Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e de Computação
Departamento de Comunicações

Autenticação Pessoal baseada no Som da Assinatura

Autor: Julio César Larco Bravo

Orientador: Prof. Dr. João Baptista Tadanobu Yabu-uti

Banca Examinadora:

Dr. Miguel Gustavo Lizárraga. ..... (Samsung)

Prof. Dr. Luiz César Martini. (FEEC/Unicamp)

Prof. Dr. Yuzo Iano. ..... (FEEC/Unicamp)

Dissertação de Mestrado apresentada à Faculdade de Engenharia Elétrica e de Computação como parte dos requisitos para obtenção do título de Mestre em Engenharia Elétrica.

Campinas, SP

Abril/2006

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

Larco Bravo, Julio César
L321a Autenticação pessoal baseada no som da assinatura
Julio César Larco Bravo. –Campinas, SP: [s.n.], 2006.

Orientador: João Baptista Tadanobu Yabu-uti.
Dissertação (Mestrado) - Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Reconhecimento de padrões. 2. Biometria. I. Yabu-uti, João Baptista Tadanobu. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

Título em Inglês:	Personal authentication based on sound of signature
Palavras-chave em Inglês:	Pattern recognition, Biometrics
Área de concentração:	Telecomunicações e Telemática
Titulação:	Mestre em Engenharia Elétrica
Banca Examinadora:	Miguel Gustavo Lizárraga, Luiz César Martini e Yuzo Iano
Data da defesa:	26/04/2006

Resumo

Uma assinatura manuscrita, é a forma mais utilizada para confirmar a identidade de uma pessoa, já que o estilo de assinar de um indivíduo é uma entidade biométrica que pode ser usada para diferenciar uma pessoa de outra.

Neste trabalho, apresenta-se uma metodologia para realizar a autenticação pessoal utilizando o som que se produz no momento de assinar. Quando uma pessoa assina, a fricção entre a ponta rígida de uma caneta e o papel produz um som que pode ser usado para verificar a identidade de uma pessoa. Esta metodologia, está baseada no fato de que o som produzido ao assinar está correlacionado com a dinâmica e a postura do assinante. Cada um dos traços que compõe a assinatura corresponde a uma parte do sinal de som. Diferentes indivíduos produzem diferentes traços ou garranchos os quais resultam em diferentes sinais de som.

Do sinal de som capturado são eliminadas as amostras que não correspondem ao som da assinatura e calcula-se a envoltória deste sinal. Cada um dos traços que formam parte da assinatura são caracterizados como variações agudas nos valores da envoltória (picos), as quais são representadas como vetores binários de características que são enviados para uma etapa de reconhecimento de padrões, a qual decidirá se o som capturado provém de uma assinatura que foi realizada por um usuário legítimo ou por um impostor.

A metodologia apresentada é avaliada utilizando um conjunto de amostras de teste e treinamento pertencentes a dois tipos de usuários: legítimo e impostor habilidoso. O usuário legítimo é o proprietário da assinatura e o impostor habilidoso conhece a forma como o usuário legítimo assina. Como parâmetros de avaliação do desempenho desta metodologia, foram obtidas as taxas de erro FAR (falsa aceitação) e FRR (falsa rejeição) de 8,55% e 8,73%, respectivamente.

Abstract

A signature is the most used way to validate a person's identity, since the style of every individual signature constitutes a biometric entity, which can be used to differentiate one person from another.

This research work presents a method to accomplish the personal authentication using the sound produced when a person is signing. During this event, the friction produced by the rigid tip of a pen rubbing the paper, generates a sound that can be used to verify the identity of a person. The reliability of this methodology is based on the fact that the sound emitted during the signature action is closely correlated with the dynamics and posture of the person who signs. Moreover, every line of the signature corresponds directly to one part of the audio signal generated. Therefore, different individuals are able to produce completely different traces or scrawls, which will generate different audio signals.

Once the audio signal is digitally captured, the samples that do not belong to the signature are discarded and the envelope of this signal is computed. Every constituent trace of the signature are characterized as sharp variations of the envelope values (peaks), which are represented as binary vectors of features. This information, is sent to a pattern recognizing stage which has the responsibility to decide whether the captured sound corresponds to an authentic user or an impostor.

The presented methodology is evaluated using a group of test and training samples belonging to two types of users: legitimate and skilled impostor. The legitimate user is the proprietor of the signature and the skilled impostor knows the form as the legitimate user signs. As parameters of evaluation of this methodology, were obtained the error rate FAR (false acceptance) and FRR (false rejection) of 8,55% and 8,73%, respectively.

A Sandy, por el amor, paciencia y comprensión.

A mis padres y hermanos.

Agradecimentos

- A Deus, porque sem ele nada é possível.
- A minha família e a família de Sandy, por sua ajuda incondicional e por sempre crer em mim.
- Ao Prof. João B. Yabu-uti, pela oportunidade para realizar meus estudos de pós-graduação. Pela amizade, ajuda, paciência e apoio incondicional para o termino deste trabalho, e durante toda minha permanência no Brasil. Muito obrigado professor.
- Ao Dr. Miguel Lizárraga, pelos conselhos e amizade. Por ser um amigo que sempre esteve presente quando precisei sua ajuda. Gracias Miguel.
- Aos professores Dr. Luiz César Martini e Dr. Yuzo Iano, membros da banca examinadora.
- Ao Prof. Dr. Lee Luan Ling e aos colegas do laboratório LRPRC (Unicamp), pela ajuda para começar este trabalho.
- A Gonzalo, Nancy e Camila, por serem a minha família no Brasil.
- A Carlos, um amigo que sempre teve palavras de motivação.
- Aos amigos: Rodrigo, Marcio, Tarciana, Andre e José, por compartilhar sua amizade e por todos os bons momentos.
- A Luis, Jaime e Rubén León e aos amigos da Escuela Politécnica del Ejército, pelo incentivo e motivação.
- À Capes e à Escuela Politécnica del Ejército no Equador pela ajuda.

Sumário

Lista de Figuras

Lista de Tabelas

Lista de Abreviaturas

DTW	-	Dynamic Time Warping.
EA	-	Emissão Acústica.
ERR	-	Taxa de Erro Igual.
FAR	-	Taxa de Falsa Aceitação.
FFT	-	Transformada Rápida de Fourier.
FRR	-	Taxa de Falsa Rejeição.
HMM	-	Hidden Markov Models.
HT	-	Transformada de Hilbert.
$NEIA$	-	Número de Tentativas de Identificação de Usuário Legitimo.
NFA	-	Número de Falsas Aceitações.
$NIIA$	-	Número de Tentativas de Identificação de Usuário Impostor.
NFR	-	Número de Falsas Rejeições.
PIN	-	Personal Identification Number.
RMS	-	Raiz Quadrática Meia.
ROC	-	Característica Operacional do Receptor.
$SVAD$	-	Sistemas de Verificação de Assinaturas Dinâmicas .
$SVAE$	-	Sistemas de Verificação de Assinaturas Estáticas.
XOR	-	Operação Lógica OR Exclusivo.

Lista de Símbolos

B	-	Amostra biométrica
d	-	Função discriminante linear
D	-	Distância Euclidiana
Di	-	Distância de similaridade de blocos de vetores vc
$Di_{l,k}$	-	Distância entre a assinatura verdadeira l e a assinatura falsa k
Dp	-	Distância calculada entre vetores
$Dt_{l,k}$	-	Distância da assinatura k respeito a assinatura verdadeira l
$Dv_{i,j}$	-	Distância entre as assinaturas verdadeiras i e j
E_k	-	Energia de cada quadro ou janela
EdB_k	-	Energia por quadro em decibéis
$EdBmax_l$	-	Valor máximo da energia EdB na janelas l
$EdBmin_l$	-	Valor mínimo da energia EdB na janelas l
$env[n]$	-	Envoltória do sinal de som discreta $s[n]$
$Env[n]$	-	Normalização do sinal env
F	-	Número de amostras do sinal de som por quadro
H_d	-	Transformada discreta de Hilbert de $s[n]$
I_j	-	Sinal da assinatura do mesmo assinante feito por um impostor
K	-	Número de blocos nos que se divide a envoltória
K_c	-	Número de quadros do sinal
L	-	Comprimento dos vetores de características
lb_i	-	Limiares para obter o vetor de características pelo método II
L_c	-	Número de quadros de m_c amostras
$lmin$	-	Limite para determinar pontos início-fim
$lmax$	-	Limite para determinar pontos início-fim
lp	-	Limiar para determinar a posição dos pontos máximos e mínimos,
m	-	Vetor médio
m_c	-	Número de amostras da energia por quadro
mi	-	Inclinação de uma reta
N	-	Número de amostras da envoltória
n_e	-	Número de pontos da envoltória normalizada Env
$p(x)$	-	Função densidade de probabilidade de x
$P(x)$	-	Função distribuição de probabilidade de x
p_c	-	Número de amostras igualmente espaçadas obtidas de $env[n]$

$pinicio$ - Posição da amostra de início do sinal de som
$pfinal$ - Posição da amostra final do sinal de som
$pm[i]$ - Posição i-ésima do ponto mínimo
$px[i]$ - Posição i-ésima do ponto máximo
q - Vetor das posições dos pontos máximos e mínimos
Q - Template
$r(t)$ - Envoltória de $x(t)$
S_n - Sinal filtrado do som da assinatura
S_i, S_j - Sinais verdadeiros do som de um mesmo assinante
s_m - Medida de semelhança dos templates
$s[n]$ - n-ésima amostra do sinal $x(t)$ digitalizado
St - Assinatura de teste
Td - Limiar de decisão
vc - Blocos do vetor de características
Vcb - Vetor de características calculado pelo método II
Vcb_i - Vetores auxiliares de características calculado com lb_i
Vcp - Vetor de características calculado pelo método I
w_i - i-ésima classe de padrões
x_i - i-ésima característica
X - Vetor de características n-dimensional
$x(t)$ - Sinal de som de entrada
$\hat{x}(t)$ - Transformada de Hilbert do sinal $x(t)$
$\|\|$ - Norma Euclidiana
μ - Média
σ^2 - Variância
$\sum$ - Matriz de covariância
$\otimes$ - Convolução
$\circ$ - Representação de um ponto mínimo
$\times$ - Representação de um ponto máximo
$\oplus$ - Operador OR exclusivo

Capítulo 1

Introdução

1.1 Identificação Pessoal

Hoje em dia, existem muitos sistemas de controle os quais restrigem o acesso a lugares ou serviços. Por esta razão, o ser humano está habituado a comprovar sua identidade para realizar muitas das atividades do dia a dia. Este fato é tão comum que é considerado como algo totalmente normal. Assim sendo, é perfeitamente aceitável utilizar senhas ou PIN (*Personal Identification Number*) e cartões magnéticos para retirar dinheiro no caixa eletrônico do banco, precisar de assinaturas ou um documento com foto para autenticar a identidade no momento de realizar uma operação com cheques bancários ou para ingressar em um determinado recinto. O propósito de todos estes procedimentos é oferecer uma evidência adicional para confirmar a identidade, e corroborar que realmente o indivíduo é quem diz ser. Esta evidência é provida de várias formas diferentes, e conseqüentemente existem vários meios para autenticar a identidade de uma pessoa: [1].

- Aparência (como a pessoa é, por exemplo altura, gênero, peso)
- Comportamento social (como uma pessoa interage com outros)
- Nome (como a pessoa é conhecida)
- Códigos (como uma pessoa é conhecida por uma organização)

- Conhecimento (o que a pessoa sabe)
- Posse (o que a pessoa possui)
- Bio-dinâmica (o que a pessoa faz)
- Fisiologia natural (quem a pessoa é, por exemplo características faciais)
- Características físicas impostas (o que está com a pessoa agora, por exemplo etiquetas, colarinhos, pulseiras).

A meta de autenticação é proteger um sistema contra o seu uso sem autorização. Esta característica também habilita a proteção dos usuários registrados negando a possibilidade de intrusos personificarem os usuários autorizados.

Existem basicamente três técnicas de autenticar uma pessoa em um sistema [3]:

1. Com base no seu conhecimento ou em algo que só o indivíduo sabe, como por exemplo, uma senha ou um PIN. Este tipo de processo de autenticação é chamado Prova por Conhecimento.
2. Com base em um dispositivo físico que só o indivíduo possui e que efetua a autenticação, como por exemplo, um cartão magnético, um crachá ou um *smart card*. Este tipo de processo de autenticação é chamado Prova por Posse;
3. Com base em algo que o indivíduo é, como uma medida fisiológica, uma característica comportamental ou um padrão ou atividade específica do indivíduo, que distingue, de forma confiável, de outros seres humanos e que pode ser utilizado para autenticar a sua identidade. Como por exemplo: impressões digitais, geometria da mão, íris, assinatura, fala, dentre outros. Este tipo de processo de autenticação é chamado Prova por Biometria.

As duas primeiras técnicas são os meios mais utilizados para autenticar a identidade de um indivíduo [1, 2] por ser de fácil manuseio e baixo custo. Por este motivo, é muito comum ser necessário memorizar os números de identificação, senha do cartão do banco; *login* do computador, número de passaporte, de CPF, entre outros. Ou ainda, o indivíduo ter que

carregar vários documentos de identificação, como por exemplo, carteira de identidade, cartões e chaves. Embora, nenhum desses métodos seja 100% confiável, já que, por exemplo, muitas vezes o usuário é displicente com sua senha, seja anotando-a ao invés de memorizá-la, ou escolhendo-a de forma óbvia, como o seu sobrenome, a sua data de nascimento ou time para o qual torce. Também os cartões magnéticos podem ser roubados, emprestados, perdidos, copiados ou falsificados.

Com o aumento do poder de processamento dos computadores a biometria vêm conquistando espaço atualmente em relação aos mecanismos clássicos de autenticação pessoal com a vantagem de que as características biométricas não podem ser roubadas, perdidas ou esquecidas. Por estas razões, tem aumentado o interesse em desenvolver métodos de verificação de identidade pessoal que levem em consideração estratégias fundamentadas em medidas biométricas, onde "biométrica" [4] é definida como: "Uma característica física ou um aspecto do comportamento de uma pessoa que pode ser usada para reconhecer ou verificar a identidade do indivíduo, ou seja, de uma pessoa registrada no sistema".

As técnicas biométricas mais comuns são [1]:

- Verificação da Assinatura.
- Análise da Retina.
- Verificação da Impressão Digital.
- Análise Facial.
- Geometria da Mão.
- Verificação da Voz.

Os métodos biométricos relativamente novos são: padrão do DNA, reconhecimento da orelha, detecção do odor, análise de suor, análise da dinâmica da digitação, análise da forma da cabeça.

As técnicas biométricas são divididas em duas categorias [1][2]. A primeira, a biometria fisiológica, que engloba características fisiológicas tais como o padrão da íris, as impressões

digitais, a forma do contorno da mão e face. E a segunda, a biometria de comportamento, que analisa características relativas ao comportamento como: o tipo da escrita, assinatura, dinâmica da digitação, voz, etc.. A escolha de um método de autenticação de identidade através de características biométricas, seja pela abordagem fisiológica ou de comportamento, pode gerar amplos debates. Nesse contexto, a abordagem fisiológica está bastante aberta a discussões. Por outro lado, a abordagem de comportamento possui um consenso geral de que a utilização de assinaturas manuscritas tem vantagens significativas, embora outras alternativas possam ser consideradas. Isso se deve aos fatos: [2]

a. A assinatura é o método mais natural e mais amplamente utilizado para confirmar nossa identidade.

b. A utilização de técnicas de verificação automática de assinaturas agilizam processos e minimizam erros em operações que solicitem a autenticação de indivíduos através de assinaturas.

c. Medidas das características de assinaturas não são invasivas (quando comparadas com outras técnicas, como por exemplo, as medidas feitas sobre a íris) e não tem conotações negativas ou de saúde indesejável (se comparadas com medidas feitas sobre impressões digitais que envolve o contato físico direto com uma superfície possivelmente contaminada).

Outra característica biométrica amplamente utilizada é a escrita ou estilo de caligrafia de um indivíduo, que pode exclusivamente diferenciá-lo dos outros. As assinaturas manuscritas são usadas em meios forenses para autenticação de documentos e trabalhos de arte. As singularidades da escrita estão situadas não somente nas características topológicas da caligrafia, i.e. formas das letras e palavras, mas também na dinâmica da postura de escrever, como a força, a velocidade e a aceleração, etc. [5]. Para realizar uma autenticação automatizada, as características topológicas são extraídas, identificadas e autenticadas usando o processamento digital de sinais associado aos métodos de reconhecimento de padrões. É bem conhecido que o uso de aspectos não topológicos como critérios extras melhoram a segurança da autenticação.

Este tipo de biometria está intimamente relacionada com a assinatura de uma pessoa, já que a forma de escrever tem influência no momento de fazer sua assinatura.

1.2 Verificação de Assinaturas

A Verificação de Assinaturas é um método de autenticação pessoal baseada em uma biometria comportamental que analisa a maneira como um usuário faz a sua assinatura. Ao assinar, características tais como a velocidade e a pressão exercida pela mão que prende a caneta são tão importantes quanto a forma gráfica de uma assinatura [4]. Este método biométrico é baseado no fato de que assinar é uma ação reflexo não influenciada pelo controle muscular deliberado, com determinadas características (ritmo, toques sucessivos na superfície da escrita, velocidade, aceleração) [1].

Os sistemas desenvolvidos baseados nesta biometria são de dois tipos [6]:

- Verificação de Assinaturas Estáticas (SVAE).
- Verificação de Assinaturas Dinâmicas (SVAD).

Os sistemas de verificação de assinaturas dinâmicas (SVAD), também chamados sistema *on-line*, se caracterizam pelo uso de informações dinâmicas do processo de escrita, tais como velocidade e aceleração. A taxa de acerto desses sistemas é elevada, mesmo quando as falsificações das assinaturas são feitas por especialistas. O bom desempenho desses sistemas é obtido porque os falsificadores procuram imitar somente a forma da assinatura original, e não conseguem reproduzir os traços da assinatura nos mesmos intervalos de tempo que o indivíduo genuíno. Desta forma, esses sistemas se valem da falta de consistência dos dados temporais nas assinaturas falsificadas, para obter uma melhor discriminação entre as assinaturas verdadeiras e as assinaturas falsas.

O principal limitante desse tipo de sistema é que a assinatura sempre deve ser escrita utilizando um equipamento que permite a aquisição das informações dinâmicas. Este equipamento é uma caneta especial para capturar a informação, uma mesa digitalizadora com uma superfície especial para coletar os dados ou uma combinação das duas. No entanto, escrever utilizando este tipo de equipamentos como a mesa digitalizadora, muitas vezes não deixa que o escritor assine de maneira natural, implicando em mudanças do estilo de sua assinatura [6].

Para evitar as imperfeições de usar uma mesa, existem canetas especiais que são úteis no momento de assinar. Estas canetas muitas vezes contem transdutores embutidos (caneta de aquisição digital) para obter a informação dinâmica sobre postura de escrever [7] [8]. Estas canetas especiais causam pequenas interferências no processo de assinar.

Os sistemas de verificação de assinaturas estáticas (SVAE), também chamados sistemas *off-line*, se caracterizam por utilizar apenas a imagem da assinatura para extrair as informações que alimentam o sistema. As taxas de acerto dos SVAE são geralmente inferiores às taxas de acerto dos SVAD. Isso se deve ao fato de que as imagens de assinaturas podem ser facilmente copiadas e a informação dinâmica que poderia ser extraída dessas imagens torna-se altamente degradada na amostra estática. Dessa forma, um bom falsificador poderia criar uma copia suficientemente fiel da assinatura original, levando o sistema a classificar erroneamente a assinatura falsificada como sendo verdadeira. A principal vantagem dos SVAE é preservar ao máximo a naturalidade do processo de escrita, pois no ato de assinar não existe nenhum tipo de dispositivo que venha interferir diretamente na escrita da assinatura.

As canetas de ponta rígida são provavelmente os instrumentos mais comuns para escrever. Ao escrever com tais instrumentos sobre o papel, a ponta rígida da caneta interage com a fina textura da superfície do papel e produz sons audíveis. Estes sons estão correlacionados com a dinâmica ou postura de escrever e portanto contêm informação útil para distinguir escritores diferentes [5]. Este tipo de verificação de assinatura é conhecido como Emissão Acústica (EA - *Acoustic Emission*). Quando um usuário escreve, o movimento da ponta de caneta sobre as fibras do papel gera emissões acústicas, que se propagam no interior do material que está abaixo do papel sobre o qual se está assinando. As ondas em materiais de estrutura sólida elástica se comportam de um modo semelhante as ondas sonoras no ar e são detectadas por um sensor preso ao bloco de escrita [4].

Utilizar o som da escrita para obter informação sobre a dinâmica no momento de escrever é algo recente. Existem dois trabalhos prévios [5] e [9] onde uma caneta com um microfone embutido foi desenvolvida com o propósito de analisar textos manuscritos e alguns resultados experimentais mostram adequadamente a utilidade de sinais de sons para a análise de texto manuscrito e autenticação.

O interesse de usar o som da escrita para a verificação e/ou autenticação do escritor, é que capturar sons com um microfone é muito mais fácil que medir a dinâmica de escrever sobre uma mesa digitalizadora. Pressão, velocidade e acelera: o sinal do som. Como resultado, o sinal do som tem menos pontos de dados que o conjunto de dados obtidos de uma mesa digitalizadora, mas ainda contem informação suficiente para identificar ao escritor.

Dentro desse contexto, neste trabalho propomos utilizar características biométricas para realizar a tarefa de identificação pessoal utilizando verificação de assinaturas através do som que produz a caneta no momento que uma pessoa assina. O sistema proposto utiliza um superfície lisa e sólida, na qual se encontra preso um microfone para a aquisição do som produzido pela fricção entre a ponta rígida da caneta e o papel sobre a superfície. Este procedimento nos leva a um sistema de aquisição mais simples e mais econômico que qualquer sistema que precise adquirir imagens para o reconhecimento. Este som que se produz no ato de assinar está relacionado com a dinâmica da assinatura, permitindo que parâmetros extraídos dele sejam uma forma de identificação pessoal.

É importante notar que o sistema proposto e utilizado neste trabalho, além das vantagens já citadas, é do tipo não intrusivo e permite obter a dinâmica da assinatura sim qualquer registro dela em forma visual. Isto é feito assinando na superfície citada sem o papel.

