signature Forgery Detection.* Canuto, J. C. e Lee, L. L. Vitória : s.n., 2010. Proceedings of the ISSNIP Biosignals and Biorobotics Conference. pp. 199-203.

[150]. Montalvão, J. R. e Freire, E. O. On the Equalization of Keystroke Timing Histograms. *Pattern Recognition Letters.* 2006, Vol. 27, 13, pp. 1440-1446.

[151]. Ejarque, P., et al. On the Use of Genuine-Impostor Statistical Information for Score Fusion in Multimodal Biometrics. *Annals of Telecommunications.* Janeiro-Fevereiro de 2007, Vol. 62, 1-2, pp. 109-129.

[152]. *Using adapted levenshtein distance for on-line signature authentication.* Schimke, S., Vielhauer, C. e Dittmann, J. 2004. Proceedings of the International Conference on Pattern Recognition. Vol. 2, pp. 931–934.

[153]. Yanikoglu, B. e Kholmatov, A. Online Signature Verification Using Fourier Descriptors. *EURASIP Journal on Advances in Signal Processing.* 2009, Vol. 2009.

[154]. Faundez-Zanuy, M. On-line signature recognition based on VQ-DTW. *Pattern Recognition.* 2007, Vol. 40, 3, pp. 981-992.

[155]. Kholmatov, A. e Yanikoglu, B. Identity authentication using improved online signature verification method. *Pattern Recognition Letters.* 2005, Vol. 26, 15, pp. 2400-2408.

[156]. Fierrez-Aguilar, J., Ortega-Garcia, J. e Gonzalez-Rodriguez, J. Target dependent score normalization techniques and their application to signature verification. *IEEE Transactions on Systems, Man, and Cybernetics, Part C.* 2005, Vol. 35, 3, pp. 418–425.

[157]. *Spectrum analysis based on windows with variable widths for online signature verification.* Z.-H. Quan, D.-S. Huang, X.-L. Xia, M. R. Lyu, and T.-M. Lok. 2006. Proceedings of the International Conference on Pattern Recognition. Vol. 2, pp. 1122–1125.

[158]. Ly Van, B., Garcia-Salicetti, S. e Dorizzi, B. On using the viterbi path along with hmm likelihood information for on-line signature verification. *IEEE Transactions on Systems, Man and Cybernetics, Part B.* 2007, Vol. 37, 5, pp. 1237–1247.

[159]. Nanni, L. e Lumini, A. Advanced methods for two-class problem formulation for on-line signature verification. *Neurocomputing.* 2006, Vol. 69, 7-9, pp. 854-857.

[160]. Nanni, L. e Lumini, A. A novel local on-line signature verification system. *Pattern Recognition Letters.* 2008, Vol. 29, 5, pp. 559–568.

[161]. Fano, R. *Transmission of information: a statistical theory of communications.* s.l. : M.I.T. Press, 1961.

[162]. Papoulis, A. e Pillai, U. *Probability, Random Variables and Stochastic Processes.* s.l. : McGraw-Hill, 2002.

[163]. Haykin, S. *Redes Neurais: Princípios e Prática.* s.l. : Bookman, 2002.

[164]. Bell, A. e Sejnowski, T. An information-maximization approach to blind separation and blind deconvolution. *Neural Computation.* 1995, Vol. 7, pp. 1129-1159.

[165]. Viola, P., Schraudolph, N. e Sejnowski, T. Empirical entropy manipulation for real-world problems. *Neural Information Processing Systems.* 1996, Vol. 8, pp. 851-857.

[166]. *Unsupervised learning for nonlinear synthetic discriminant functions.* J., Fisher e Principe, J. C. 1995. Proceedings of the SPIE. pp. 1-13.

[167]. *A Nonparametric Method for Information Theoretic Feature Extraction.* J., Fisher e Principe, J. C. 1997. Proceedings of the DARPA Image Understanding Workshop.