1.3 Objetivos do Trabalho

O objetivo desta dissertação é apresentar uma metodologia que utilize medidas biométricas para realizar a tarefa automática de autenticação pessoal. Esta tarefa, é feita através da verificação automática de assinaturas utilizando a denominada Emissão Acústica (EA). Em nosso caso, essa metodologia esta baseada em medidas biométricas feitas a partir do som que se produz no momento que uma pessoa assina. Isto é, verificar que uma pessoa é realmente quem diz ser, capturando e processando o som que produz a ponta rígida da caneta sobre a superfície do papel no ato de assinar.

1.4 Estrutura da Dissertação

Com o propósito de apresentar de forma estruturada o trabalho realizado, esta dissertação estará constituída dos seguintes capítulos:

1. Introdução
2. Conceitos de Biometria e Reconhecimento de Padrões
3. Verificação de Assinaturas
4. A Metodologia de Verificação de Assinaturas baseada em Emissão Acústica
5. Resultados
6. Conclusões

No capítulo 2 são apresentados, de forma geral, os conceitos básicos das áreas de biometria e reconhecimento de padrões, revisando os principais métodos biométricos de identificação pessoal. Além disso é feita uma introdução sobre reconhecimento de padrões e dos conceitos referentes à extração de características e classificadores. A seguir são apresentados os componentes de um sistema automático de verificação/identificação de assinaturas.

No capítulo 3 é discutido o processo da autenticação por assinaturas, com base em suas características e tipos de falsificações. Ainda nesse capítulo faz-se uma revisão da literatura sobre a verificação de assinaturas utilizando EA.

No capítulo 4 é descrito o método de verificação de assinaturas proposto, detalhando o sistema de aquisição do som da assinatura e as diferentes características que são extraídas deste som, bem como as etapas de pré-processamento do som e classificação.

No capítulo 5, os resultados dos experimentos realizados são apresentados e discutidos. Finalmente, no capítulo 6 são apresentadas as conclusões do trabalho e propostas para a continuidade desta pesquisa.

Capítulo 2

Conceitos de Biometria e Reconhecimento de Padrões

Na sociedade moderna, as pessoas que se conectam através de uma rede de comunicação para compartilhar informação buscam sempre a maior segurança possível. Nos dias de hoje, há uma necessidade crescente para determinar ou verificar a identidade de uma pessoa, com o fim de obter uma autorização para acessar a produtos e serviços.

O propósito de qualquer sistema de autenticação pessoal é garantir a segurança do usuário legítimo. Exemplos de tais aplicações incluem acesso seguro a edifícios, sistemas de computação, telefones celulares, entre outros. Na ausência de um reconhecimento pessoal robusto, estes sistemas se tornam vulneráveis à vontade de um impostor. Uma maneira para robustecer o sistema é utilizar as características biométricas para o reconhecimento pessoal.

2.1 Biometria

O que é Biometria?

Biometria é o ramo da ciência que estuda a mensuração dos seres vivos [10]. Reconhecimento Biométrico, ou simplesmente biométrica, refere-se ao reconhecimento automático de indivíduos baseado em suas características fisiológicas e/ou características de comportamento [11]. Exemplos de características fisiológicas são: padrão da íris, impressões digitais, formas dos contornos da mão e da face, etc, e os exemplos de características de comportamento são:

tipo da escrita, assinatura, dinâmica da digitação, voz, etc..

Uma característica fisiológica é uma propriedade física relativamente estável tais como as impressões digitais, geometria da mão, padrão da íris, entre outras. Esse tipo de característica é basicamente imutável. Por outro lado, uma característica de comportamento é mais um reflexo de atitudes psicológicas do indivíduo, por exemplo, a maneira como se digita nos teclados, a maneira de falar ou a maneira de assinar que são as características de comportamento mais utilizadas para autenticação.

Qualquer característica humana fisiológica ou comportamental pode ser uma característica biométrica sempre que satisfaça os seguintes requerimentos:

- Universalidade: todas as pessoas devem ter a característica.
- Distinguibilidade: a característica deve distinguir de maneira única a cada pessoa.
- Permanência: a característica não deve sofrer mudança com o tempo.
- Coletividade: a característica pode ser medida qualitativamente.

As características de comportamento tendem a variar com o tempo. Por esse motivo, muitos sistemas biométricos permitem que sejam feitas as atualizações de seus dados biométricos de referência à medida que esses vão sendo utilizados.

Usando biometria é possível confirmar ou estabelecer a identidade de um indivíduo baseado no "quem ele é" (prova por biometria), em lugar de "o que ele possui" (prova por posse) ou "o do que ele se lembra" (prova por conhecimento).

A figura 2.1 apresenta um diagrama dos métodos de autenticação associados a tecnologias biométricas. Ela também mostra alguns exemplos de tipos de biometrias para as características fisiológicas e de comportamento.

As diferenças entre métodos de comportamento e fisiológicos são importantes por vários motivos. Primeiro, o grau de variação intra-pessoal numa característica física é menor do que em uma característica de comportamento. Exemplificando, isto significa que, com exceção de algum ferimento, suas impressões digitais são as mesmas ao longo da sua vida. Uma assinatura, por outro lado, é influenciada tanto por fatores físicos como por fatores emocionais. Assim,

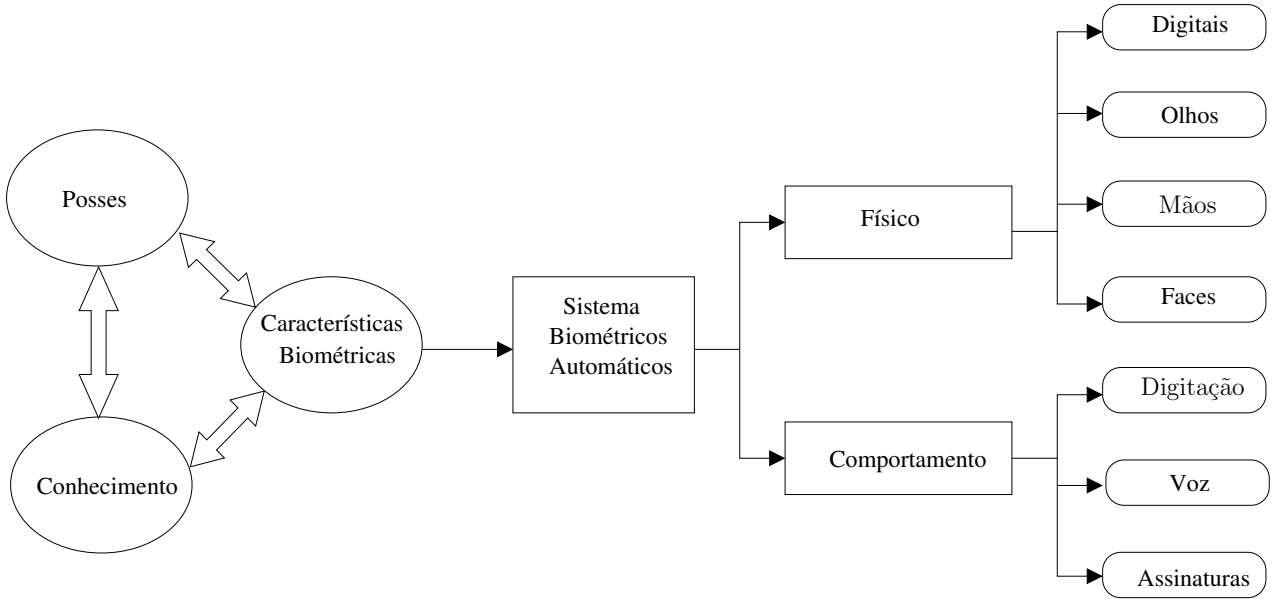


Fig. 2.1: Métodos de autenticação associados aos sistemas baseados em características biométricas

sistemas baseados em comportamento tem um grande trabalho em ajustar as variações intra-pessoais. Por esse motivo, oferece melhor desempenho um sistema que, por exemplo, guie o usuário a colocar a palma de sua mão sempre em determinada posição, que implementar um algoritmo que traduza o estado emocional de uma pessoa [6].

Em um sistema prático que utiliza biometria para reconhecimento de pessoas, há alguns outros aspectos que devem ser considerados:

- Desempenho: refere-se à precisão no reconhecimento, os recursos utilizados para alcançar esta velocidade e precisão, assim como os fatores operacionais que as afetam.
- Aceitabilidade: aceitação das pessoas para utilizar um identificador biométrico em sua vida diária.
- Fraudabilidade: revela a fragilidade do sistema aos métodos fraudulentos.

Outra característica importante dos sistemas de reconhecimento biométrico é que o sistema deve ser capaz de determinar que a informação biométrica que está sendo amostrada pertence a uma pessoa que está viva e presente, evitando assim fraudes com a utilização de amostras artificiais. Como exemplo, pode ocorrer que diante de um sistema de verificação de locutor, um

indivíduo tente se passar por outro através da reprodução do som da voz de um dos usuários do sistema que tenha sido previamente gravada. Uma das soluções para este tipo de fraude é que os dispositivos de captura que fazem parte dos sistemas biométricos incluam meios para determinar se existe uma característica "viva". Um exemplo disso já pode ser encontrado em alguns sistemas de reconhecimento de faces. Neste caso, o sensor que faz a captura da imagem não é uma câmera de vídeo comum. Trata-se de dispositivo que além de capturar a imagem da face como um matriz de valores de intensidade de luz, capta também a distribuição de temperatura sobre as diferentes regiões do rosto. Dessa forma, ao apresentar uma foto comum como entrada para o sistema, mesmo que as características referentes à intensidade de luz casem com as da base de dados, aquelas referentes à distribuição de temperatura com certeza serão diferentes e, portanto, o resultado do pedido de autenticação de identidade será falho.

2.1.1 Sistemas Biométricos

Um sistema biométrico é um sistema de reconhecimento de padrões que opera através da aquisição de dados biométricos de um indivíduo, extraindo um conjunto de características a partir desses dados, e comparando-as com um conjunto de modelos em uma base de dados.

Um sistema biométrico pode ser classificado em relação à maneira como os dados de entrada são comparados junto à base de dados. Neste caso, duas categorias podem ser definidas:

- Verificação: Uma comparação "1:1" é realizada; a informação biométrica apresentada por um individuo é comparada com o *template* correspondente àquele individuo. Este tipo de verificação é chamado de *reconhecimento positivo*, ou seja, prevenção de uso da mesma identidade por várias pessoas.

- Identificação: Uma comparação "1:N" é realizada; a informação biométrica apresentada por um individuo é comparada com todos os *templates* ou um conjunto deles armazenados na base de dados. Este tipo de identificação é chamado de *reconhecimento negativo*, ou seja, prevenção de várias identidades serem usadas por uma mesma pessoa.

Nessas definições, *Template* é a representação das informações extraídas das amostras biométricas fornecidas pelo indivíduo no processo de cadastramento (*enrollment*). *Cadastramento* é o processo de coletar amostras biométricas de uma pessoa, e a subseqüente preparação

e armazenamento dos templates biométricos de referência que representam a identidade dessa pessoa [4].

Um sistema biométrico é composto por 4 módulos principais: Coleta de Dados (Sensor), Extrator de Características, Comparador e Armazenamento (Base de Dados) [11]. A figura 2.2 apresenta um diagrama de blocos para sistemas de verificação e identificação, assim como também para o cadastramento de usuários, o qual é um processo comum a ambos os sistemas.

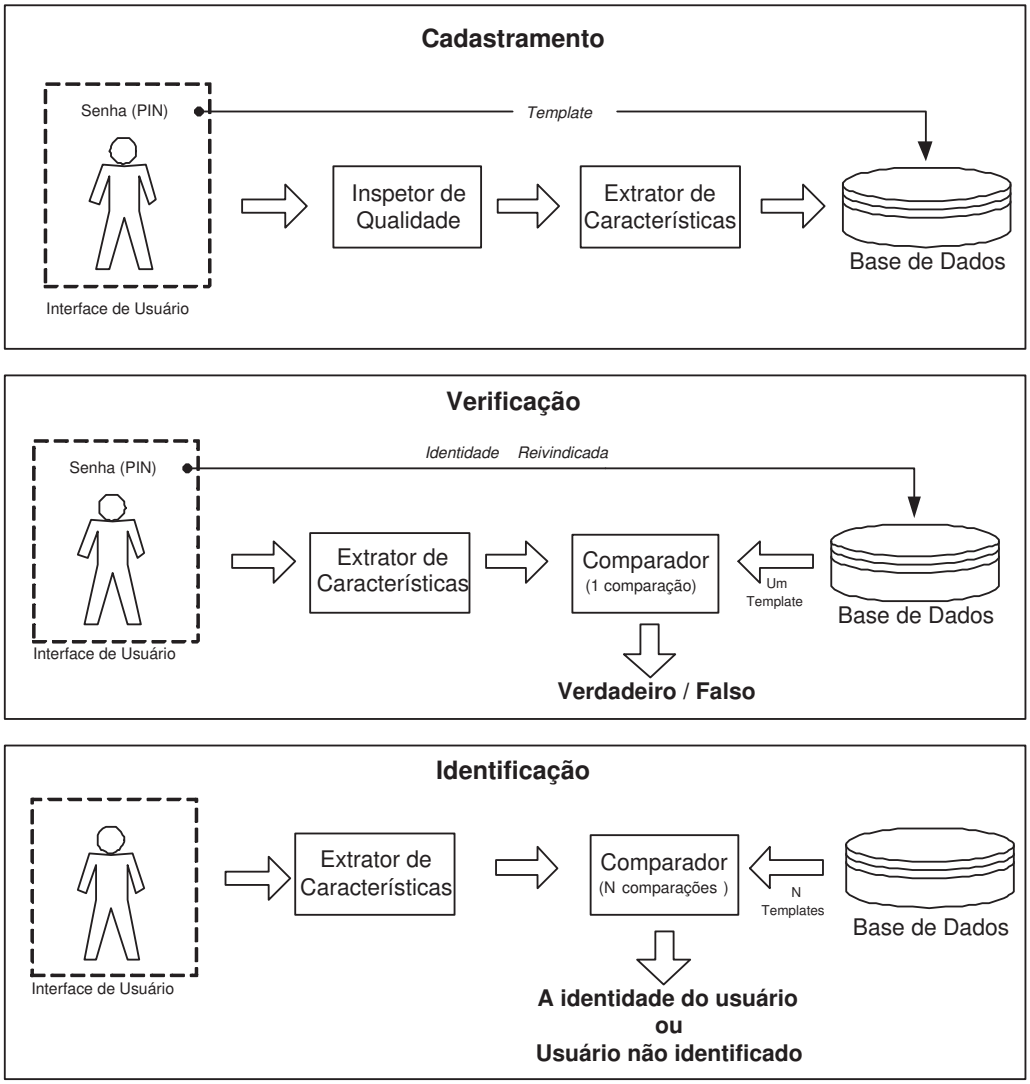


Fig. 2.2: Diagrama em blocos de cadastramento, verificação e identificação.

- Coleta de Dados: contém um dispositivo ou sensor que capta as amostras biométricas de um usuário e converte-as em um formato adequado (sinal eletrônico) para ser utilizadas

pelas outras partes do sistema. O desempenho de todo o sistema é afetado pela qualidade da amostra fornecida e pelo desempenho do próprio sensor ou dispositivo de coleta [3].

- Extrator de Características/Processamento de Sinais: é responsável por receber a amostra biométrica fornecida pelo subsistema de coleta de dados e convertê-la em uma forma adequada para o processamento pelo módulo de comparação. Este processamento pode aplicar uma análise da qualidade da amostra fornecida para determinar se ela pode ser passada adiante, ou aplicar uma filtragem para remover ruídos e outras informações que podem afetar o julgamento do módulo de comparação, ou ainda normalizar o sinal. Uma vez que a amostra tenha sido processada, este subsistema extrai características da amostra que são passadas ao comparador [3].

- Comparador: esta etapa do sistema faz a comparação da amostra biométrica apresentada com um *template* da base de dados. Ele verifica quanto ambas as amostras são similares para tomar a decisão, que identifica se a amostra apresentada pertence ou não ao proprietário do *template* selecionado da base de dados. Para tomar esta decisão, um limiar deve ser estabelecido para poder delimitar até que valor de similaridade é considerado como uma amostra autêntica ou uma amostra falsa [3].

- Armazenamento: este módulo mantém os *templates* dos usuários cadastrados no sistema biométrico. Ele disponibiliza a adição, eliminação ou atualização dos *templates* cadastrados. Ele pode conter para um único usuário um único *template* ou vários deles, dependendo para qual função o sistema foi desenvolvido. Além do *template*, outras informações também podem ser incluídas [3].

Cada *template* é armazenado com um identificador do usuário que permita determinar a que pessoa ele pertence. Estes *templates* podem ser armazenados em cartões de memória, em uma base de dados central ou em cartões magnéticos. O tipo de armazenamento se dará pela aplicabilidade prática a que se destina o sistema biométrico e pelo tamanho dos *templates* gerados [1].

Um típico processo de autenticação utilizando tecnologias biométricas consiste dos seguintes passos básicos [3]:

1. Capturar a(s) amostra(s) biométrica(s).

2. Avaliar a qualidade da amostra biométrica capturada e, se é necessário, recapturá-la.

3. Processar a amostra biométrica capturada para realizar a comparação;

4. Comparar a amostra biométrica processada com o *template* previamente armazenado na base de dados para determinar se a amostra biométrica processada pertence ao usuário do *template* armazenado. Esta comparação pode ser feita fazendo uso do processo de autenticação para a verificação ou identificação.

2.1.2 Medidas de Desempenho em Sistemas Biométricos

A captura de amostras de uma característica biométrica feita por um dispositivo ou sensor biométrico não é um processo que permite a extração de características perfeitamente precisas e idênticas após sucessivas tentativas. Duas amostras de uma mesma pessoa podem ser diferentes devido a muitas variáveis que influenciam nas características biométricas. Mudanças no estado físico, emocional e/ou psicológico do usuário; condições ambientais (temperatura, umidade, ruído no sensor); interação do usuário como sensor (posição do usuário), são algumas variáveis que afetam na captura das amostras biométricas. Estas variações nas características ocasionam erros de autenticação dos usuários no sistema biométrico, rejeitando usuários legítimos ou aceitando usuários impostores.

Um erro biométrico é caracterizado pela autenticação de uma amostra que não deveria ser autenticada. Para entender isto, suponha que exista um sistema que utiliza biometria para autenticar a identidade de uma pessoa (por exemplo: impressão digital). Uma amostra B desta biometria é capturada e dela se obtém um *template* que é armazenado em $Q = S(B)$. Uma nova amostra desta mesma biometria B' é obtida e é calculado seu *template* $Q' = S(B')$. Agora definimos duas hipóteses:

$$\begin{aligned} H_0 &: \texttt{B' vem da mesma pessoa que B}, \\ H_1 &: \texttt{B' não vem da mesma pessoa que B}. \end{aligned} \tag{2.1}$$

Em seguida, uma medida de semelhança dos *templates* é calculada por $s_m = Sim(Q, Q')$. Se $s_m \geq Td$ o sistema escolhe H_0 e a nova amostra é considerada verdadeira (a identidade reivindicada é correta) ou se $s_m < Td$ o sistema escolhe H_1 e a nova amostra é considerada falsa (a identidade reivindicada não é correta), onde Td é um limiar de decisão.

Quando $Q = Q'$, B e B' são chamados de par casado e quando $Q \neq Q'$ de par não-casado. A distribuição (histograma) gerada por pares casados é chamada de distribuição genuína e a de pares não-casados de distribuição impostora, conforme mostradas na figura 2.3, onde $TERR$ é a Taxa de Igual Erro, que será explicada em breve..

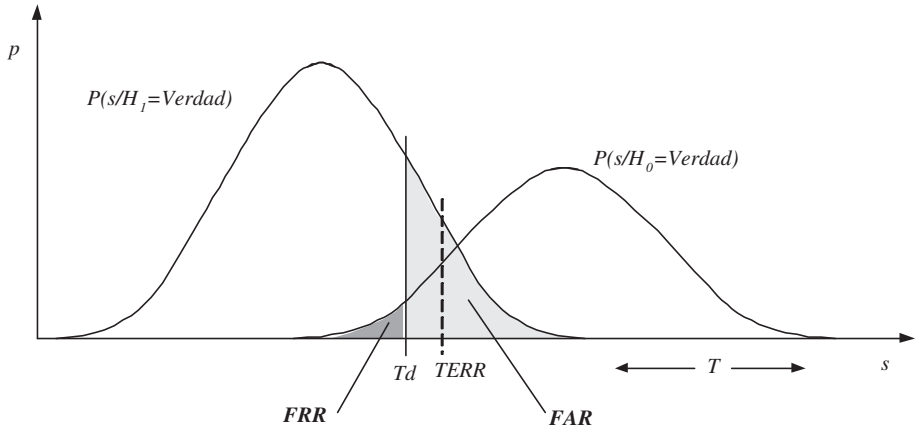


Fig. 2.3: Curvas de FAR e FRR

Se o sistema decide por H_0 quando H_1 é verdade resulta em uma falsa aceitação mas se o sistama decide por H_1 quando H_0 é verdade, o resultado é uma falsa rejeição.

- Falsa Aceitação - quando o sistema biométrico aceita um indivíduo impostor.
- Falsa Rejeição - quando o sistema biométrico rejeita um usuário genuino.

2.1.3 Erros nos Sistemas Biométricos

O desempenho de um sistema de autenticação pode ser avaliado por dois tipos de erro:

- Erro Tipo I, também conhecido como Taxa de Falsa Aceitação FAR (*False Acceptance Rate*) ou FMR (*False Match Rate*).

- Erro Tipo II, também conhecido como Taxa de Falsa Rejeição FRR (*False Rejection Rate*) ou FNMR (*False Non-Match Rate*)

Taxa de Falsa Aceitação (FAR), é a probabilidade de um sistema biométrico aceitar um indivíduo incorretamente ou não rejeitar um impostor. A FAR pode ser calculada como:

$$\texttt{FAR} = \frac{\texttt{NFA}}{\texttt{NIIA}}, \tag{2.2}$$

onde, NFA é o número de falsas aceitações e NIIA é o número de tentativas de identificação de impostor.

Taxa de Falsa Rejeição (FRR), é a probabilidade de um sistema biométrico rejeitar um usuário cadastrado legitimo. A Taxa de Falsa Rejeição pode ser calculada como segue:

$$\texttt{FRR} = \frac{\texttt{NFR}}{\texttt{NEIA}}, \tag{2.3}$$

onde, NFR é o número de falsas rejeições e NEIA é o número de tentativas de identificação de usuário legitimo.

Na figura 2.3, o FAR é a área baixo da função densidade H_1 à direita do limiar e o FRR é a área baixo da função densidade H_0 à esquerda do limiar.

$$FAR\,(Td) \;=\; P\,(s \geq Td\,|H_1 = true\,) = \left[1 - \int_0^{Td} p\,(s\,|H_1 = true\,)\,ds\right] \tag{2.4}$$

$$FRR\,(Td) \;=\; P\,(s < Td\,|H_0 = true\,) = \int_0^{Td} p\,(s\,|H_0 = true\,)\,ds. \tag{2.5}$$

Taxa de Erro Igual EER (*Equal Error Rate*), é o ponto no qual os valores de FRR e FAR são iguais, onde as áreas baixo as duas curvas na figura 2.3 são iguais. Este ponto é importante, pois especifica a separabilidade que o sistema oferece entre os acessos permitidos e os não-permitidos.

Em lugar de mostrar as taxas de erro em termos de densidades de probabilidade como em figura 2.3, é desejável apresentar a precisão do sistema utilizando uma curva Característica Operacional do Receptor (ROC). A ROC é um desenho da taxa FAR vs. FRR para vários

valores do limiar (ver fig.2.4).

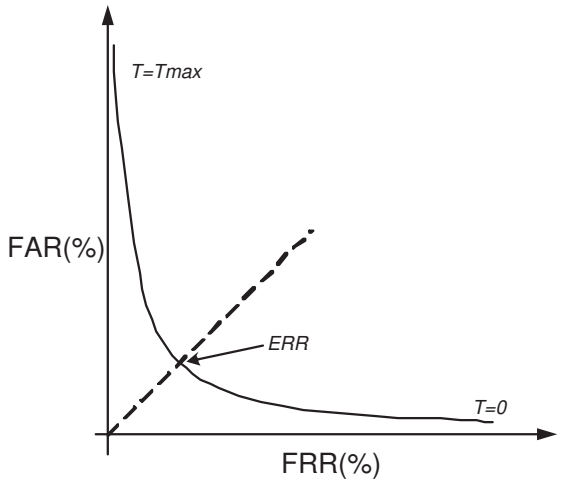


Fig. 2.4: Curva Característica Operacional do Receptor (ROC)

Numa situação ideal o valor de EER é igual a zero o que significa que as curvas de FAR e FRR estão completamente separadas, como conseqüência não haverá nenhum caso de rejeição de usuários legítimos ou de aceitação de usuários impostores. Este é um objetivo que todos os pesquisadores buscam em um sistema de autenticação. Em uma aplicação real, um sistema de autenticação raramente consiga operar neste ponto.

2.2 Reconhecimento de Padrões

A facilidade com que as pessoas reconhecem um rosto, entendem as palavras faladas, lêem os caracteres escritos, identificam suas chaves no bolso pelo tato, e decidem que uma maça é madura pelo cheiro, contradiz com o assombrosamente complexo processo que é a base destes atos de reconhecimento de padrões.

Para começar com o estudo de reconhecimento de padrões, inicialmente precisa se saber é, o que é um padrão. Padrão é: "Modelo oficial de pesos e medidas; aquilo que serve de base ou norma para a avaliação de qualidade ou quantidade; qualquer objeto que serve de modelo à feitura de outro; modelo, exemplo, protótipo, arquétipo" [10]. Por exemplo, na biometria, um padrão poderia ser a imagem de uma impressão digital, uma palavra manuscrita, uma face humana ou um sinal de voz.

Os seres humanos são capazes de reconhecer dígitos, letras ou figuras, sejam essas pequenas,

grandes, manuscritas, inclinadas ou rotacionadas. Essas formas podem estar dispersas no fundo de uma imagem ou parcialmente ocultas entre outras informações e, mesmo assim, tem-se a habilidade de reconhecê-las. O melhor reconhecedor de padrões é o ser humano, mesmo que ainda não saibamos com exatidão como a tarefa do reconhecimento ocorre internamente em nosso cérebro.

O reconhecimento de padrões é o estudo de como as máquinas podem observar o ambiente que as rodeia, aprender a distinguir padrões de seu interesse neste meio e tomam decisões razoáveis sobre as categorias a que pertencem tais padrões [12].

O reconhecimento, descrição, classificação e agrupamento de padrões feitos automáticamente (por máquinas) são problemas importantes a serem resolvidos em uma variedade de disciplinas científicas tais como biologia, psicologia, medicina, visão computacional, inteligência artificial e sensoriamento remoto.

Ha duas maneiras em que se pode reconhecer/classificar um padrão:

1. Classificação supervisionada: na qual o padrão de entrada é identificado como membro de uma classe predefinida.

2. Classificação não supervisionada: na qual o padrão é assinado para uma classe não previamente definida.

Note que o problema de reconhecimento nesses casos está sendo definido em relação as tarefas de classificação ou categorização, onde as classes podem ser definidas a priori pelo projetista do sistema (no caso da classificação supervisionada) ou estão baseadas em um aprendizado feito sobre a similaridade dos padrões (no caso de classificação não supervisionada).

Conforme já dito, o projeto de um sistemas de reconhecimento de padrões envolve essencialmente a aquisicão de dados, o pré-preprocessamento e a tomada de decisão (classificação).

De maneira geral, um problema de reconhecimento de padrões bem definido e corretamente delimitado pode apresentar pequenas variações intraclasses (elementos de uma mesma classe) e grandes variações interclasses (elementos de classes diferentes) levando-nos a uma representação compacta do padrão e uma estratégia de decisão simples. Dependendo do tipo de problema

a ser resolvido dita a escolha do(s) sensor(es) para aquisição de dados, as técnicas de pré-processamento, os esquemas de representação e os modelos de tomada de decisão.

Uma das maneiras de fazer com que o sistema reconheça os padrões de entrada das classes para as quais terá que classificar é apresentar ao sistema um conjunto de exemplos (amostras), também chamado de conjunto de treinamento. A partir deste conjunto, o sistema poderá delimitar o espaço de características a que pertencem os padrões.

As abordagens mais conhecidas em reconhecimento de padrões são [12]:

- "Casamento" de modelos (*Template Matching*): É uma das mais simples abordagens para reconhecer padrões. O "casamento" é uma operação genérica usada para determinar a similaridade entre duas entidades de um mesmo tipo. Nesta abordagem o padrão a ser reconhecido é comparado com os modelos armazenados, observando todas as variações possíveis em termos de: translação, rotação e mudanças de escala. A medida de similaridade é freqüentemente uma correlação ou uma função de distância.

- Estatística: Nesta abordagem, o padrão é representado por um conjunto de características chamado vetor de características. Um conjunto de treinamento é fornecido para cada classe com o objetivo de estabelecer fronteiras de decisões em um espaço de características, separando os padrões pertencentes a classes diferentes. As fronteiras de decisão são determinadas pelas distribuições de probabilidade dos padrões pertencentes a cada classe.

- Sintática: Nesta abordagem o padrão é visto como uma composição de sub-padrões, que por sua vez são constituídos por sub-padrões mais simples, e esta subdivisão continua até que se alcance os sub-padrões elementares chamados de primitivas. O padrão é representado em termos da inter-relação entre estas primitivas. A abordagem sintática, por exemplo uma linguagem gramatical, faz uma analogia entre a estrutura dos padrões e a sintaxe de uma linguagem, onde os padrões são as sentenças, as primitivas são as letras do alfabeto, e as sentenças são geradas de acordo com uma gramática.

- Redes Neurais: Os modelos de redes neurais tentam usar alguns princípios de organização, como aprendizagem, generalização, adaptação e processamento distribuído, em uma rede

de grafos direcionados e com pesos, nos quais os nós são neurônios, e as extremidades direcionadas são conexões entre os neurônios de entrada e saída. As principais características desta abordagem são: a capacidade de aprender relações não-lineares complexas de entrada e saída, a utilização de procedimentos de treinamento seqüencial e a adaptação aos dados.

Em sistemas de reconhecimento de padrões, estas abordagens não são necessariamente independentes e existem vários trabalhos que aplicam sistemas híbridos.

Neste trabalho, é feita uma autenticação do tipo verificação, que idealmente, pode ser apresentada como um problema de classificação com duas classes: uma classe com as amostras pertencentes a uma mesma pessoa e outra classe com as amostras não-pertencentes a esta pessoa. Para conseguir este objetivo é utilizada uma abordagem estatística, associando um padrão de entrada a uma determinada classe.

2.3 Modelo de Classificação

Seja um conjunto de objetos os quais estão classificados em diferentes classes e um objeto aleatório a ser classificado. Uma forma de classificar este objeto é fazer uma comparação dele com os modelos padrões que representam cada uma das classes e então escolher aquele que apresente o melhor "casamento". Às vezes, o que torna difícil a classificação correta do objeto de entrada é seu elevado grau de variabilidade com relação à classe que realmente pertence. Uma das maneiras de minimizar esse problema é representar um padrão de entrada através de características que sejam discriminantes.

2.3.1 Características

Para classificar um objeto ou evento é possível fazer medidas de suas propriedades ou características. Por exemplo, para classificar um objeto pode ser útil saber: sua área, seu perímetro; pode-se também medir seu grau de simetria com relação ao eixo horizontal, entre outras.

Na maioria das vezes, pode-se medir um conjunto fixo de características de qualquer objeto ou evento que se deseja classificar. Seja X um vetor de características n-dimensional composto

por $x_1, x_2, ..., x_n$, para c classes de padrões $w_1, w_2, ..., w_c$. Pode-se pensar que X representa um ponto em um espaço de n características, por exemplo, um espaço tri-dimensional, figura 2.5.

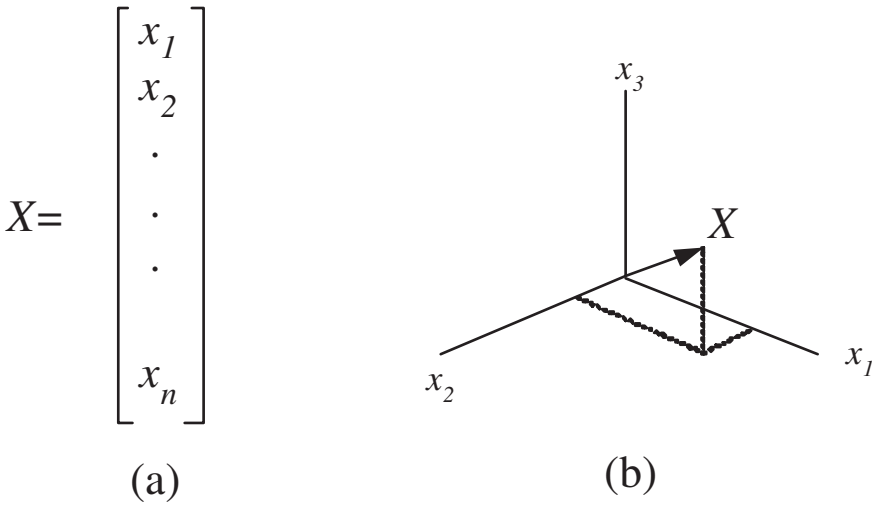


Fig. 2.5: Características de classificação. a) Vetor de características X. b) Representação no espaço tri-dimensional.

Na figura 2.6 é mostrado um diagrama de blocos de um sistema de reconhecimento de padrões simplificado, onde o módulo de extração de características processa a informação de entrada com o objetivo de determinar os valores numéricos para o conjunto de n características, $x_1, x_2, ..., x_n$, que compõe o vetor de características X. A seguir, o módulo de classificação recebe o vetor X e associa-o a uma de suas classes: $w_1, w_2, ..., w_c$.

O limite conceitual entre extração de característica e classificação adequada é um pouco arbitrário. Um extrator de característica ideal produziria uma representação que faz o trabalho do classificador trivial e reciprocamente, um classificador ideal não precisaria da ajuda de um extrator de característica sofisticado [13].

A implementação do módulo de extração de características é dependente do problema. Por exemplo, um bom extrator de características para identificar impressões digitais não é útil na classificação de fotografias de células sangüíneas.

Em geral, um módulo extrator de características ideal deveria produzir o mesmo vetor de características X para todos os padrões que pertencem à mesma classe w_i, e diferentes vetores de características para padrões de classes diferentes w_j. Na prática, dados de entrada diferentes

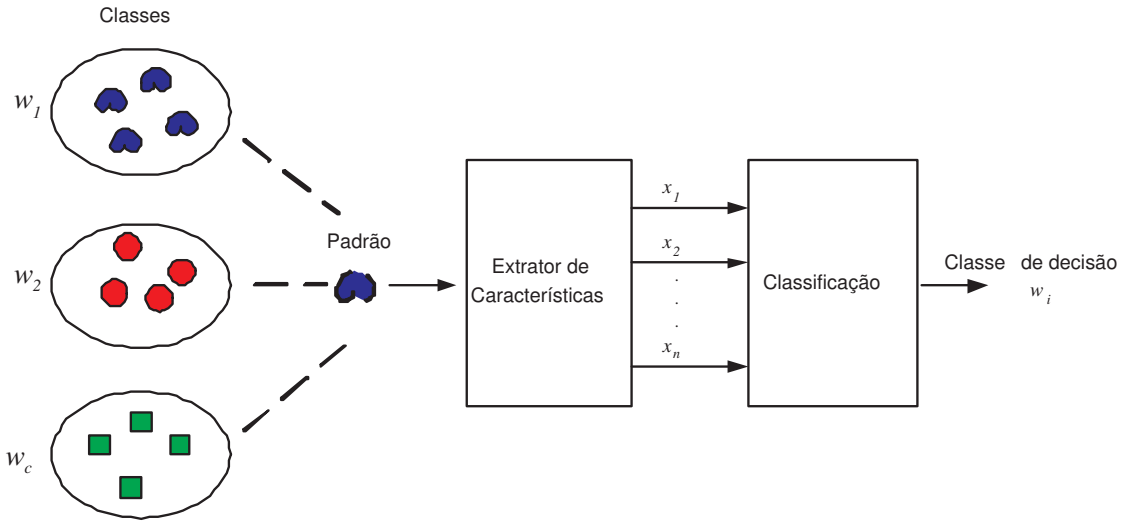


Fig. 2.6: Módulos de extração e classificação de características.

no módulo de extração de características produzem diferentes vetores de características, porém, espera-se que a variabilidade intraclasse seja pequena.

2.3.2 Um Classificador Simples

O modelo mais simples para reconhecimento de padrões é o casamento de padrões ou *template matching* o qual é uma operação genérica em reconhecimento de padrões que é usada para determinar a similaridade entre duas entidades do mesmo tipo. O padrão a ser reconhecido é comparado com os modelos previamente armazenados na base de dados do sistema, observando todas as variações possíveis em termos de: translação, rotação e mudanças de escala. A medida de similaridade é freqüentemente realizada por uma correlação ou uma função de distância [12].

O casamento de modelos faz parte das abordagens de decisão teórica que se baseiam na utilização de *funções de decisão* (ou discriminantes). Seja $X = (x_1, x_2, ..., x_n)^T$ um vetor de características n-dimensional, para c classes de padrões $w_1, w_2, ..., w_c$. O problema básico é encontrar c funções de decisão $d_1(X), d_2(X), ..., d_c(X)$, com a propriedade de que, se o padrão X pertence à classe w_i, então:

$$d_i(X) > d_j(X) \qquad j = 1, 2, ..., c; j \neq i. \tag{2.6}$$

Ou seja, um padrão desconhecido X pertencerá à i-ésima classe de padrões se a substituição

de X em todas as funções de decisão for tal que $d_i(X)$ tenha o maior valor numérico. Empates são resolvidos arbitrariamente. O limite de decisão que separa as classes w_i e w_j é dada pelos valores de X para os quais $d_i(X) = d_j(X)$ ou, equivalentemente, pelos valores de X tais que:

$$d_i(X) - d_j(X) = 0 \tag{2.7}$$

É comun identificar o limite de decisão entre duas classes pela função $d_{ij}(X) = d_i(X) - d_j(X) = 0$. Por tanto, $d_{ij}(X) > 0$ para padrões de classe w_i e $d_{ij}(X) <= 0$ para padrões de classe w_j [14].

Classificador de distância mínima

Suponha que cada uma das c classe de padrões $w_1, w_2, ..., w_c$ seja representada por um vetor protótipo ou médio (*template*):

$$m_j = \frac{1}{N_j} \sum_{X \in w_j} X \qquad j = 1, 2, ..., c, \tag{2.8}$$

onde N_j é o número de vetores de padrões da classe w_i e a soma é realizada sobre esses vetores. Uma maneira de definir a pertinência de uma vetor padrão X desconhecido é atribui-lo à classe de seu protótipo mais próximo. A distância Euclidiana ou a de Hamming, pode ser usada para determinar a proximidade, reduzindo o problema à computação das distâncias:

$$D_j(X) = \|X - m_j\| \qquad j = 1, 2, ..., c, \tag{2.9}$$

onde $\|a\| = (a^T a)^{1/2}$ é a norma euclidiana. Atribuimos então X à classe w_i se $D_i(X)$ for a menor distância, ou seja, a menor distância implica no melhor casamento nessa formulação. Se:

$$\begin{aligned} \|X - m_j\|^2 &= (X - m_j)^T (X - m_j) \\ &= X^T X - m_j^T X - X^T m_j + m_j^T m_j \\ &= -2 \left[m_j^T X - \frac{1}{2} m_j^T m_j \right] + X^T X. \end{aligned} \tag{2.10}$$

Note que o termo $X^T X$ é o mesmo para todas as classes, ou seja para todo j. Para se encontrar o template m_j que minimiza $\|X - m_j\|$ é suficiente encontrar m_j que maximize a expressão entre colchetes da equação 2.10. Define-se a função de discriminante linear por:

$$d_j(X) = X^T m_j - \frac{1}{2} m_j^T m_j \qquad j = 1, 2, ..., c, \tag{2.11}$$

e atribui X à classe w_i se $d_i(X)$ for o maior valor numérico. Esta formulação está de acordo com o conceito de função de decisão, como foi definido na equação 2.11. A partir das equações 2.11 e 2.7, pode-se ver que o limite de decisão entre as classes w_i e w_j para o classificador de distância mínima é:

$$\begin{aligned} d_{ij} &= d_i(X) - d_j(X) \\ &= X^T(m_i - m_j) - \frac{1}{2}(m_i - m_j)^T(m_i - m_j) = 0. \end{aligned} \tag{2.12}$$

A superfície dada pela Equação 2.12 é a bisseção perpendicular do segmento de linha entre m_i e m_j. Para $n = 2$ a bisseção perpendicular é uma linha, para $n = 3$ é um plano, e para $n > 3$ é chamado de *hiperplano* [14].

A figura 2.7 apresenta um diagrama de blocos de um classificador de distância mínima.

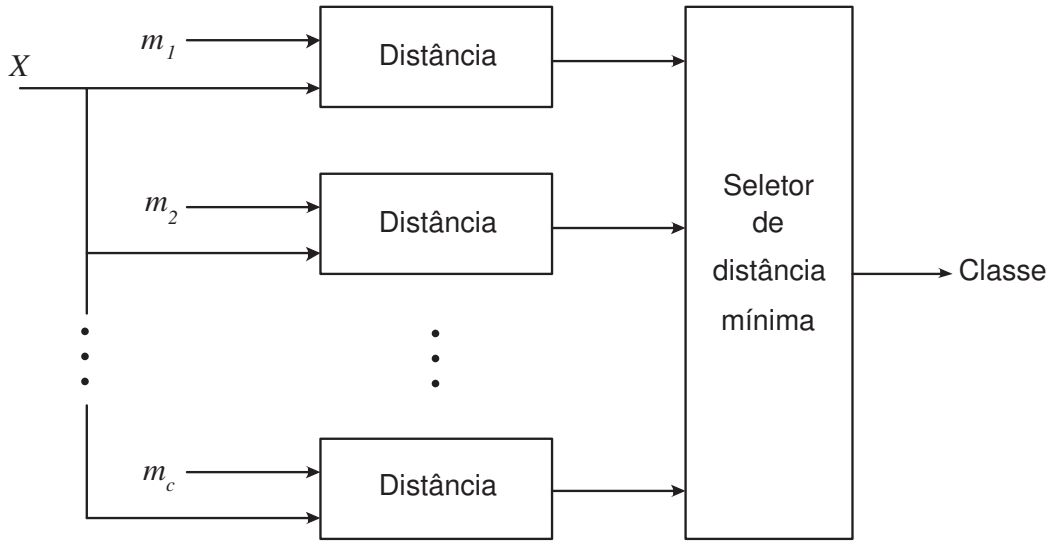


Fig. 2.7: Diagrama em blocos de um classificador de distância mínima.

2.3.3 Classificador Estatístico

Seja um padrão representado por um vetor $\mathtt{X} = (x_1, x_2, ..., x_d)$ com d características, ele pertence a uma das c possíveis classes $w_1, w_2, ..., w_c$. Supõe-se que cada uma das características apresente uma densidade de probabilidade ou função massa (dependendo das características serem contínuas ou discretas) condicionada a cada classe. Assim, que um padrão `x` pertença à classe w_i é visto como uma observação extraída aleatoriamente a partir de uma função de probabilidade condicional de classe $P(x|w_i)$. É claro que o reconhecimento dos membros de uma classe pode ser efetuado baseado nas diferenças das características das c classes. Expressamos estas diferenças, assumindo que o vetor de medições `x` é obtido de uma das c populações descritas pelas probabilidades condicionais $P(x|w_1), P(x|w_2), ..., P(x|w_c)$. É importante enfatizar que não necessariamente assume-se que as $P(x|w_i)$ são conhecidas com anterioridade. De fato, o coração de um modelo realista de reconhecimento de padrões é a habilidade de aprender propriedades das $P(x|w_i)$, observando as anteriores amostras obtidas desde algumas populações estatísticas [15].

Por exemplo, suponha que existem um conjunto de objetos que pertencem a uma de duas classes w_1, w_2; estes objetos são estocados aleatoriamente. Um objeto é pego, e se tem que decidir a que classe ele pertence; para isso se conhece que $P(w_1) = 0.8$ e $P(w_2) = 0.2$, onde $P(w_1)$ é chamada de *probabilidade a priori* e é a probabilidade do objeto pertencer à classe w_1. $P(w_2)$ é a probabilidade do objeto pertencer à classe w_2. Se a decisão de definir a que classe pertence o objeto deve ser feita baseada só nesta informação, então tem sentido utilizar a seguinte regra de decisão [16]:

Decidir por w_1 se $P(w_1) > P(w_2)$, no outro caso decidir por w_2.