[168]. *Entropy Manipulation of Arbitrary Nonlinear Mappings.* J., Fisher e and Principe, J. C. 1997. p. Proceedings of the Neural Networks for Signal Processing Workshop.

[169]. *A methodology for information theoretic feature extraction.* Fisher, J. W. e Principe, J. C. 1998. Proceedings of the IEEE International Joint Conference on Neural Networks. pp. 1712-1716.

[170]. Cover, T. e Thomas, J. *Elements of Information Theory.* s.l. : John Wiley & Sons, 1991.

[171]. Fisher, J. W., Ihler, A. e Viola, P. Learning Informative Statistics: a Nonparametric Approach. *Proceedings of Neural Information Processing Systems.* 1999.

[172]. *Estimation of signal information content for classification.* Fisher, J. W., Siracusa, M. e Tieu, K. 2009. Proceedings of the IEEE DSP Workshop.

[173]. Hellman, M. e Raviv, J. Probability of error, equivocation, and the chernoff bound. *IEEE Transactions on Information Theory.* Julho de 1970, Vol. 16, 4, pp. 368-372.

[174]. *Entropy-base Variable Frame Rate Analysis of Speech Signals and Its Application to ASR.* H. You, Q. Zhu, and A. Alwan. Montreal : s.n., 2004. Proceedings of the ICASSP. pp. 549-552.