Esta decisão depende dos valores das probabilidades a priori. Se $P(w_1)$ é muito maior que $P(w_2)$, a decisão a favor de w_1 será certa a maioria do tempo. Se $P(w_1) = P(w_2)$, obtem-se só uma chance de cinqüenta por cento de estar certos. Esta regra, faz sentido quando se julgar só um objeto, mas não é correta para julgar muitos objetos, afinal de contas, sempre se tomará a mesma decisão embora se saiba que ambos os tipos de objetos podem se apresentar. Desta forma ao escolher w_1 estaria certo só o 80% das vezes, mas a regra de decisão está baseada na pouca informação disponível das classes e do objeto. Este exemplo é simples, mas pode

ser usado para entender a idéia básica do problema de reconhecimento de padrões envolvendo alguma informação probabilística.

Na maioria das circunstâncias há alguma informação acerca do objeto que se vai tentar classificar. Por exemplo se poder ter a distribuição da probabilidade de uma ou umas das características do objeto; se w_1 representa uma classe, X é uma variável aleatória contínua que representa as medições feitas das características do objeto. Então a expressão $p(x|w_1)$ é a função densidade de probabilidade para $\mathtt{x}$, dado que o seu estado de natural é w_1 (probabilidade condicional de estado). i.e. a probabilidade daquela medida condicional do objeto venha da classe w_1. Então a diferença entre $p(x|w_1)$ e $p(x|w_2)$ descreve a diferença das características entre populações de w_1 e w_2 [16].

Em um típico problema, pode-se conhecer o estar em capacidade de calcular a densidade condicional $p(x|w_i)$ para qualquer i em w_1 ou w_2 e também conhecer a probabilidade $P(w_1)$ e $P(w_2)$, agora é preciso procurar uma fórmula que diga acerca da probabilidade de um objeto pertencer à classe w_1 ou w_2, dado que se observa um certo valor de $\mathtt{x}$. Este uso de probabilidades condicionais para nomear um objeto a uma classe é a base de toda a teoria de classificação e é conhecida como regra de Bayes [17].

$$P\left(w_j\left|x\right.\right) = \frac{p\left(x\left|w_j\right.\right)P\left(w_j\right)}{p\left(x\right)} \tag{2.13}$$

Isto significa, que usando a informação apriori, pode-se calcular a probabilidade a posteriori que é a probabilidade que dado um conjunto de medições $\mathtt{x}$, o objeto pertença a classe w_j. Assim, observa-se um certo valor de $\mathtt{x}$ para um objeto, então por cálculo de $P\left(w_j\left|x\right.\right)$ pode-se decidir a que classe o objeto pertence se seu valor de probabilidade $P(w_1|x)$ é maior a $P(w_2|x)$

$$\textit{Decidir por}\, w_1,\ \textit{Se}\, P\left(w_1\left|x\right.\right) > P\left(w_2\left|x\right.\right);\ \textit{caso contrário decidir por}\ w_2, \tag{2.14}$$

dado que $p(x)$ está em ambos lados da comparação, a regra é equivalente a dizer:

$$\textit{Decidir por}\, w_1,\ \textit{Se}\, p\left(x\left|w_1\right.\right)P\left(w_1\right) > p\left(x\left|w_2\right.\right)P\left(w_2\right);\ \textit{caso contrário decidir por}\ w_2. \tag{2.15}$$

Em um caso mais geral, há muitas diferentes características que são medidas, $\mathtt{X} = (x_1, x_2, ..., x_d)$

e existem diferentes classes que representam os c distintos estados da natureza. Então a fórmula de Bayes, pode ser calculada da seguinte forma:

$$P\left(w_{\mathrm{j}}\left|x\right.\right)=\frac{p\left(x\left|w_{j}\right.\right)P\left(w_{j}\right)}{p\left(x\right)}\quad j=1,2,...,c. \tag{2.16}$$

Mas agora, $p(x)$ pode ser calculado usando a lei de probabilidade total:

$$p\left(x\right)=\sum_{\mathrm{j=1}}^{\mathrm{c}}p\ \left(x\left|w_{j}\right.\right)P\left(w_{j}\right)\ . \tag{2.17}$$

Ao medir o vetor de características `x`, pretende-se classificar um objeto dentro da classe w_j, se $P(w_j|x)$ é o máximo de todas as densidades de probabilidade para $j = 1, 2, ..., c$.

$$P\left(w_{i}\left|x\right.\right)>P\left(w_{j}\left|x\right.\right)\ para\,todo\ j\neq i. \tag{2.18}$$

Uma forma fácil de estimar $P(w_j|x)$ é assumir que a probabilidade de obter uma dada medição dentro de uma classe vem de uma distribuição conhecida. $p(x|w_j)$ é em princípio algo que pode ser determinado experimentalmente. Tudo o que se tem que fazer é construir um histograma da freqüência relativa da ocorrência dos diferentes valores de `x` para cada classe [17]. Por exemplo no caso de existir duas classes $p(x|w_1)$ e $p(x|w_2)$ poderia aparecer assim:

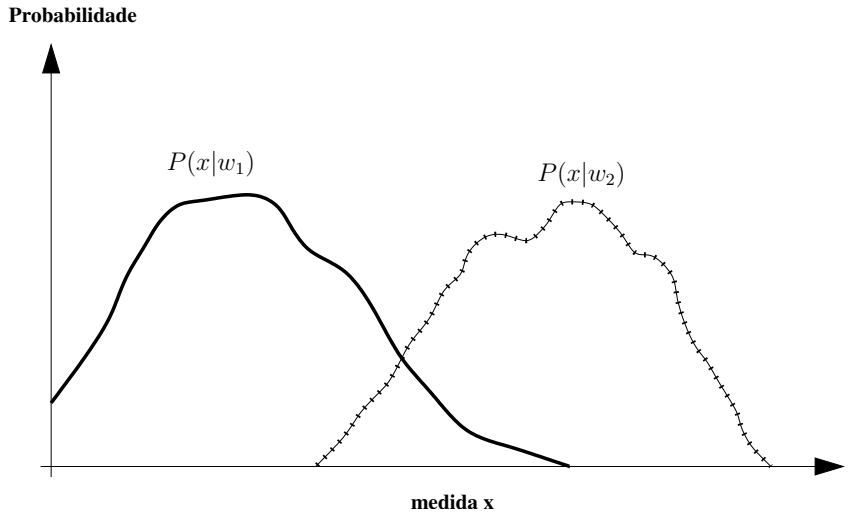


Fig. 2.8: Histograma de freqüência relativa da ocorrência de valores de `x`.

A Densidade Normal

A estrutura de um classificador de Bayes é determinada pelas densidades condicionais $p(x|w_i)$ como também pelas probabilidades a priori. Das várias funções de densidade que se são utilizadas em reconhecimentos de padrões, nenhuma recebeu mais atenção que a normal multivariável ou densidade Gaussiana [16].

A densidade normal multivariável é tipicamente um modelo apropriado para a maioria de problemas de reconhecimento de padrões, onde o vetor de características `x` para uma dada classe w_i são valores contínuos, versões mediamente corrompidas de um único vetor médio μ_i. Neste caso, a densidade condicional $p(x|w_j)$ e a probabilidade a priori $P(w_j)$ são normalmente distribuídas.

Assim a forma mais comum de distribuição assumida por $p(x|w_j)$ é a distribuição normal com média μ e variância σ^2, que são parâmetros que definem completamente a distribuição normal.

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}. \tag{2.19}$$

A distribuição normal é útil por as seguintes razões: esta ocorre freqüentemente na prática exatamente ou aproximadamente; é uma boa aproximação para um amplo conjunto de outras distribuições; é a regra base da classificação simples como também é uma das poucas distribuições que é fácil para trabalhar [17].

A densidade normal multivariável geral é dada por um vetor de médio d-dimensional e uma matriz de covariância $d \times d$:

$$p(x) = \frac{1}{(2\pi)^{d/2}\sqrt{|\sum|}}\, e^{-\frac{1}{2}(x-\mu)^T \sum^{-1}(x-\mu)}. \tag{2.20}$$

O vetor médio é uma coleção de médias simples μ_i, onde a i-esima média representa a média da i-esima característica que é medida. A matriz de covariância $\sum$ é similar à variância no caso univariável. Os elementos da diagonal representam a variância para diferentes características que são medidas. Os elementos fora da diagonal representam a covariância entre duas diferentes características.

$$\mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_d \end{bmatrix} \qquad \sum = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1d} \\ \vdots & \ddots & \vdots \\ \sigma_{d1} & \cdots & \sigma_{dd} \end{bmatrix} \tag{2.21}$$

Neste trabalho se vai desenvolver um método de verificação pessoal com reconhecimento supervisionado que utilizará como dados de entrada uma característica comportamental (assinatura) que se fundamenta nos conceitos explicados neste capítulo.

Capítulo 3

Visão Geral da Verificação de Assinaturas

A verificação de assinatura pertence a uma classe específica de processadores automáticos de escrita, que realizam a comparação de uma assinatura de teste com uma ou mais assinaturas de referência que foram coletadas quando um usuário cadastrou-se no sistema. Verificar uma assinatura requer a extração de informação específica do sinal da entrada ou imagem, independentemente de seu conteúdo. Esta informação tem que ser quase invariante no tempo e efetivamente discriminante.

3.1 A Assinatura Humana

Assinar é "o ato ou efeito de subscrever o próprio sinal ou nome num documento" e assinatura é "o estilo, marca ou sinal que permite identificar a autoria de algo" [10]. A assinatura é tradicionalmente a forma mais confiável de autenticar a identidade de um indivíduo, já que contém características únicas da escrita que são o reflexo de um conjunto de fatores físicos e psicológicos do assinante durante a escrita. Com base nessas características, especialistas procuram quantificá-la com o objetivo de definir, de maneira consistente, se uma certa assinatura foi feita por um indivíduo genuíno ou por outra pessoa [18]. Por este motivo a assinatura manuscrita é exigida em transações financeiras, de forma que o indivíduo ateste o conhecimento e sua concordância com os termos de um documento.

Muitas vezes as assinaturas não correspondem à escrita legível do nome de uma pessoa. O que ocorre, na prática, é uma junção entre componentes da escrita manuscrita com uma série de traços estilísticos, que tem por objetivo individualizar a assinatura através de um sinal gráfico. Mesmo que a assinatura seja totalmente ilegível, ela é suficiente para ser reconhecida e determinar a que indivíduo pertence [19].

Apesar de ter uma certa estabilidade, a escrita não é um processo preciso. A assinatura de uma pessoa não apresenta uma forma única e bem definida, pois aparecem variações em seus traços a cada momento que ela é reproduzida, de fato, grandes diferenças podem ser observadas em assinaturas de acordo com o país, idade, tempo, hábitos, estado psicológico ou mental e condições físicas. A assinatura também evolui à medida que vai sendo feita repetidamente, aparecendo pequenas diferenças na sua forma e em seu estilo cada vez que o escritor a reproduz. Por isso constata-se que duas assinaturas verdadeiras de uma pessoa nunca são exatamente iguais, as pequenas alterações ou variações que apresentam são chamadas de variações intra-pessoais [20].

- As variações intra-pessoais ou intra-classes, são variações observadas dentro de uma mesma classe, entre espécimes de assinaturas genuínas de um mesmo autor;
- As variações inter-pessoais ou inter-classes, são diferenças que existem entre classes de autores distintos.

Em teoria, uma variação intra-classe deve ser a mínima possível e uma variação inter-classes deve ser a máxima possível. Na prática, as classes não são bem separadas.

Quando um assinante reproduz sua assinatura mantendo um mesmo padrão de letras e de traços, sem a necessidade de um grande esforço de concentração, refere-se a ela como uma assinatura verdadeira. Utilizando um grupo de assinaturas verdadeiras é possível obter-se hábitos e qualidades da escrita que são úteis para determinar se a assinatura de uma pessoa pode ser aceitada ou rejeitada como autêntica.

Uma assinatura aceita pode ser de um dos seguintes tipos [24]:

- Autêntica, se é escrita pelo mesmo autor do modelo e se possui uma boa semelhnaça com o modelo de referência;

- Imitação, se é escrita por algum escritor que não é o autor e se possui semelhanças com o modelo de referência.

Igualmente, uma assinatura rejeitada pode ser de um dos seguintes tipos:

- Degenerada, se é escrita pelo mesmo autor do modelo e não é similar à assinatura de referência. O termo disfarçada é freqüentemente usado quando a degeneração é voluntária;
- Falsa, se é escrita por algum escritor que não é o autor do modelo de referência.

Partindo da utilização de um conjunto de assinaturas verdadeiras, especialistas se valem das características de forma, de movimento e de possíveis influências externas do meio para estabelecer sua autenticidade. Esses especialistas definem dois tipos de fatores que influenciam de maneira direta no processo de escrita de uma assinatura. O primeiro fator se refere às influências internas, que são qualidades inerentes do assinante, e o segundo fator, diz respeito às influências externas ou do meio, que atuam no momento de realizar a assinatura [22].

3.2 Análise de Assinaturas

Na análise de assinaturas procura-se características de sua escrita que sejam específicas como: os toques inicias e finais da assinatura, a rapidez da execução, a forma da letra, os alinhamentos vertical e horizontal, a razão de distância entre as várias letras, o espaçamento entre palavras, a qualidade das linhas da escrita, o tamanho total da assinatura e os traços de estilo, entre outras características [23].

Para o mesmo indivíduo, a maioria dessas características variam levemente e se enquadram dentro de um limite máximo. É por causa disso que, quando se tenta determinar a legitimidade de uma assinatura, o examinador do documento tentará conduzir a verificação pela comparação da assinatura em questão, com um conjunto de assinaturas verdadeiras. Esse conjunto deve, de preferência, ter sido adquirido de forma semelhante e mais ou menos na mesma época da assinatura que se deseja examinar [21].

O objetivo do análise utilizando um conjunto de assinaturas verdadeiras é determinar o grau de variação, ou seja, quanto podem variar as assinaturas de uma mesma pessoa. De maneira geral, o grau de variação será diferentemente afetado por influências internas e externas que

atuam no momento de assinar. Assim sendo, pode-se esperar um alto grau de variação quando influências internas e externas se apresentam, e uma variação pequena quando as influências externas e internas se mantém constantes.

Podemos citar como influências externas o tamanho do instrumento de escrita, seu peso e a maneira como sua ponta desliza sobre o papel. A posição do escritor em relação ao papel onde se assina possui uma relevância importante no grau de variação. É dizer, se o escritor se encontra em pé ou não, se está confortavelmente sentado ou de maneira incomoda, se está numa posição mais inclinada ou reta, se a mão, pulso, braço e cotovelo estão na posição normal de escrita ou não. O tipo de papel que se está utilizado pode também criar alguns problemas, como por exemplo, se a folha de papel em que se está escrevendo é muito áspera ou lisa. A inclinação do papel com relação à posição do escritor e o espaço disponível para realizar a assinatura são também outras influências [6] [24].

As influências internas são principalmente do tipo psicológico. O estado emocional, a pressa, a idade do assinante influem no resultado final da escrita da assinatura. O conjunto dessas influências resultam, em variações nos traços, na proporção entre as palavras, na inclinação e arredondamento das letras, na ornamentação, na legibilidade, etc.. Ocorrem também variações nas características ligadas à dinâmica dos movimentos da escrita, como a não uniformidade da velocidade e interrupções durante a escrita das palavras [6].

3.2.1 Assinaturas Pessoais

Quando uma pessoa deseja criar sua assinatura, começa a fazer uma série de esboços e depois de algum tempo de exercício constante, consegue reproduzir estes traços de forma automática. Esta assinatura irá tornar-se personalizada à medida que o tempo passa e ela for produzida regularmente.

Dependendo da aparência final que o indivíduo deseja dar para sua assinatura, ela pode ser classificada em dois grupos. O primeiro chamado assinatura cursiva ou contextual, é aquele no qual a aparência final reflete o próprio nome do escritor. Nesse caso, a pessoa assina utilizando quase exclusivamente a semântica do seu nome e do seu estilo padrão de escrita, como nas assinatura a) e b) da figura 3.1. O segundo grupo chamado de não contextual, é aquele em que a aparência final da assinatura toma a forma de um desenho estilizado que pode ou não conter

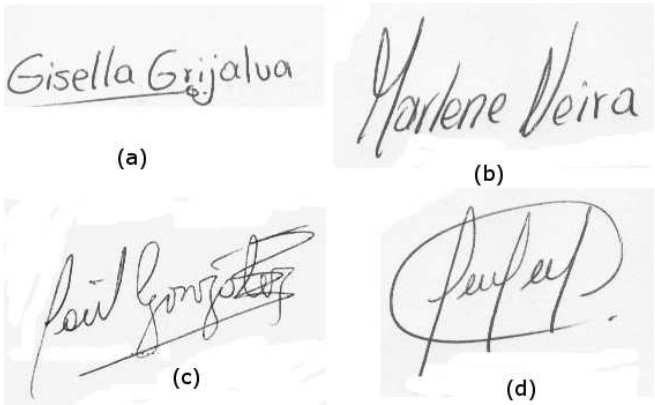


Fig. 3.1: Estilos de assinaturas manuscritas.

características de um texto, por exemplo as iniciais do nome do autor, mas isso não é uma regra. Nesse caso, a habilidade gráfica do escritor é em geral mais aprimorada, e se preocupa mais com a forma dos traços do que com a semântica do próprio nome, como nas assinaturas c) e d) da figura 3.1.

3.2.2 Falsificações de Assinaturas

Uma assinatura falsificada é aquela feita com o fim de imitar uma assinatura verdadeira para se passar por legítima. Tomando como referência a semelhança existente entre uma assinatura original e sua falsificação, encontramos três tipos de falsificações: as aleatórias, as simples e as habilitadas.

Falsificação aleatória

As falsificações aleatórias chamadas também falsificações de zero-esforço, são aquelas onde o falsificador reproduz a assinatura original sem conhecê-la. Estas se caracterizam por ter sua forma gráfica e constituintes semânticos completamente diferentes da assinatura original. Nesse caso, o falsificador faz uma assinatura no documento sem se importar em imitar os traços básicos da assinatura original, inclusive chegando a escrever seu próprio nome ou qualquer outro grafismo para indicar que se trata da assinatura genuína. Na maioria das vezes, a assinatura forjada não possui semelhança com a original, conforme mostra a Fig. 3.3.

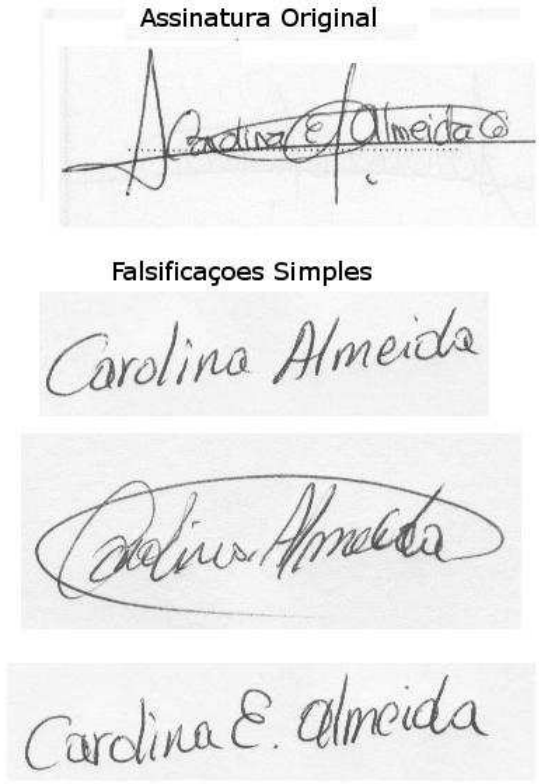


Fig. 3.2: Exemplo de falsificações simples

Falsificação simples

Este tipo de falsificação ocorre quando um falsificador escreve o nome da pessoa corretamente mas não consegue imitar sua forma gráfica. Dessa maneira, a falsificação pode se assemelhar ou não com a assinatura original. Esse tipo de falsificação geralmente ocorre quando o falsificador é inábil e possui apenas o conhecimento do nome da pessoa, mas não tem nenhuma cópia impressa da assinatura verdadeira para se basear e assim poder desenhar uma falsificação mais aprimorada. A figura 3.2 mostra um exemplo de falsificação simples, onde temos a assinatura genuína na parte superior e as falsificações na parte inferior.

Falsificação habilidosa

Esse tipo de falsificação também é conhecido como servil ou qualificada, é produzida quando o falsificador tem acesso a uma cópia ou conhece como é feita a assinatura original. O falsificador faz um esforço para obter a reprodução fiel dessa assinatura, trabalhando da maneira mais

Fig. 3.3: Exemplo de falsificações aleatórias

detalhada possível traço após traço, até conseguir uma falsificação de excelente qualidade. A figura 3.4 mostra o exemplo de duas falsificações habilidosas.

3.3 Um Sistema de Verificação de Assinaturas

As assinaturas são um caso particular de manuscritos, em que aparecem desenhos e caracteres com forma especial ou distorcida. É comum não haver regularidade quanto ao tamanho e distribuição dos caracteres. Em muitos casos, as assinaturas são ilegíveis (parte semântica desconhecida). Porém, inúmeras características próprias do autor são consciente e inconscientemente depositadas no papel quando o mesmo assina. Assim sendo, é possível uma identificação posterior através do processamento de tais características [25].

3.3.1 Um Modelo Geral de um Sistema de Verificação de Assinaturas

O desenho de um sistema para verificação de assinaturas requer a execução de 5 tipos de processos [26]:

1. Aquisição de dados

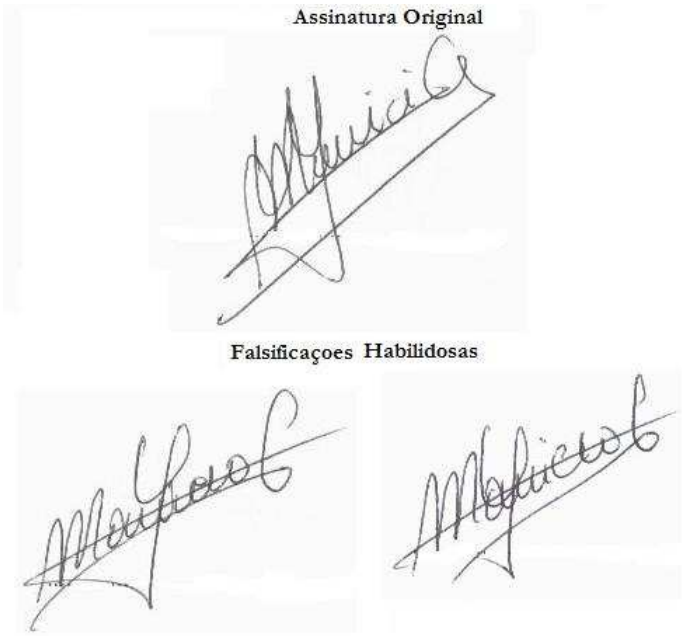


Fig. 3.4: Exemplo de falsificações habilidosas

2. Pré-processamento

3. Extração de características

4. Processo de comparação

5. Decisão (Aceitação/Rejeição).

Um modelo geral de um sistema de verificação de assinaturas está representado na figura 3.5.

Aquisição de dados

A unidade de aquisição da dados executa a operação de digitalização da informação subministrada por um dispositivo de entrada ao sistema de verificação de assinaturas. O tipo de informação que é utilizada determina se o sistema de verificação de assinaturas é estático ou dinâmico.

Em um sistema estático de verificação de assinaturas, (também conhecido como sistemas *off-line*), normalmente captura a imagem de uma assinatura com ajuda de uma câmara, scanner, ou mesa digitalizadora. A assinatura ou texto escrito no papel aparece como uma imagem bidimensional adquirida através destes meios ópticos.

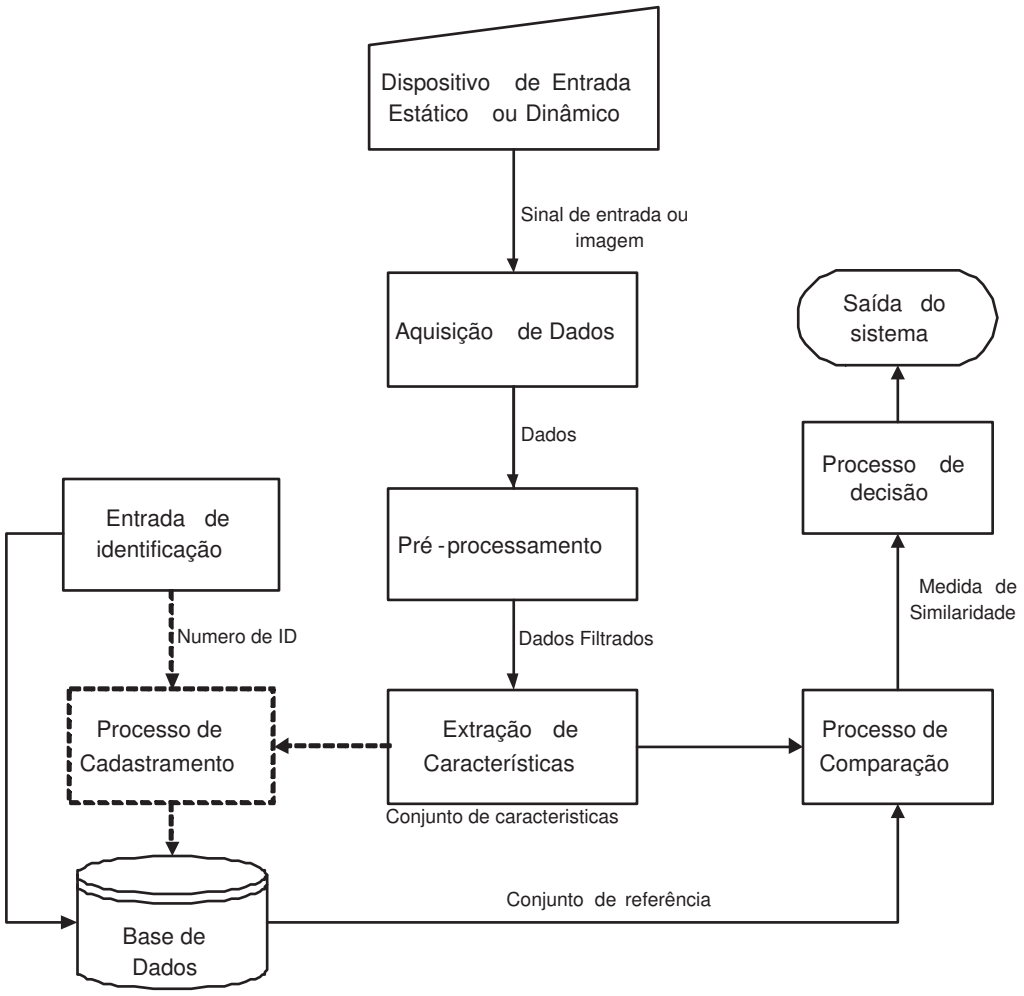


Fig. 3.5: Modelo geral de um sistema de verificação de assinaturas.

Em um sistema dinâmico de verificação de assinaturas, (também conhecido como sistema *on-line*), na maioria dos casos utiliza uma mesa digitalizadora, uma caneta eletrônica ou outro dispositivo para produzir os sinais de entrada. A assinatura é representada por um ou mais sinais que variam com o tempo. Velocidade, pressão, aceleração e posição são alguns exemplos de sinais dinâmicos.

A informação produzida por esta etapa do sistema é enviada à etapa de pré-processamento.

Pré-processamento

A unidade de pré-processamento recebe os dados da etapa de aquisição com o objetivo de eliminar os dados não relevantes e extrair as características importantes para a análise do sinal de entrada. Para isto, executa processos de filtragem, segmentação e normalização do sinal.

Fig. 3.6: Mesa digitalizadora

Os dados de entrada são filtrados com o fim de reduzir o ruído ou componentes não desejados do sinal, a seguir, são executadas algumas tarefas de normalização que podem ser executadas facilmente através interpolação e um processo de reamostragem. Isto é preciso para que toda a informação das diferentes amostras de entrada estejam dentro de um mesmo padrão para que seja possível uma posterior comparação [5].

Em alguns casos esta etapa também realiza um procedimento de segmentação da assinatura que consiste em separar pedaços de assinatura incluídos entre o momento que o usuário do sistema pressiona a caneta até que esta é levantada, depois estes pedaços ou segmentos da assinatura são processados.

Para o caso em que se realiza a verificação de assinaturas através da captura de uma imagem, a fase de pré-processamento realiza os processos de localização da assinatura no papel, a extração da assinatura propriamente dita do fundo da região onde esta contida, a seguir, se realiza a filtragem para melhorar a qualidade da imagem e finalmente a normalização da imagem para estabelecer um tamanho padrão que facilitará sua comparação.

Extração de características

A etapa de extração de características está encarregada da representação de uma assinatura através de propriedades ou relações extraídas dentre as partes de um sinal ou imagem que

pertence à assinatura que está sendo analisada. Estas propriedades ou relações são chamadas de características e o conjunto das características é denominado de vetor de características. Normalmente este vetor é calculado a partir de dados de entrada filtrados e normalizados.

É claro que ainda fica a pergunta "Que aspectos ou características de uma assinatura são importantes?". Não há nenhuma resposta simples para isto, mas duas diferentes abordagens são comuns. Na primeira, se assume que todos os valores coletados de uma assinatura são importantes e assinaturas de teste e de referência são comparadas ponto a ponto usando um ou mais conjuntos destes valores. Nesta aproximação o principal assunto que surge é como a comparação será levada a cabo. Talvez as assinaturas poderiam ser comparadas calculando o coeficiente de correlação entre os valores da assinatura de teste e os valores da assinatura de referência correspondente, mas a comparação ponto a ponto não trabalha bem dado que partes de duas assinaturas genuínas da mesma pessoa podem variar significativamente e a correlação pode ser afetada seriamente pelo deslocamento, rotação ou tamanho variável da assinatura. Na segunda abordagem, não todos os valores disponíveis são usados, só uma coleção de valores é calculada e comparada.

A representação da assinatura é caracterizada como sendo global ou local [30] [31] [29]. A representação global é conveniente quando se estuda as características da assinatura como um todo. Um exemplo de tal abordagem é o uso dos coeficientes da transformada de uma imagem ou ainda a medida da inclinação de vários segmentos da assinatura. A representação local é conveniente quando se requer um detalhamento maior entre os componentes da assinatura. Nesse caso geralmente a assinatura é dividida em segmentos ou regiões onde são extraídas características referentes apenas àquela região. Posteriormente para comparar duas assinaturas se teria que comparar as características dos segmentos correspondentes usando algum alinhamento de segmentos de ser necessário.

A seguir se apresentam alguns exemplos das principais características que poderiam ser utilizadas na verificação de assinaturas [27], mas existem trabalhos que utilizam outras características [28].

- Tempo total levado pelo escrito ao assinar.
- Comprimento do caminho da assinatura: deslocamento nas direções x e y e o desloca-

mento total.

- Tangentes dos ângulos do caminho: perfil de sua variação e média ou valores da raiz quadrática meia (RMS).
- Velocidade da Assinatura: perfil de variações em velocidade horizontal, vertical e total como também sua média ou valor RMS.
- Acelerações da Assinatura: variações em acelerações horizontais e verticais, aceleração centrípeta, aceleração tangencial, acelerações totais, como também suas médias ou valores RMS.
- Tempo no qual a caneta é levantada (*penup*): tempo total de *penup* ou a relação de tempo *penup* com o tempo total.
- Coeficientes de Fourier, Walsh, Haar ou transformada Wavelet.
- Inclinação e coeficientes derivados da transformada de Hadamard.

Uma vez que um conjunto de características é selecionado, não há nenhuma necessidade de manter a assinatura de referência, só é preciso armazenar os valores de suas características. Assim, quando uma assinatura de teste é processada, somente os valores de suas características são necessários, não a assinatura. Isto é freqüentemente utilizado para economizar em armazenamento, por exemplo, quando a assinatura de referência precisa ser armazenada em um cartão.

Os vetores de características de teste e de referência que foram extraídos de uma ou um conjunto de assinaturas devem ser armazenados para cada usuário do sistema. Estas informações são incorporados a uma base de dados através de uma informação que identifique cada usuário. Essa identificação é depois utilizada para extrair o conjunto de referência apropriado da base de dados.

Processo de comparação

Esta etapa do sistema de verificação de assinaturas executa o processo de comparação entre os vetores de características das assinaturas, que encontram-se armazenados numa base de

dados. Esta comparação e obtida através de uma distância de similaridade (quanto menor a distância, maior a semelhança) entre os vetores de características de teste e referência [32].

Dependendo do método utilizado na verificação de assinatura *on-line*, o processo de comparação pode ser classificado em dois grupos: O primeiro grupo contém métodos que utilizam funções como características; os sinais (normalmente posição, velocidade, aceleração, pressão. vs tempo, etc.) são representados por funções matemáticas cujos coeficientes diretamente constituem o conjunto de características. No segundo grupo, o método utiliza parâmetros do sinal como características (tempo total, médias, número de cruzamentos por zero , etc.) que são calculados dos sinais medidos.

No processo de comparação, a assinatura de referência está baseada em um grupo de amostras das quais se obtêm um conjunto de características que é calculado usando a média e o desvio padrão.

Para obter uma boa estimativa da média e do desvio padrão dos valores das características de uma população de assinaturas genuínas é necessário ter várias amostras. Com um grupo maior de amostras é provável que se tenha uma assinatura de referência melhor e então resultados melhores, mas é sabido que nem sempre isto é possível.

A distância entre uma assinatura de teste e uma de referência correspondente pode ser calculada de várias formas. Por exemplo: função de discriminante Linear, classificador de distância Euclideana, técnicas de casamento dinâmico, Modelos ocultos de Markov e redes neurais (ver 2.2).

Como em qualquer sistema automático de reconhecimento de padrões, um sistema de verificação precisa adquirir algum conhecimento sobre as assinaturas verdadeiras de seus escritores, antes de poder executar sua tarefa. Tal processo de aquisição de conhecimento é executado durante a fase chamada de treinamento.

Na fase de treinamento, um certo número de assinaturas verdadeiras é utilizado para gerar um vetor de características, que contém a média de cada uma das características que foram extraídas. Somente depois que esses dados são calculados o sistema poderá ser colocado em operação. Então, quando uma assinatura desconhecida é processada, o sistema determina sua autenticidade ou não, através de um processo de comparação.

Finalmente, um processo de decisão avalia a saída da etapa de comparação com respeito a um limiar, para determinar se a assinatura deverá ser considerada como verdadeira ou falsa.

Processo de Decisão

O processo de decisão consiste em classificar as assinaturas, através de um vetor de características. Ou seja, nos deparamos com um problema de classificação de apenas dois padrões, um deles sendo as assinaturas verdadeiras e o outro, qualquer conjunto que não tenha as características da assinatura genuína.

A seleção do conjunto ótimo de assinaturas de referência para cada usuário do sistema assim como também a obtenção de um limiar que permita diferenciar entre assinaturas pertencentes a um usuário ou não, afeta fortemente o resultado do processo de decisão no momento de aceitar ou rejeitar uma assinatura.

No caso de um classificador estatístico, o limiar normalmente é obtido comparando as características das assinaturas do conjunto de referência. Esta comparação é realizada calculando a semelhança entre as amostras deste conjunto. Pode se calcular um só limiar para todo o conjunto de usuários ou um limiar para cada um dos usuários do sistema.

O conjunto de amostras de referências pode estar formado por assinaturas verdadeiras e falsas, mas de não ser isto possível utiliza-se só assinaturas verdadeiras, já que pode ocorrer que as falsificações sejam de baixa qualidade ou não se conte com pessoas dispostas a colaborar falsificando assinaturas.

No caso ideal, as assinaturas verdadeiras e as falsas estariam muito bem definidas e poderiam ser separadas em duas classes. Na prática, devido à variabilidade que existe numa mesma assinatura, assim como a habilidade que pessoas têm de falsificar assinaturas, os vetores de características das assinaturas verdadeiras e falsas podem chegar a valores muito próximos. Assim sendo, dependendo do limiar que se escolha para a classificação das assinaturas, poderá ocorrer um erro de classificação.

Existem diversas técnicas de classificação que tem sido utilizadas nos últimos anos dentre as mais importantes podemos citar: as redes neurais onde na maioria das vezes se utiliza um classificador perceptron multicamada [5]. Por outro lado, abordagens tradicionais como classificação utilizando o classificador de distância mínima de vizinhos mais próximos ou também

Dynamic Time-Warping, vem sendo adotados [32]. Abordagens baseadas em classificadores que utilizam modelos ocultos de Markov tem sido recentemente descritos na literatura [33]. A principal diferença entre esses estudos se encontra nas características utilizadas para representar uma dada assinatura.

Base de Dados

Outra parte do sistema de verificação de assinaturas é a unidade de base de dados a qual contém informações sobre as assinaturas de referência dos usuários e outros dados relativos à sua identidade.

Como pode ser visto na figura 3.5, a entrada de identificação do sistema é um número, um código ou o próprio nome do indivíduo cuja assinatura vai ser armazenada ou analisada. Esta informação é necessária para poder localizar os dados de referência de um usuário na base, sem essa informação, o sistema teria que fazer a identificação da assinatura comparando-a uma a uma, entre todas as assinaturas da base de dados.

3.4 Verificação de Usuário pelo Som

Quando um usuário escreve, o movimento da ponta de caneta sobre as fibras do papel gera emissões acústicas, que se propagam no interior do material que está abaixo do papel. Estas ondas se comportam de um modo semelhante às ondas sonoras no ar e podem ser detectadas por um sensor preso no material sobre o qual se está assinando. As ondas de som capturadas formam um sinal que pode ser utilizado para desenvolver uma técnica de verificação pessoal. A esta forma de verificação baseada no som da assinatura denomina-se Emissão Acústica (EA - *Acoustic Emission*) [4].

Fricções entre uma caneta de ponta rígida e o papel resultam em sons audíveis que são correlacionados com a dinâmica da escrita. Tais sons da escrita podem ser usados como uma identidade biométrica para alcançar a autenticação do escritor.

As singularidades da escrita a mão não estão situadas somente nas características das formas das letras e palavras, senão também na dinâmica da postura de escrever, como a força, velocidade e aceleração, etc. A pressão, velocidade e aceleração dos movimentos da caneta são transformadas em um único sinal no tempo - o sinal do som. Num sistema de autenticação ou

identificação automática, estas características podem ser extraídas usando processamento de sinais associado a métodos de reconhecimento de padrões.

O interesse de usar o som da escrita para a verificação e/ou autenticação do escritor, é que capturar sons com um microfone é muito mais fácil que medir a dinâmica de escrever sobre uma mesa digitalizadora. Estes sons estão correlacionados com a dinâmica ou postura de escrever pelo que contêm informação útil para distinguir escritores diferentes.

A utilidade de sinais de sons para a análise de textos e autenticação é mostrada nos trabalhos realizados por: Soule no ano 2003 [9] e Li no ano 2004 [5], onde se utiliza o som da escrita para obter informação sobre a dinâmica do escritor.

No trabalho realizado por Soule, apresenta-se uma caneta a qual possui um microfone embutido para aquisição do som produzido durante a escrita acima de um papel comum (ver figura 3.7). Neste trabalho também são discutidas características do dispositivo como também uma avaliação dos sinais produzidos. Os resultados experimentais indicam como possíveis áreas de aplicação deste dispositivo a verificação de assinatura, a identificação de escritor e o reconhecimento de texto manuscrito.

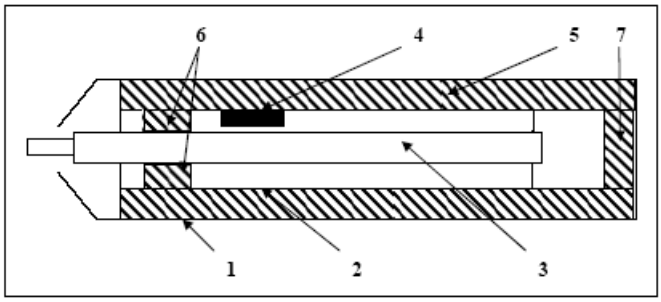


Fig. 3.7: Caneta com um microfone integrado. 1) cobertura exterior; 2) cobertura interna; 3) refill; 4) microfone; 5), 6), 7) isolamento de som.

No trabalho desenvolvido por Li apresenta-se uma alternativa para autenticação da caligrafia baseada no som. Características extraídas dos sons da escrita, são utilizadas para diferenciar topologicamente similares caracteres escritos por diferentes indivíduos. Uma etapa

de pré-processamento em união com uma rede neural supervisionada que é projetada para este propósito é treinada com exemplos para diferenciar efetivamente padrões de sons da escrita e assim alcançar a autenticação do escritor, o que constitui uma alternativa aos métodos existentes.

Nestes trabalhos foram feitas experiências para investigar os parâmetros que têm um impacto maior num sistema de reconhecimento que utiliza o som do sinal registrado. Os aspectos seguintes foram considerados:

1. Som indesejável como barulho de fundo
2. A superfície usada para a escrita
3. O material e a geometria da caneta
4. Os parâmetros dos componentes eletrônicos (microfone)
5. A dinâmica da escrita.

Os sons indesejados causam a degradação do sinal capturado e fazem seu processamento mais difícil. Para resolver isto o dispositivo de captura tem seu próprio isolamento para suprimir o barulho de fundo e o barulho que resulta do movimento da mão. Também, procede-se a filtrar o sinal capturado e a eliminar as amostras que correspondem ao ruido de fundo, isto é, elimina-se as amostras que não são parte da escrita do usuário. Para conseguir este propósito se trabalha com a energia do sinal a fim de determinar quais amostras são ruído de fundo e quais são sinal da escrita.

Um dos aspectos mais importantes do sistema é a influência da superfície na que o usuário escreve. A aspereza e dureza da superfície determinam a amplitude do sinal de som durante a escrita. A ponta da caneta que rola na superfície é uma fonte de fortes vibrações geradas pelo movimento acima do papel. Estas vibrações definidas pela aspereza da superfície e a velocidade do movimento, são capturadas pelo microfone que age como um sensor acústico. Experiências com superfícies diferentes mostraram que todos os tipos de papel são úteis para registrar a dinâmica da escrita. Os sinais medidos são suficientes para reconhecer mudanças nas características do escritor.

O material e a geometria da caneta podem influenciar essencialmente nas características do sinal registrado no tempo e no domínio de freqüência. A amplitude do sinal é dada pela construção mecânica da caneta e a posição do microfone. As experiências realizadas têm provado que a configuração da caneta com um microfone embutido provê os melhores sinais acústicos da escrita.

Os parâmetros dos componentes eletrônicos determinam as características do sinal resultante. Assim é desejável ter os parâmetros elétricos dos componentes como o microfone o mais constantes possível.

Para uma determinada configuração de: superfície, papel, geometria da caneta e microfone as amplitudes do sinal são principalmente determinadas pela velocidade do movimento durante a escrita. Assim o comportamento do sinal capturado corresponde a dinâmica da escrita dada pelas características biométricas do escritor.

Capítulo 4

Verificação de assinaturas: Modelo Acústico

Neste capítulo, descreve-se detalhadamente as diferentes etapas que compõem o sistema de verificação de assinaturas proposto no presente trabalho.

O diagrama da figura 4.1 apresenta as etapas de um sistema de verificação pessoal baseado no som da assinatura. Neste diagrama pode-se observar que o sinal de entrada é o som da assinatura que foi capturado. Este sinal é ingressado nas etapas de pré-processamento onde é filtrado, eliminadas as amostras que não são parte da assinatura utilizando a detecção de pontos de início e fim, e em seguida, a envoltória deste sinal é obtida e extraiem-se os vetores de características que são enviados à etapa final formada pelo processo de decisão onde os vetores de características de uma assinatura verdadeira e uma assinatura falsa são comparados utilizando uma medida de semelhança. Cada uma destas etapas serão detalhadas neste capítulo.

4.1 Aquisição de Assinaturas

A definição de uma base de dados que seja representativa do universo de assinaturas que podem ser encontradas em um problema real de reconhecimento é um requisito fundamental para a avaliação experimental das técnicas aplicadas neste trabalho.