Apêndice A – Tabelas de Resultados

BASE: MCYT-100	DIST		STAT	
	EER_{DADOS} (%)	EER_{PONTOS} (%)	EER_{DADOS} (%)	EER_{PONTOS} (%)
$\boldsymbol{X}$	16,4±0,3	–	8,5±0,1	–
$\boldsymbol{Y}$	6,9±0,3	–	7,2±0,1	–
$\boldsymbol{P}$	12,0±0,4	–	8,5±0,2	–
$\boldsymbol{X}^{\sim}$	13,9 ±0,5	–	9,4±0,4	–
$\boldsymbol{Y}^{\sim}$	15,1±0,2	–	9,8±0,5	–
$\boldsymbol{P}^{\sim}$	16,6±0,2	–	11,3±0,2	–
$\boldsymbol{X}^{\psi}$	13,0±0,5	–	10,1±0,3	–
$\boldsymbol{Y}^{\psi}$	14,1±0,2	–	10,6±0,3	–
$\boldsymbol{P}^{\psi}$	9,6±0,1	–	10,9±0,7	–
$\boldsymbol{C}(\boldsymbol{X},\boldsymbol{X}^{\sim})$	9,7±0,5	10,9±0,6	7,2±0,4	8,3±0,4
$\boldsymbol{C}(\boldsymbol{Y},\boldsymbol{Y}^{\sim})$	6,6±0,2	7,5±0,4	6,5±0,3	7,8±0,5
$\boldsymbol{C}(\boldsymbol{P},\boldsymbol{P}^{\sim})$	9,5±0,2	10,7±0,4	7,9±0,4	9,3±0,3
$\boldsymbol{C}(\boldsymbol{X},\boldsymbol{X}^{\psi})$	9,3±0,3	8,6±0,6	7,5±0,5	8,6±0,3
$\boldsymbol{C}(\boldsymbol{Y},\boldsymbol{Y}^{\psi})$	6,1±0,4	6,7±0,4	5,8±0,1	7,5±0,3
$\boldsymbol{C}(\boldsymbol{P},\boldsymbol{P}^{\psi})$	7,7±0,3	8.0±0,4	7,2±0,7	9,5±0,7
$\boldsymbol{C}(\boldsymbol{X},\boldsymbol{X}^{\sim},\boldsymbol{X}^{\psi})$	8,4±0,6	8,6±0,8	6,6±0,1	7,9±0,5
$\boldsymbol{C}(\boldsymbol{Y},\boldsymbol{Y}^{\sim},\boldsymbol{Y}^{\psi})$	6,2±0,2	7,7±0,4	5,6±0,1	7,7±0,6
$\boldsymbol{C}(\boldsymbol{P},\boldsymbol{P}^{\sim},\boldsymbol{P}^{\psi})$	8,4±0,4	8,5±0,5	7,7±0,2	9,5±0,8
$\boldsymbol{C}(\boldsymbol{X},\boldsymbol{Y})$	6,3±0,6	7,0±0,4	5,7±0,3	7,8±0,1
$\boldsymbol{C}(\boldsymbol{X}^{\sim},\boldsymbol{Y}^{\sim})$	9,3±0,6	11,6±0,5	7,8±0,1	9,2±0,6
$\boldsymbol{C}(\boldsymbol{X}^{\psi},\boldsymbol{Y}^{\psi})$	9,6±0,5	10,7±0,5	9,1±0,4	9,6±0,4
$\boldsymbol{C}(\boldsymbol{X},\boldsymbol{Y},\boldsymbol{X}^{\sim},\boldsymbol{Y}^{\sim})$	5,6±0,1	7,1±0,7	4,6±0,1	7,6±0,7
$\boldsymbol{C}(\boldsymbol{X},\boldsymbol{Y},\boldsymbol{X}^{\psi},\boldsymbol{Y}^{\psi})$	5,1±0,3	5,7±0,7	4,6±0,1	7,4±0,4
$\boldsymbol{C}(\boldsymbol{X},\boldsymbol{Y},\boldsymbol{X}^{\sim},\boldsymbol{Y}^{\sim},\boldsymbol{X}^{\psi},\boldsymbol{Y}^{\psi})$	4,8±0,3	6,6±0,9	4,3±0,1	7,2±0,8
$\boldsymbol{C}(\boldsymbol{S})$	4,3±0,2	5,4±0,6	4,2±0,3	6,8±0,2
$\boldsymbol{C}(\boldsymbol{S}^{\sim})$	7,1±0,4	9,5±0,6	6,3±0,3	8,6±0,7
$\boldsymbol{C}(\boldsymbol{S}^{\psi})$	6,4±0,1	7,2±0,5	6,5±0,4	8,9±0,8
$\boldsymbol{C}(\boldsymbol{S},\boldsymbol{S}^{\sim})$	4,0±0,2	5,6±0,8	3,5±0,1	6,6±0,7
$\boldsymbol{C}(\boldsymbol{S},\boldsymbol{S}^{\psi})$	3,7±0,2	4,3±0,8	3,6±0,2	6,7±0,9
$\boldsymbol{C}(\boldsymbol{S},\boldsymbol{S}^{\sim},\boldsymbol{S}^{\psi})$	3,9±0,3	5,2±1,0	3,3±0,1	6,8±1,1

BASE: SVC2004	DIST		STAT	
	EER_{DADOS} (%)	EER_{PONTOS} (%)	EER_{DADOS} (%)	EER_{PONTOS} (%)
$\boldsymbol{X}$	25,3±1,9	–	15,3±0,3	–
$\boldsymbol{Y}$	17,1±0,8	–	13,7±0,3	–
$\boldsymbol{P}$	26,9±1,2	–	15,8±0,6	–
$\boldsymbol{X}^{\sim}$	23,1±0,6	–	16,2±0,8	–
$\boldsymbol{Y}^{\sim}$	25,2±0,7	–	17,2±0,4	–
$\boldsymbol{P}^{\sim}$	29,7±1,1	–	17,4±0,8	–
$\boldsymbol{X}^{\psi}$	31,4±1,2	–	17,8±0,7	–
$\boldsymbol{Y}^{\psi}$	30,6±1,0	–	19,8±0,6	–
$\boldsymbol{P}^{\psi}$	31,9±0,6	–	17,9±0,8	–
$\boldsymbol{C}(\boldsymbol{X}, \boldsymbol{X}^{\sim})$	19,7±0,3	20,8± 2,0	12,2±1,6	18,9±0,9
$\boldsymbol{C}(\boldsymbol{Y}, \boldsymbol{Y}^{\sim})$	19,1±0,4	21,5±1,1	14,6±1,0	16,9±0,5
$\boldsymbol{C}(\boldsymbol{P}, \boldsymbol{P}^{\sim})$	21,9±0,7	26,0±1,6	15,8±0,6	15,5±1,0
$\boldsymbol{C}(\boldsymbol{X}, \boldsymbol{X}^{\psi})$	28,1±0,8	28,7±2,2	15,2±0,4	19,4±0,8
$\boldsymbol{C}(\boldsymbol{Y}, \boldsymbol{Y}^{\psi})$	23,7±0,9	26,7±1,3	16,5±0,8	18,3±0,7
$\boldsymbol{C}(\boldsymbol{P}, \boldsymbol{P}^{\psi})$	22,6±0,7	26,1±1,3	17,2±0,5	16,6±1,0
$\boldsymbol{C}(\boldsymbol{X}, \boldsymbol{X}^{\sim}, \boldsymbol{X}^{\psi})$	21,2±0,5	22,8±2,3	14,1±0,5	18,6±1,1
$\boldsymbol{C}(\boldsymbol{Y}, \boldsymbol{Y}^{\sim}, \boldsymbol{Y}^{\psi})$	20,4±1,0	21,2±1,5	15,2±0,8	17,6±0,8
$\boldsymbol{C}(\boldsymbol{P}, \boldsymbol{P}^{\sim}, \boldsymbol{P}^{\psi})$	22,9±0,7	26,7±1,7	18,1±0,8	15,4±1,3
$\boldsymbol{C}(\boldsymbol{X}, \boldsymbol{Y})$	16,2±0,8	17,0±2,1	10,8±0,5	18,5±0,4
$\boldsymbol{C}(\boldsymbol{X}^{\sim}, \boldsymbol{Y}^{\sim})$	16,1±0,6	21,9±0,9	15,8±1,2	16,9±0,9
$\boldsymbol{C}(\boldsymbol{X}^{\psi}, \boldsymbol{Y}^{\psi})$	25,7±0,8	30,4±1,6	18,2±1,1	22,2±0,9
$\boldsymbol{C}(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{X}^{\sim}, \boldsymbol{Y}^{\sim})$	14,9±0,6	16,9±2,3	12,6±1,0	17,1±1,0
$\boldsymbol{C}(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{X}^{\psi}, \boldsymbol{Y}^{\psi})$	21,8±0,7	25,9±2,6	14,1±1,1	18,4±1,0
$\boldsymbol{C}(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{X}^{\sim}, \boldsymbol{Y}^{\sim}, \boldsymbol{X}^{\psi}, \boldsymbol{Y}^{\psi})$	15,8±0,4	20,7±2,7	13,0±1,1	17,7±1,4
$\boldsymbol{C}(\boldsymbol{S})$	14,9±0,4	18,9±2,4	9,4±0,6	16,4±0,7
$\boldsymbol{C}(\boldsymbol{S}^{\sim})$	15,4±1,3	21,7±1,4	15,0±1,1	16,1±1,2
$\boldsymbol{C}(\boldsymbol{S}^{\psi})$	15,5±1,0	30,4±1,7	20,0±0,7	20,4±1,2
$\boldsymbol{C}(\boldsymbol{S}, \boldsymbol{S}^{\sim})$	12,8±0,8	16,9±2,8	10,5±0,8	15,6±1,4
$\boldsymbol{C}(\boldsymbol{S}, \boldsymbol{S}^{\psi})$	19,9±1,1	24,0±2,9	14,8±0,9	17,2±1,4
$\boldsymbol{C}(\boldsymbol{S}, \boldsymbol{S}^{\sim}, \boldsymbol{S}^{\psi})$	15,7±0,5	21,0±3,2	12,3±0,7	15,7±1,9