Nos problemas de verificação, técnicas de amostragem são utilizadas para extrair um subconjunto de elementos (amostras) de uma população que seja representativo, no sentido de que

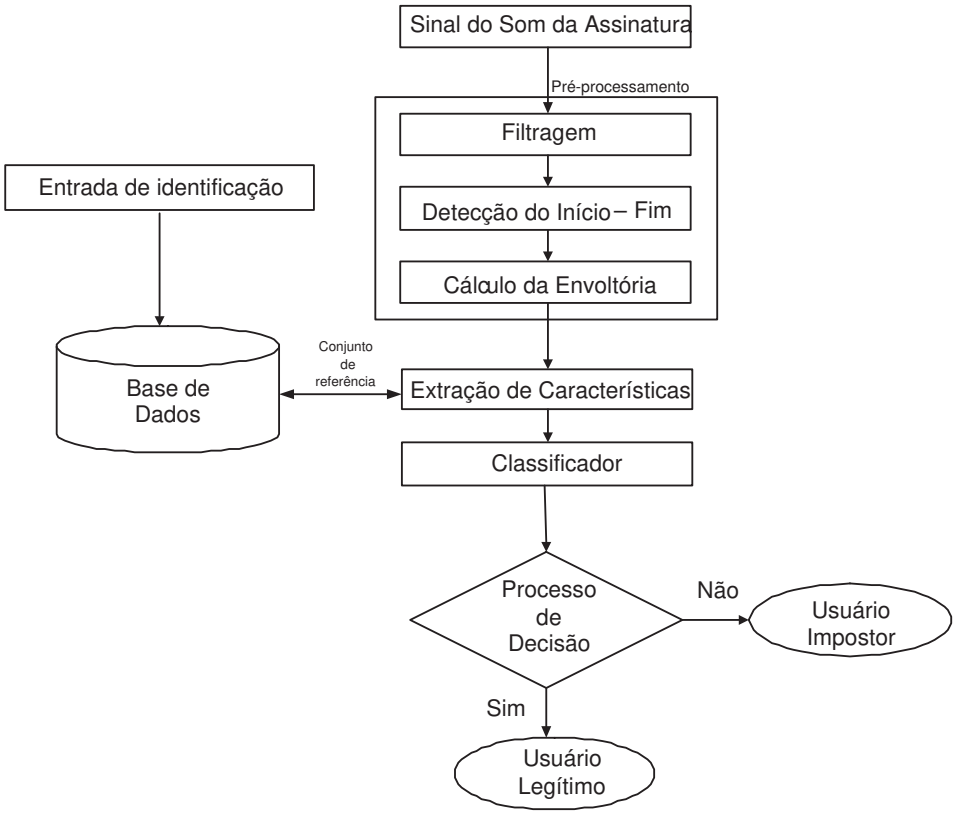


Fig. 4.1: Diagrama em blocos do método proposto

propriedades obtidas a partir da observação de certas variáveis ou característica da amostra possam ser extrapoladas para a população como um todo.

Um método de amostragem apropriado deve ser utilizado de maneira que leve em conta a possibilidade que todos os elementos da população ou alguns desses elementos façam parte das amostras. Se todos os componentes da população tiverem igual probabilidade de participar da amostragem, se diz que o método usado é de amostragem causal, caso contrario, fala-se de amostragem não causal.

Vários critérios podem ser utilizados num método de amostragem não causal para garantir que uma amostra não seja tendenciosa ou não representativa da população. No entanto, quando esse tipo de informação não está disponível ou é difícil de ser obtida, a adoção de tais critérios torna-se proibitiva. Nessa situação, uma opção seria obter uma amostra de conveniência, ou seja, uma amostra que esteja naturalmente disponível e que não dependa de critérios difíceis para a seleção de seus elementos. Assim, o espaço amostral poderia ser o local de trabalho, a

universidade, uma cidade, etc.. Para este trabalho, adotou-se um método de amostragem de conveniência para obtenção das amostras da base de dados de assinaturas. A Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas foi o local de conveniência escolhido para a coleta das amostras.

O problema seguinte a resolver, é quais tipos de objetos deverão constituir a base de dados. Como o objetivo deste trabalho é realizar a verificação pessoal utilizando o som que se produz no ato de assinar, a base de dados é constituída pelos sons digitalizados e informação do assinante que permita a sua identificação. A fim de permitir a avaliação de sistemas de verificação, além dos sons das assinaturas genuínas (verdadeiras) de cada escritor, são incluídos sons de falsificações para essas assinaturas.

Durante a realização dos experimentos, um subconjunto de assinaturas verdadeiras e impostoras são utilizados como conjunto de referência, com o objetivo de treinar os classificadores com exemplos e contra-exemplos de assinaturas de um autor. As assinaturas verdadeiras e as falsas restantes constituiriam o conjunto de teste.

Utilizar assinaturas verdadeiras e falsas para formar o conjunto de referência de um autor é uma abordagem bastante útil, mas isto não é sempre possível já que é difícil encontrar voluntários dispostos a praticar até conseguir imitar uma assinatura (falsificação habilidosa).

Outro aspecto a levar em conta para prosseguir com o desenvolvimento deste trabalho é definir a quantidade ideal de classes a serem consideradas em uma amostra, pois mesmo em estatística, não segue uma regra muito precisa. Nesse contexto, costuma-se falar de conjunto pequeno ou grande de amostras. A partir desse elemento, quanto maior for o número das amostras, melhor será a sua representatividade, desde que se tenha adotados critérios de amostragem adequados. Em problemas de classificação envolvendo uma população infinita, uma amostra contendo menos de 50 padrões com variáveis de distribuição de probabilidade aleatória é considerada pequena. Dentro desta visão, neste trabalho, escolheu-se 55 como o número total de classes que constituem a base de dados.

Com referência à quantidade de padrões ou amostras por classe, neste trabalho, realiza-se os testes para analisar as diferentes respostas do sistema de verificação baseadas no número de assinaturas utilizadas para gerar o conjunto de referências. Assim, no tocante à quantidade de

amostras por classe, escolheu-se onze sons de assinaturas verdadeiras e onze sons de assinaturas falsas como sendo o número de assinaturas capturadas para cada usuário.

Analisando a maioria dos problemas de verificação de assinaturas, pode-se identificar um certo limite prático do número de assinaturas verdadeiras utilizadas para gerar o conjunto de referência. Esse limite está relacionado com a habilidade do ser humano em reconhecer as assinaturas. Num sistema bancário, ou num sistema do reconhecimento de firmas em cartório, por exemplo, uma amostra de teste é conferida com base em três assinaturas de referência. Neste trabalho utilizou-se um número máximo de onze sons de assinaturas no conjunto de referência para cada autor.

As assinaturas estão divididas em dois grupos: os sons das assinaturas verdadeiras e os sons das falsificações habilidosas.

O grupo de assinaturas verdadeiras totaliza 605 sons. Essas assinaturas foram obtidas junto a cinqüenta e cinco pessoas, sendo que cada uma delas contribuiu com 11 assinaturas. As assinaturas verdadeiras foram coletadas num período de seis meses.

Para este conjunto de assinaturas verdadeiras foram adquiridas 11 falsificações habilidosas por assinante, resultando em um grupo com 605 imitações. Assim, o número de sons de assinaturas adquiridas para os testes do sistema é de 605 sons verdadeiros e de 605 sons falsos, totalizando 1210 sons de assinaturas.

4.1.1 Equipamento para Coleta de Assinaturas

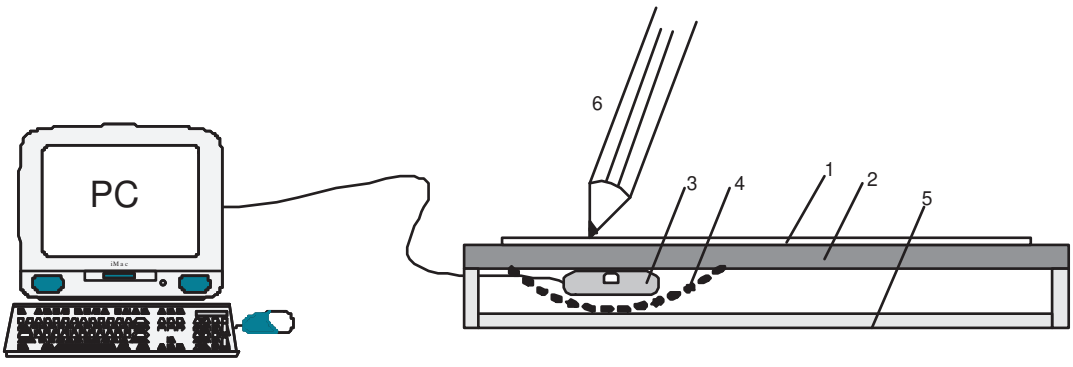


Fig. 4.2: Estrutura interna do dispositivo de captura. 1)papel 2)tábua 3)microfone 4)isolamento 5)cobertura exterior 6)caneta.

Os equipamentos utilizados para a execução deste trabalho foram um computador pessoal

e um dispositivo para captura de sons das assinaturas. O microcomputador utilizado foi um IBM-PC Intel Celeron de 1.8 Ghz, com 256 Mbytes de RAM. Ele foi dotado de um software desenvolvido que permite armazenar o som capturado num arquivo de formato WAV.

O objetivo de realizar este sistema de verificação, é utilizar o som que se produz quando a ponta rígida de uma caneta percorre a superfície do papel durante o ato de assinar e, por essa razão, foi desenvolvido um dispositivo que ajudasse nesta tarefa.

Fig. 4.3: Dispositivo de captura de som.

A figura 4.2 mostra um desenho do dispositivo utilizado para captura do som no presente trabalho, o qual é formado por uma tábua de madeira, sob o qual está fixado um microfone para capturar o som. Este microfone foi isolado de maneira que capture apenas o som produzido pela ponta de uma caneta que se desliza na superfície de uma folha de papel, a qual está presa na tábua. O isolamento minimiza os sons provenientes de outras fontes de som. O sinal capturado pelo microfone é ingressado para o computador para ser digitalizado.

O tamanho da tábua utilizada como base para o papel e sobre o qual o usuário assina é de 25 cm. de largura por 22 cm. de comprimento. Este tamanho foi escolhido para que o assinante tenha a comodidade necessária para assinar e tal que o dispositivo não interfira na

captura das amostras. A tábua de madeira é suficientemente firme para suportar a pressão do assinante e também oferece um bom isolamento do microfone frente aos ruídos ambientais (ver figura 4.3).

Um fato importante durante o processo de coleta das assinaturas é que o som resulta da fricção da ponta da caneta com a superfície do papel, e então, o som produzido depende do tipo de caneta e do papel utilizado, assim como também da posição onde o assinante começa a sua assinatura com relação à posição do microfone.

Para minimizar tais dependências no processo de captura de amostras, sempre se utilizou o mesmo tipo de caneta com ponta rígida de diâmetro 0,5 mm e as folhas utilizadas sempre foram do mesmo tamanho e de papel branco de 75 gramas. Para padronizar a região do papel onde a pessoa deve assinar, foi desenhado um retângulo de 6 cm por 4 cm dentro do qual o usuário assina e, desta forma, tentar manter a mesma posição da assinatura com relação à posição do microfone (ver figura 4.4).

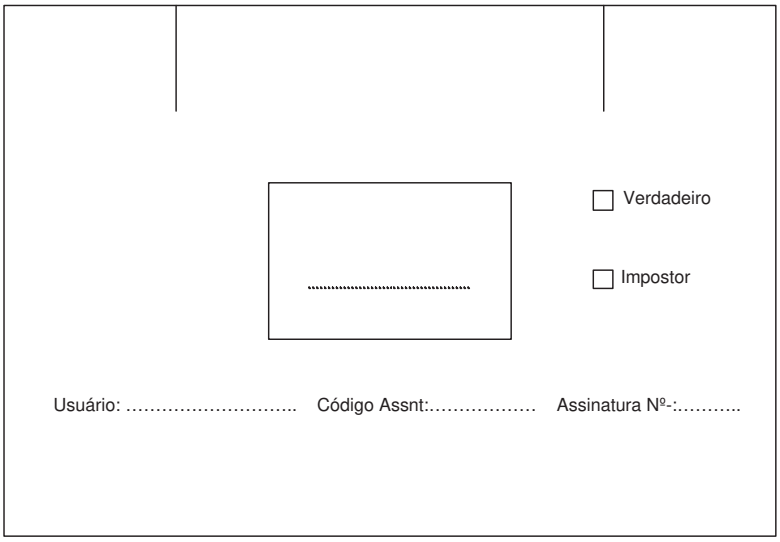


Fig. 4.4: Folha de captura da assinatura.

4.1.2 Software para Capturar o Som da Assinaturas

Para formar a base de dados do sistema de reconhecimento, se utiliza um software que foi desenvolvido na linguagem de programação Java. Está aplicação está composta por uma tela principal e opções de menu que permitem a automatização do cadastramento e o armazena-

mento da assinatura como sinal sonora (ver figura 4.5).

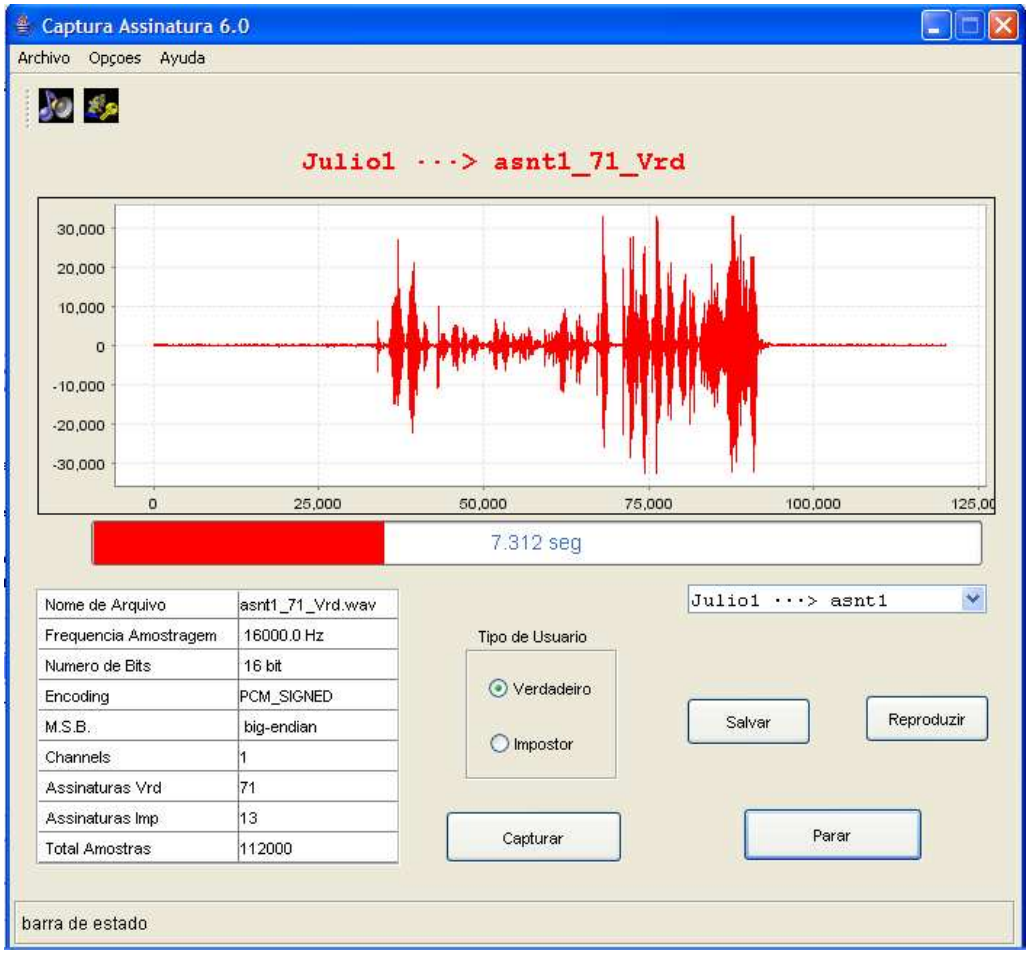


Fig. 4.5: Programa para capturar o som da assinatura.

A opção de menu chamada “registro” realiza o processo de cadastramento do usuário, onde se armazena o nome e se designa um código que permitirá identificar as amostras capturadas para este usuário (ver 4.6). Além disso pode-se eliminar um usuário da base de dados usando a opção de menu “eliminar”.

Mediante a opção de menu “Configurar” é possível especificar as características como: freqüência de amostragem, número de bits e sinal da codificação e posição do bit mais significativo (ver figura 4.7(a)).

Na tela principal da aplicação, pode-es verificar a configuração do amostragem e especificar o tipo de usuário (verdadeiro ou impostor), informação que determina se o som capturado da assinatura foi realizada pelo dono da conta ou um impostor habilidoso. O programa também

Fig. 4.6: Janela de cadastramento do nome do usuário.

apresenta informações como: um desenho do sinal capturado, o nome do usuário ao qual pertence esta conta, nome do arquivo no qual se armazenou o som capturado, número de amostras verdadeiras e impostoras da conta (ver figura 4.7(b)).

O software possui uma interface que facilita seu uso por parte de um usuário cadastrado. Cada vez que o assinante pressiona o botão "Capturar" dispõe de 15 segundos para assinar. Durante este período de tempo, qualquer sinal que se produz sobre a superfície do dispositivo de captura é digitalizado. Este tempo pode ser mudado se for necessário.

O sinal de assinatura apresenta uma energia cujas componentes de freqüência variam aproximadamente entre 200 Hz e 5000 Hz e, portanto, uma freqüência de amostragem de 16KHz. se faz necessária. O sinal de som capturado e convertido num sinal analógico e ingressa no computador onde é digitalizado a uma freqüência de amostragem de 16 KHz. e armazenado pelo programa como um arquivo no formato WAV de 16 bits (ver figura 4.8).

Nas figuras 4.9 e 4.10 se apresentam a impressão gráfica de uma assinatura e seu respectivo sinal digitalizado.

(a) Janela que permite escolher as características do amostragem.

(b) Parte da tela da aplicação que amostra as características do sinal capturado.

Fig. 4.7: Informação das características do sinal de som.

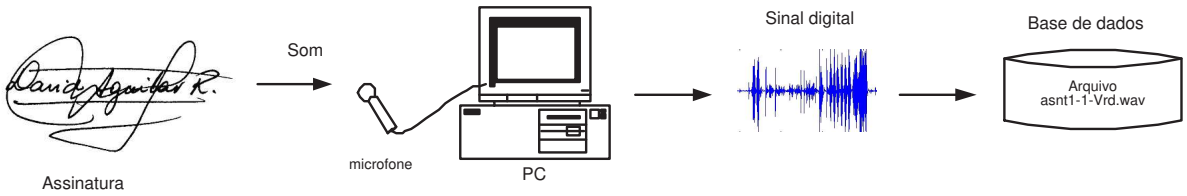


Fig. 4.8: Diagrama de captura da assinatura

Fig. 4.9: Imagem da assinatura

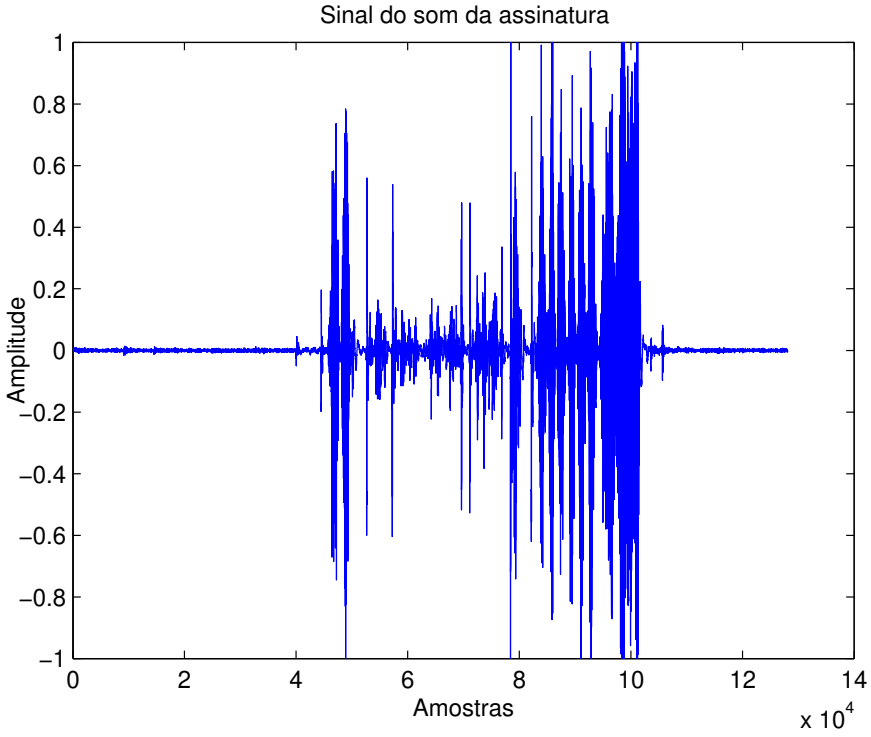


Fig. 4.10: Sinal de som da assinatura

4.2 Pré-processamento do Som da Assinatura

Nesta seção são abordadas as técnicas de pré-processamento aplicadas nos sinais de som das assinaturas. Estas técnicas são utilizadas com o objetivo de tratar os sinais e deixá-los num formato que permita fazer a extração de características que minimize a variabilidade intra-classe e maximize a variabilidade inter-classe. Este pré-processamento foi dividido em quatro etapas que são de (1) filtragem, (2) detecção dos pontos inicial e final, (3) cálculo da envoltória e (4) normalização dessa envoltória, conforme mostrado na figura 4.11.

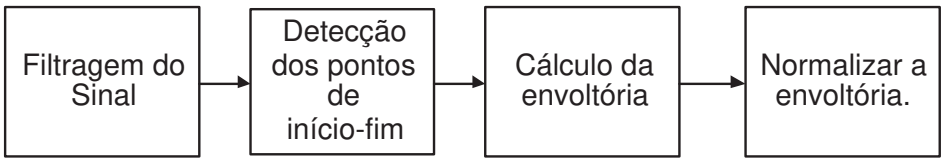


Fig. 4.11: Diagrama em blocos das quatro etapas que compõem o pré-processamento.

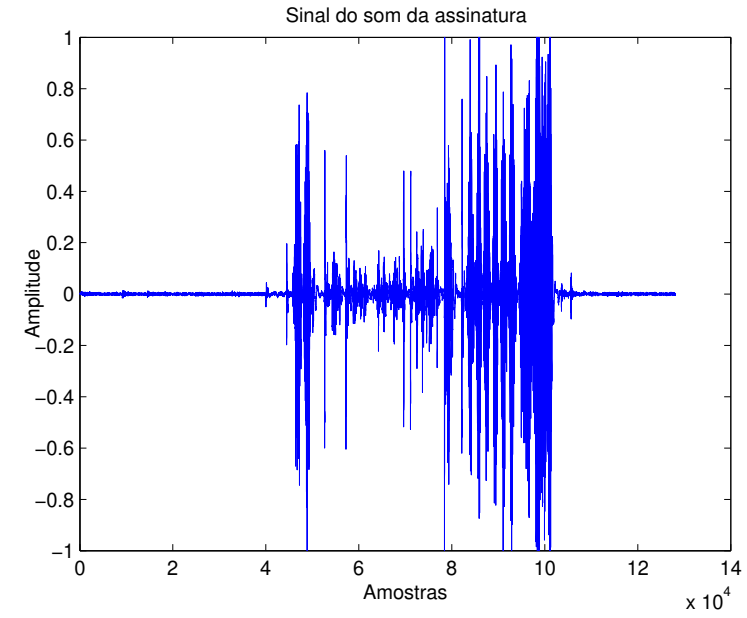
4.2.1 Filtragem do Sinal

Nesta etapa, o objetivo é minimizar o ruído externo ao sinal de som da assinatura que também foi capturado e digitalizado.

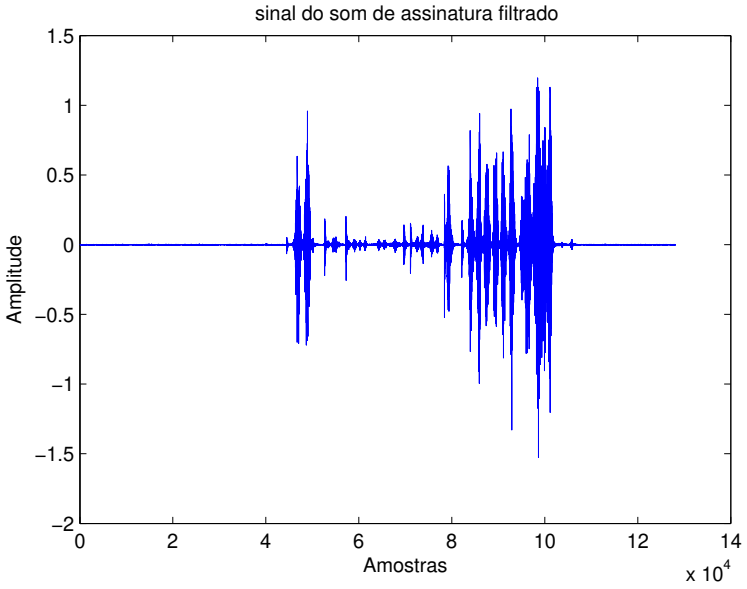
Como explicou-se anteriormente, o dispositivo de captura do som da assinatura é formado de por um microfone o qual está isolado para evitar que o ruído externo ao dispositivo também seja capturado mas esta isolação não é total. Por exemplo, cada vez que o dispositivo de captura de som é utilizado, produz-se um ruído proveniente da fricção da mão do assinante com a superfície da tábua do dispositivo e este ruído apresenta freqüências menores do que 300Hz. Este valor foi obtido depois de analisar amostras do som da fricção da mão de diferentes usuários.

Para eliminar este ruído, o sinal de entrada é passado por um filtro passa-faixa de 380 a 3500 Hz, que é a faixa onde o sinal do som tem concentrada a sua energia.

Na figura 4.12(b) mostra-se o sinal de entrada após do processo de filtragem.



(a) Sinal de som da assinatura.



(b) Sinal de som da assinatura filtrado.

Fig. 4.12: Sinal de som da assinatura. (a) Sinal de som da assinatura. (b) Sinal de som da assinatura filtrado.

4.2.2 Detecção dos Pontos de Início-Fim

Quando o som de uma assinatura é capturado, existe um lapso de tempo antes e depois do usuário assinar, onde se encontram as amostras de ruído de fundo digitalizadas. Estes intervalos de tempo são conhecidos como de silêncio inicial e final. As amostras capturadas nesses tempos não trazem nenhuma informação relevante para o sistema de verificação. Além disso, sua incorporação aos dados de um usuário no momento de calcular seu vetor de características pode ser prejudicial, sendo então necessário a detecção e eliminação destes intervalos de silêncio.

Para eliminar estes intervalos, utiliza-se os algoritmos de detecção dos pontos de início-fim, os quais também são conhecidos como *endpoint detection*, geralmente empregam-se os seguintes parâmetros: energia, estimação espectral e restrições temporais [5] [34] [35]; cruzamento por zero ou por nível [36] [38]. Todos estes trabalhos estão orientados a eliminar esses intervalos para aplicações que trabalham com reconhecimento de voz.

Neste trabalho, foram testados os métodos de detecção dos pontos de início-fim utilizando as técnicas de energia e de cruzamento por zero, sendo que o primeiro apresentou melhores resultados.

Nos sistemas de verificação, por exemplo, de locutor, a energia é bastante empregada para detectar os pontos onde começa e termina o sinal de voz. Para este propósito, calcula-se a energia por janela ou quadro do sinal digitalizado e estes valores são comparados com os valores limite de energia. Os sons do ruído de fundo ou silêncio têm menos energia que sons da voz. Contudo, para que a energia seja útil na tarefa de detecção dos pontos de inicio-fim, ela deve ser apropriadamente normalizada para cada elocução. Este mesmo método é utilizado para a detecção dos pontos de início-fim do sinal de som da assinatura.

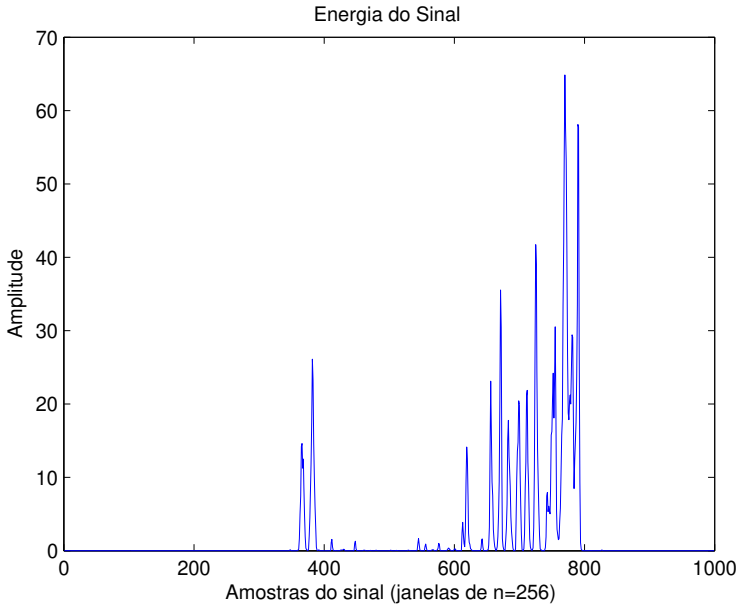
A energia por quadro é calculada pela seguinte equação:

$$E_k = \sum_{n=(k-1)F/2+1}^{(k+1)F/2} (S_n)^2 \tag{4.1}$$

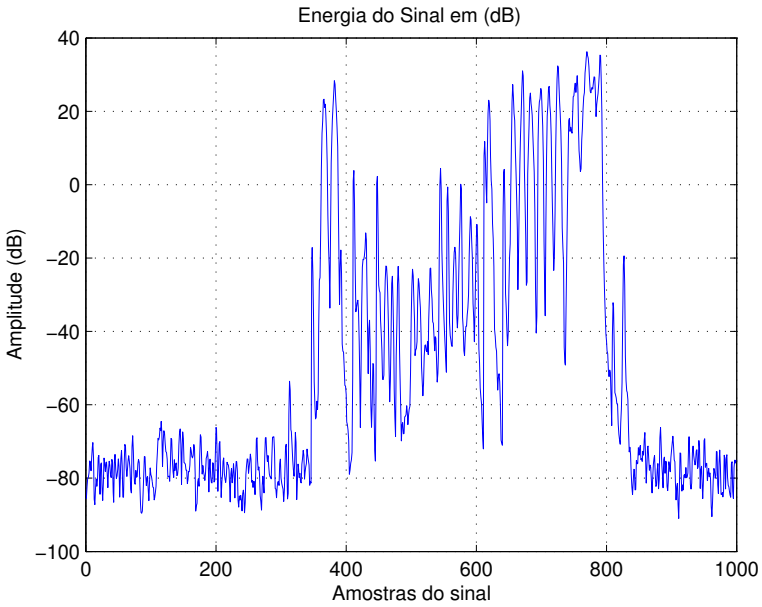
onde $k = 1, 2, ..., K_c$, E_k é a energia de cada quadro ou janela, S_n representa uma sinal do som da assinatura filtrado, K_c é número de quadros nos que se divide o sinal, e F o comprimento do quadro ou o número de amostras do sinal do som por quadro. Para este caso $F = 256$. A

figura 4.13(a) mostra a energia por quadro de um sinal de som [36].

Para se ter uma melhor visualização dos valores pequenos do sinal, pode-se representar a energia por quadro em decibéis. Está energia é definida como $EdB_k = 20 * log(E_k)$ (ver figura 4.13(b)).



(a) Energia do som do sinal E_k.



(b) Energia do som em decibéis EdB_k.

Fig. 4.13: Energia por quadro e Energia por quadro representada em dB do sinal.

A seguir, calcula-se os valores máximos e mínimos de janelas com $m_c = 15$ amostras de comprimento da energia por quadro representada em decibéis EdB. Estes valores são chamados de EdBmax (vide figura 4.14(a)) e EdBmin (vide figura 4.14(b)). O valor com $m = 15$ amostras foi escolhido já que a maioria de picos ou traços do sinal que faz parte de som da assinatura tem um cumprimento de 10 a 20 amostras. Dessa forma, cada ponto do sinal de $EdBmax$ ou $EdBmin$ representa um pico do sinal EdB,

$$EdBmax_l = \max\left(EdB_{i\times l}\right) \;\; i = 1, ..., m_c; \;\; l = 1, ..., Lc, \tag{4.2}$$

$$EdBmin_l = \min\left(EdB_{i\times l}\right) \;\; i = 1, ..., m_c; \;\; l = 1, ..., Lc, \tag{4.3}$$

onde $Lc = \frac{número\ de\ amostras\ de\ EdB}{15}$ é o número de quadros de 15 amostras em que se divide EdB.

Para se detectar os pontos inicial e final do sinal de som, trabalha-se com os sinais auxiliares $EdBmax$ e $EdBmin$. Inicialmente, encontra-se os índices dos quadros de valores de EdB chamados $piniaux$ e $pfimaux$, que são os valores tal que $EdBmax_l > lmax$ em pelo menos dois quadros consecutivos. Encontrados estes índices, retorna-se ao sinal $EdBmin$ até encontrar os valores que $EdBmin_l < lmin$ em pelo menos dois quadros consecutivos. Estes são os índices $pini$ e $pfim$, os quais tem que ser ajustados aos índices de E_k através das equações (4.4) e (4.5) para determinar os pontos início ($pinicio$) e fim ($pfinal$) do sinal de som (ver figura 4.15).

$$pinicio = m\left(\frac{F}{2}\right) pini, \tag{4.4}$$

$$pfinal = m\left(\frac{F}{2}\right) pfim, \tag{4.5}$$

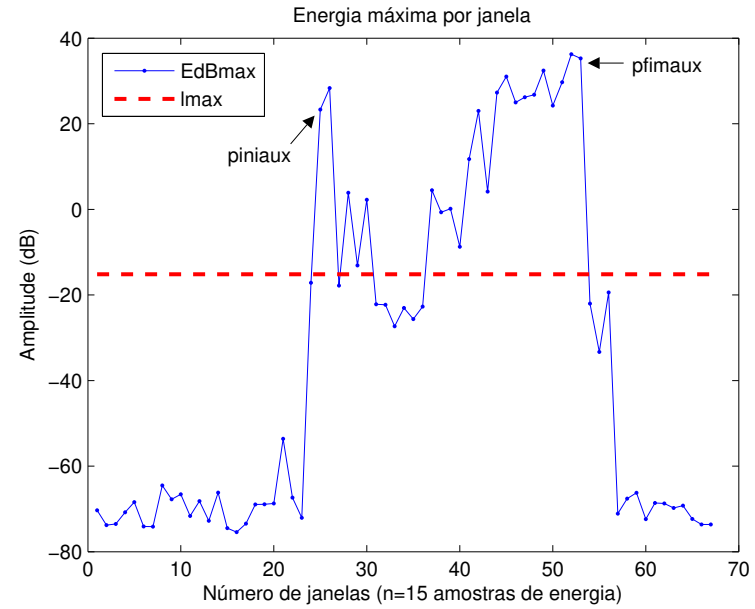
Os valores de $lmin$ e $lmax$ são os limites que permitem encontrar os pontos de início-fim do sinal. Estes são calculados pelas equações (4.6) e (4.7):

$$lmax = \alpha \sum_{i=1}^{m} \frac{EdB_i max}{L}, \tag{4.6}$$

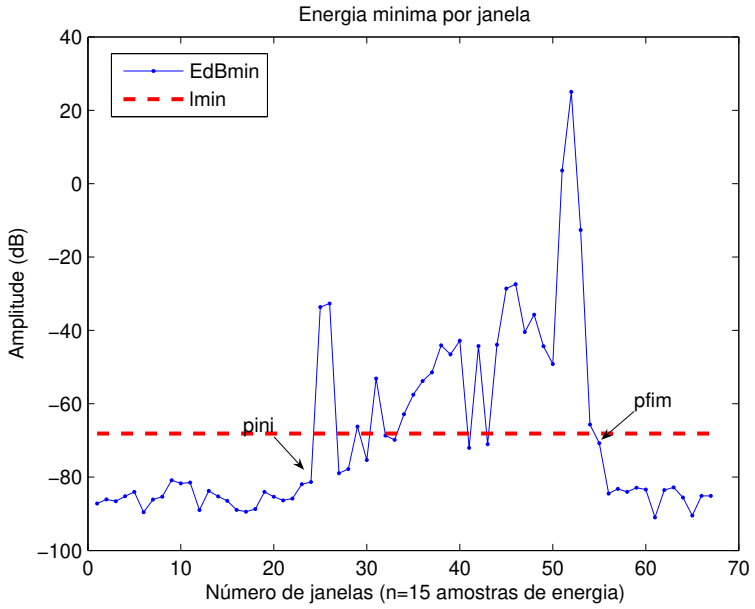
$$lmin = \beta \sum_{i=1}^{m} \frac{EdB_i min}{L} \tag{4.7}$$

As constantes α, β são constantes obtidas empiricamente e, para este caso,foram usadas α=0,45 β=1.

Finalmente todas as amostras do sinal S_n anteriores ao ponto *pinicio* e as amostras posteriores ao ponto $pfinal$ são eliminadas, ficando só as amostras de som da assinatura (ver figura 4.16).



(a)



(b)

Fig. 4.14: Energia máxima e mínima. (a) Energia máxima por quadro com 15 amostras EdB_kmax. (b) Energia mínima por quadro com 15 amostras EdB_kmin.

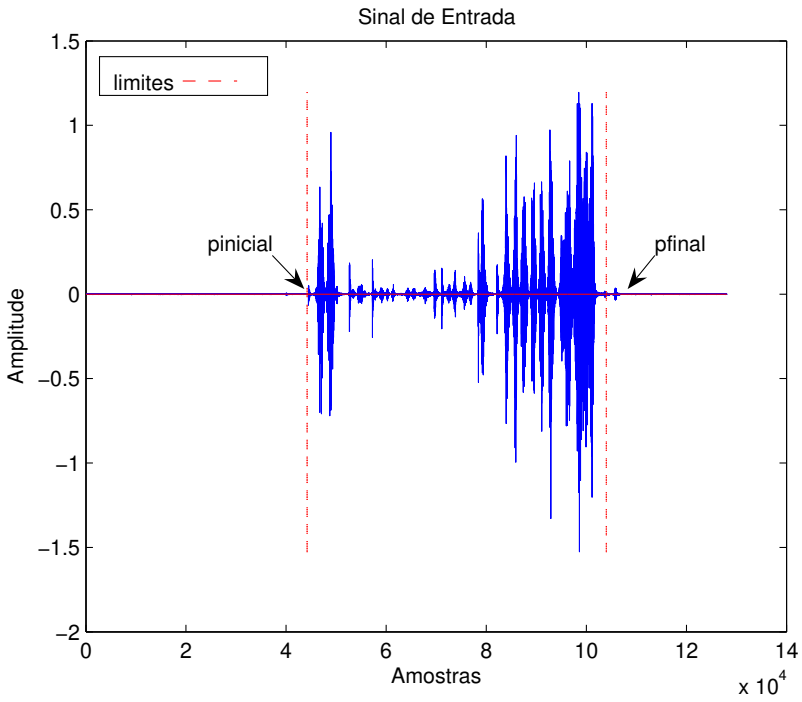


Fig. 4.15: Pontos inicial e final do sinal de som

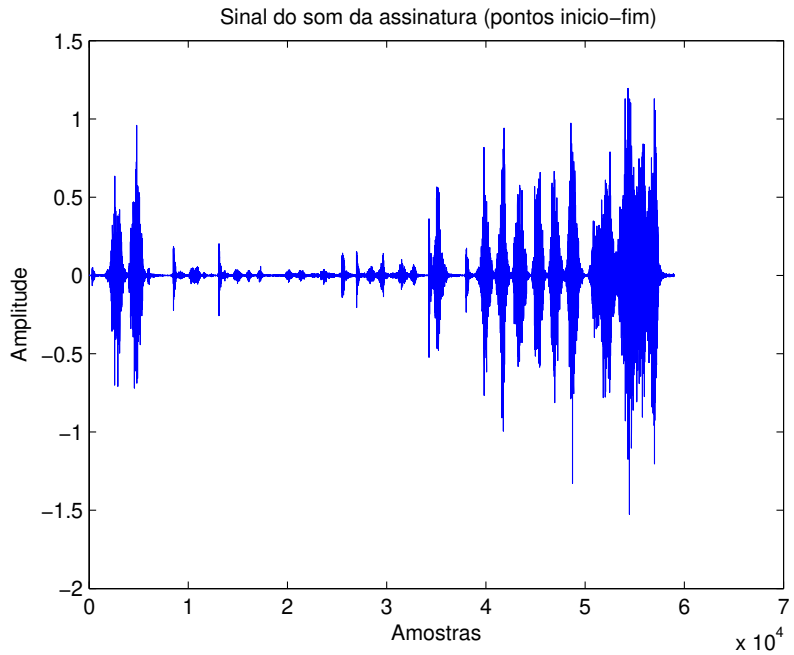


Fig. 4.16: Sinal do som da assinatura (início-fim)

4.2.3 Cálculo da Envoltória

O sinal de som da assinatura capturado, após uma filtragem e uma eliminação de amostras indesejáveis, é aplicado à etapa de pré-processamento onde se determina a sua envoltória normalizada.

Uma técnica comum e muito eficiente para a detecção da envoltória é baseada na transformada de Hilbert. O detector que utiliza esta transformada é considerado como um detector ideal de envoltória [39]

O método de reconhecimento proposto neste trabalho utiliza a envoltória normalizada dos sons da assinatura como o espaço de característica, onde os detalhes de alta freqüência presentes nos sons da assinatura podem ser vistos como sinais de portadora. Diferentes papéis e canetas exibem conteúdos de freqüência diferentes, isto é, portadoras diferentes. O movimento da ponta da caneta na superfície de papel impõe um efeito de modulação nestas portadoras. Espera-se que o algoritmo que calcula o vetor de características extraia as características da dinâmica do movimento de mão, sem grande influência da informação sobre os papéis e as canetas usadas. Então as envoltórias dos sons das assinaturas são consideradas como o espaço de característica do qual pode-se extrair um vetor de características conveniente ao reconhecimento do assinante [5].

Transformada de Hilbert:

Em processamento digital de sinais, precisamos freqüentemente encontrar a relação entre a parte real e imaginária de um sinal complexo. Esta relação geralmente é descrita pela transformada de Hilbert HT.

A HT gira de 90 graus, no sentido horário todas as componentes de freqüência do sinal em 90 graus ($\pi/2$):

$$\hat{x}(t) = x(t) \otimes \frac{1}{\pi t} = \int_{-\infty}^{\infty} \frac{x(\tau)}{t-\tau} d\tau, \tag{4.8}$$

onde $\otimes$ representa a convolução e $\hat{x}(t)$ é a transformada de Hilbert do sinal $x(t)$. O sinal $\hat{x}(t)$ pode ser considerado como a resposta de um filtro linear à entrada $x(t)$, cuja resposta ao impulso é $1/(\pi t)$.

Um sinal $x(t)$ pode ser representado como um sinal complexo $x_c(t) = x(t) + j\hat{x}(t)$ onde a parte imaginária nesta representação é a transformada de Hilbert de $x(t)$.

A envoltória $r(t)$ é obtida calculando-se a magnitude de $x_c(t)$ [39]:

$$r(t) = |x(t) + j\hat{x}(t)| \tag{4.9}$$
$$r(t) = \sqrt{x^2(t) + \hat{x}^2(t)}$$

A Envoltória:

A envoltória do sinal de som digitalizado é obtida calculando-se a transformada discreta de Hilbert através da transformada rápida de Fourier (FFT). Dado que $x(t)$ é o sinal de som da assinatura, $s[n]$ denota a n-ésima amostra desta sinal digitalizado,

$$s[n] = x(nT) \ \ para \ n \in [0, \mathrm{n}_1, \mathrm{n}_2, ..., \mathrm{N} - 1], \tag{4.10}$$

onde T é o período de amostragem e n denota o número de amostras. A transformada discreta de Hilbert de $s[n]$ é então calculado por [5]:

$$H_d\{s[n]\} = \sum_{k=0}^{N} \left\{ A[k] \sin\left(\frac{2\pi kn}{N}\right) - jB[k] \cos\left(\frac{2\pi kn}{N}\right) \right\} \tag{4.11}$$

onde os coeficientes A e B são determinados pela transformada de Fourier.

$$A[k] = \mathrm{Re}\left\{ \sum_{n=0}^{N} s[n] e^{-j2\pi kn/N} \right\} \tag{4.12}$$

$$B[k] = \mathrm{Imag}\left\{ \sum_{n=0}^{N} s[n] e^{-j2\pi kn/N} \right\} \tag{4.13}$$

Estes são eficientemente calculados utilizando a FFT. Finalmente a envoltória do sinal $s[n]$ é obtida por:

$$env[n] = |s[n] + jH_d\{s[n]\}|, \tag{4.14}$$

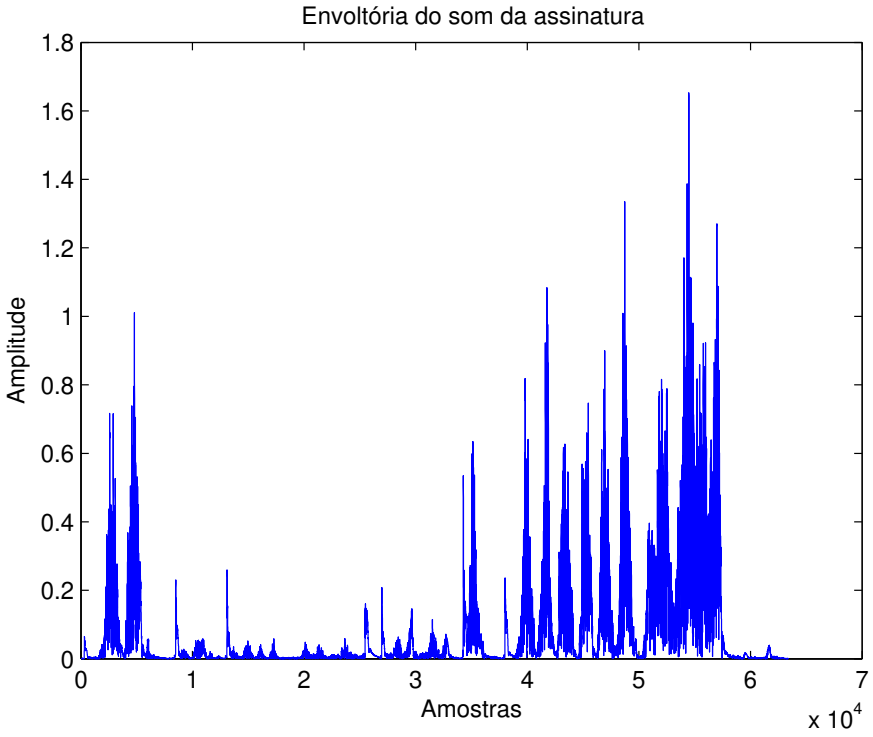


Fig. 4.17: Envoltória do sinal de som.

onde $env[n]$ é a envoltória do sinal do som da assinatura discreta $s[n]$ e H_d é a transformada discreta de Hilbert de $s[n]$.

Na figura 4.17 pode-se observar a envoltória *env* do sinal de som apresentado na figura 4.16. Esta envoltória é calculada utilizando a equação 4.14.

A forma da envoltória do sinal de som (picos e vales) está relacionada com as características de assinar de uma pessoa. Os picos e vales de *env* representam os traços que compõem a assinatura e podem ser utilizados como uma forma de representação. Para obter a forma da envoltória, o sinal *env* pode ser dizimada com o fim de reduzir o número de amostras que vão ser processadas, mas sem perder informação sobre a forma da envoltória (ver fig.4.18). Nesta etapa do pré-processamento a dizimação utilizada é de 500 amostras/seg.

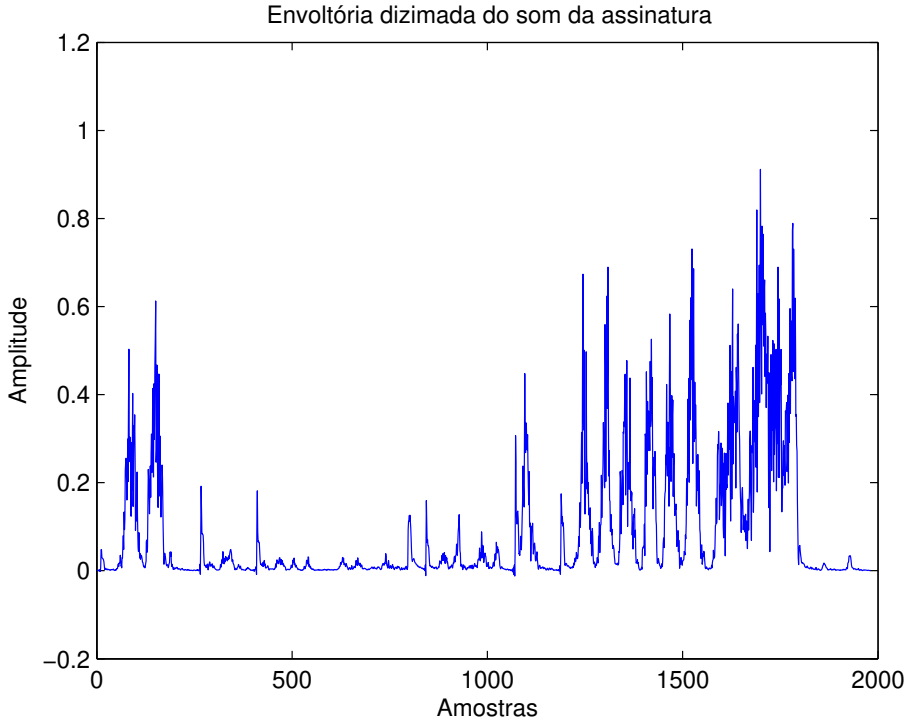


Fig. 4.18: Envoltória do sinal de som dizimada no tempo.

Normalização da Envoltória:

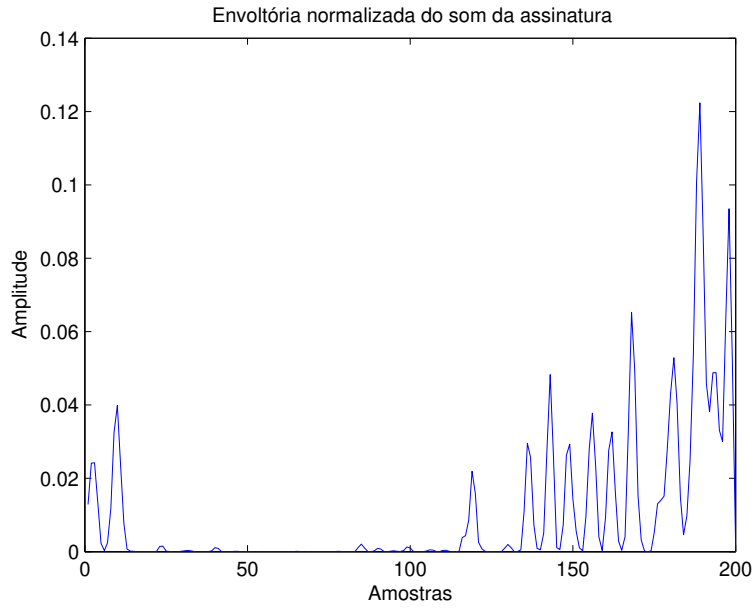
Não é raro encontrar envoltórias de assinaturas feitas pelo mesmo assinante com diferentes amplitudes e comprimentos. Isto justifica um processo de normalização antes que a envoltória seja enviada à etapa onde se calcula o vetor de características.

O primeiro passo para fazer esta normalização é sempre ter o mesmo número de amostras para qualquer assinatura. Para isso, faz-se uma sub-amostragem do sinal $env[n]$ tomando sempre p amostras igualmente espaçadas. A normalização do sinal env é feita utilizando a equação 4.15 [5],

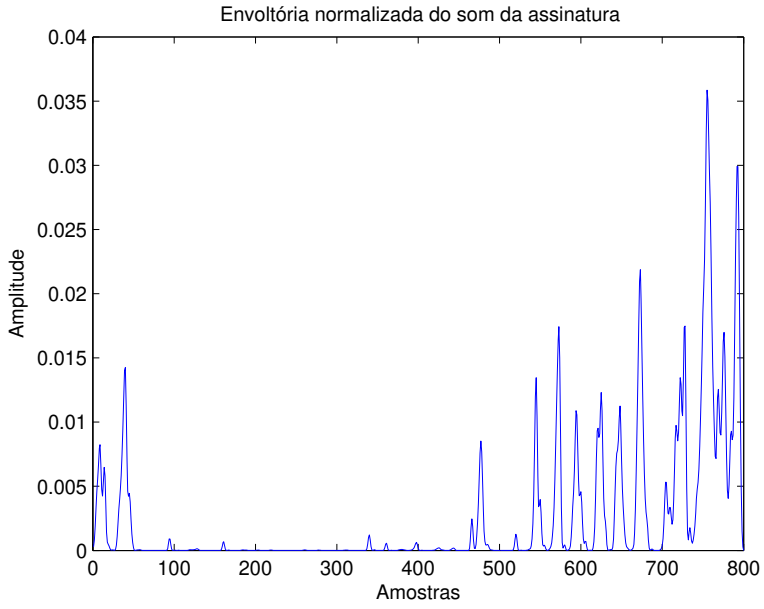
$$Env[n] = \frac{env[n]}{\sqrt{\sum_{n=1}^{p_c} env^2[n]}}, \tag{4.15}$$

onde Env é a envoltória normalizada, p_c é o número de amostras igualmente espaçadas obtidas de $env[n]$. Neste trabalho, fez-se testes com diferentes valores de p_c. Nas figuras 4.19(a) e 4.19(b), são mostradas as envoltórias normalizadas da mesma assinatura com valores de $p_c = 200$ e $p_c = 800$, respectivamente.

Na figura 4.20 são apresentadas duas envoltórias normalizadas de uma mesma assinatura feita pelo mesmo assinante e na figura 4.21 as envoltórias da mesma assinatura feita por um assinante original e um impostor. Deve-se notar que as envoltórias da figura 4.20 são similares enquanto que as envoltórias da figura 4.21 apresentam diferenças. Essas características extraídas das envoltórias do som das assinaturas podem ser utilizadas para classificar entre um usuário legítimo e um impostor.



(a)



(b)

Fig. 4.19: Envoltórias normalizadas do sinal de som. (a) Envoltória normalizada com p_c=200 amostras. (b) Envoltória normalizada com p_c=800 amostras.

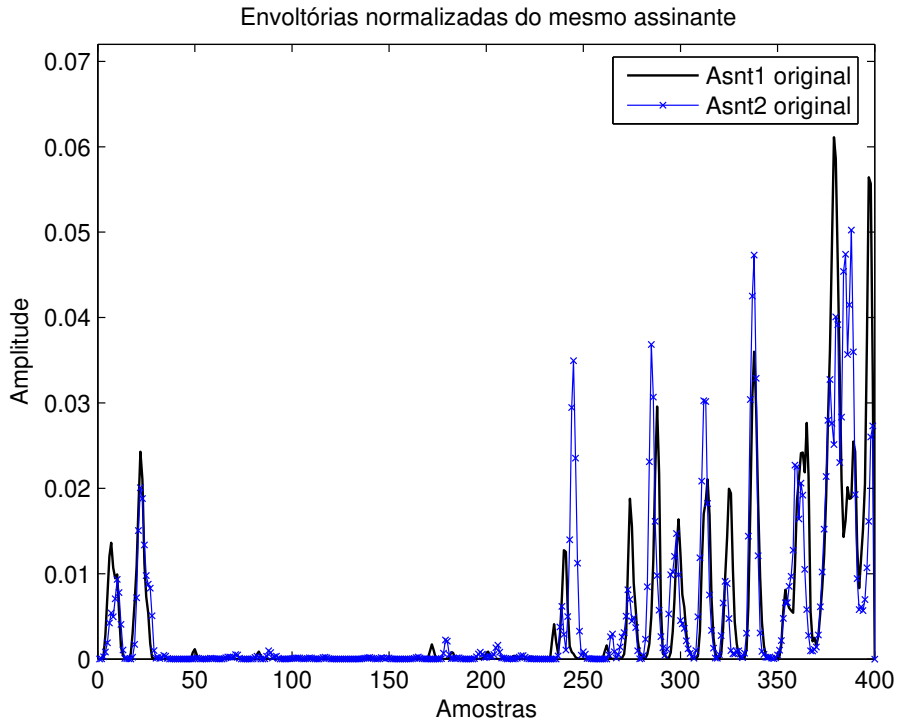


Fig. 4.20: Envoltórias normalizadas da mesma assinaturas feitas pelo mesmo assinante

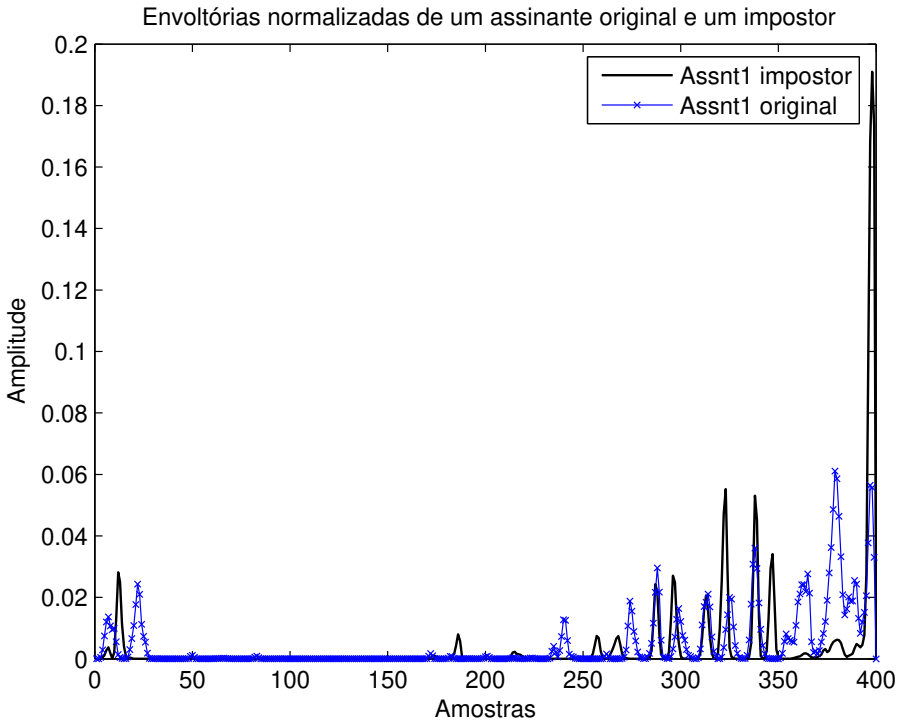


Fig. 4.21: Envoltórias normalizadas da mesma assinatura feitas por um assinate original e um impostor

4.3 Extração de Características

O propósito desta é extrair um vetor de características da envoltória normalizada do sinal de som que representa a assinatura. Cada letra, palavra, linha, rabisco ou traço que compõe uma assinatura é traduzida em picos que formam a envoltória do som da assinatura, isto é, os picos que são parte da envoltória representam os traços produzidos pelo movimento da caneta no momento de assinar. Cada assinatura apresenta uma forma de envoltória dependendo da velocidade, aceleração e pressão do assinante. Assinaturas feitas pelo mesmo assinante têm pequenas variações em suas envoltórias mas assinaturas de diferentes assinantes têm grandes diferenças. Assim vetores de características extraídos da envoltória do sinal de som da assinatura podem ser utilizados para classificar as assinaturas dos diferentes assinantes.

Para alcançar este propósito, neste trabalho são utilizados dois métodos que se fundamentam na representação binária da envoltória do sinal de som.

Método I:

A forma da envoltória do sinal de som de uma assinatura apresenta um conjunto de picos que representam as variações deste sinal no tempo. Estas variações dependem dos traços da assinatura. Um traço de curta duração no tempo é representado por um pico da envoltória de pequena duração e vice-versa. Se o assinante exerce maior pressão na caneta, o pico da envoltória que corresponde a este traço será de amplitude maior. Por esta razão, pode-se dizer que as variações da amplitude e do tempo do conjunto de picos que formam a envoltória estão relacionadas com as características depositadas pelo assinante no momento de fazer a sua assinatura.

Então, para se obter o vetor de características poderia ser suficiente somente capturar as variações locais de maior importância ao longo da direção horizontal da envoltória normalizada que caracteriza a assinatura. Cada variação local da envoltória é usada para caracterizar um traço importante da assinatura (ver figura 4.22).

Neste método, o propósito é localizar as variações locais da envoltória as quais geralmente indicam o aparecimento ou desaparecimento de um traço da assinatura. Isto significa que os pontos mais extremos das variações do sinal (pontos máximos e mínimos) correspondem aos

pontos onde começa e termina as variações locais do sinal original da assinatura.

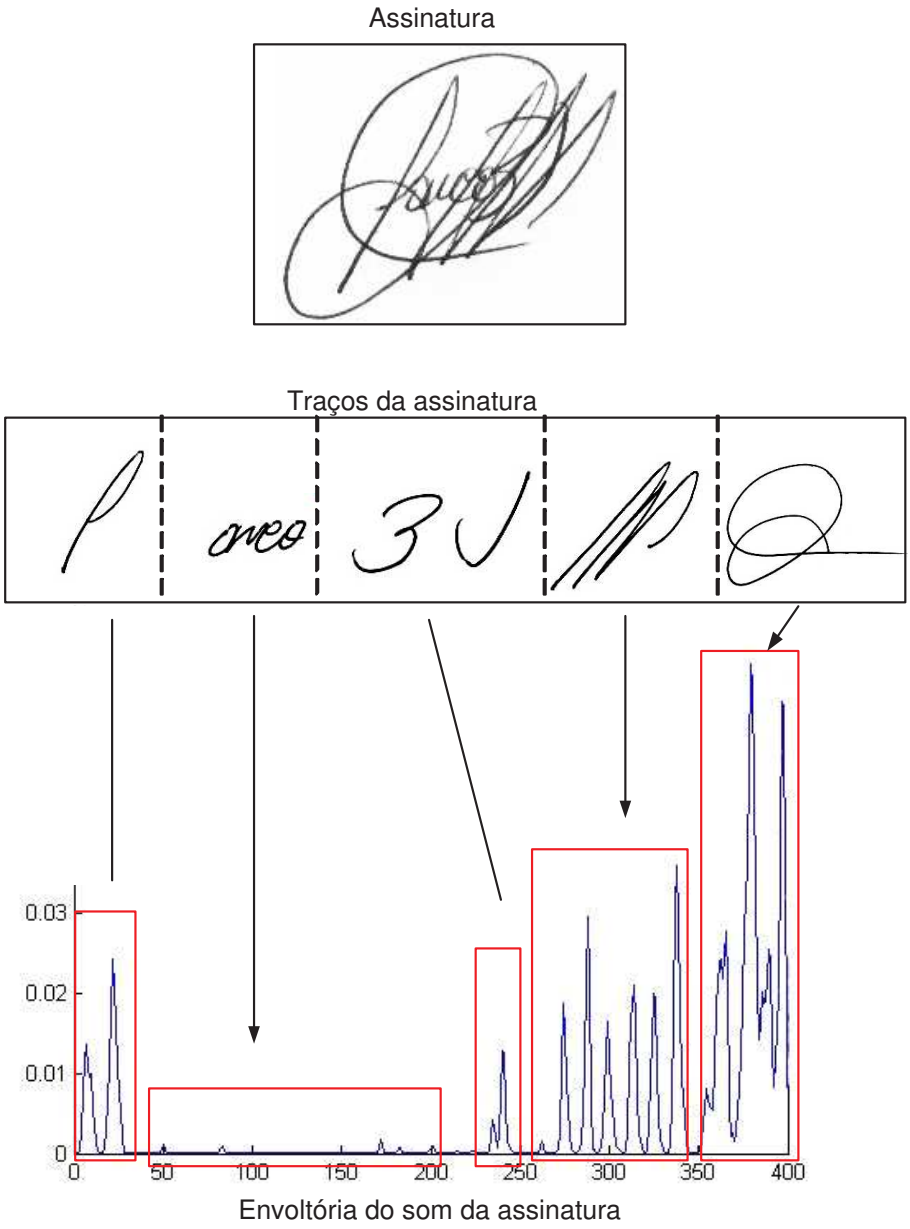


Fig. 4.22: Traços da assinatura que conformam a envoltória

Uma forma de determinar os pontos locais das variações é encontrar os pontos máximos e mínimos da forma da envoltória. Para conseguir isto pode-se utilizar o cálculo da inclinação da envoltória ou inclinação de uma reta que une dois pontos da envoltória em relação a plano horizontal. Esse parâmetro é também conhecido como coeficiente angular de uma reta (ver figura 4.23).

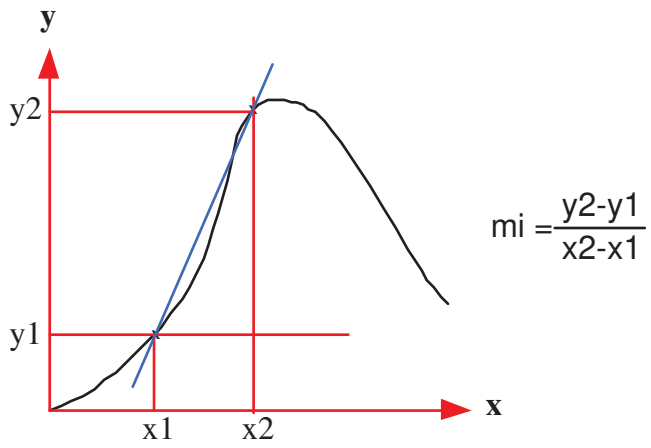


Fig. 4.23: Coeficiente angular de uma reta

Uma vez calculada a inclinação, percorre-se todos os pontos para determinar os pontos máximos e mínimos da envoltória. Um ponto máximo é indicado pela variação da inclinação de um valor positivo a um valor negativo e um mínimo pela variação de negativo para positivo. A figura 4.24 mostra parte da envoltória de uma sinal de som com seus pontos máximos e mínimos.

Um ponto mínimo do sinal denota o aparecimento de um traço da assinatura. Um par de pontos adjacentes, um máximo e um mínimo, indica que existe um traço da assinatura entre eles. Existem alguns pontos máximos e mínimos adjacentes entre os quais a diferença de amplitude da envoltória é muito pequena. Estes pontos locais podem corresponder a características de um sinal relativamente fraco ou de baixa variação em intensidade que dificulta o processo de reconhecimento. Um esquema baseado num limiar é usado para suprimi-los. Neste caso, o limiar é calculado pela média aritmética das amostras da envoltória,

$$lp = \frac{1}{n_e} \sum_{i=1}^{n_e} Env[i], \tag{4.16}$$

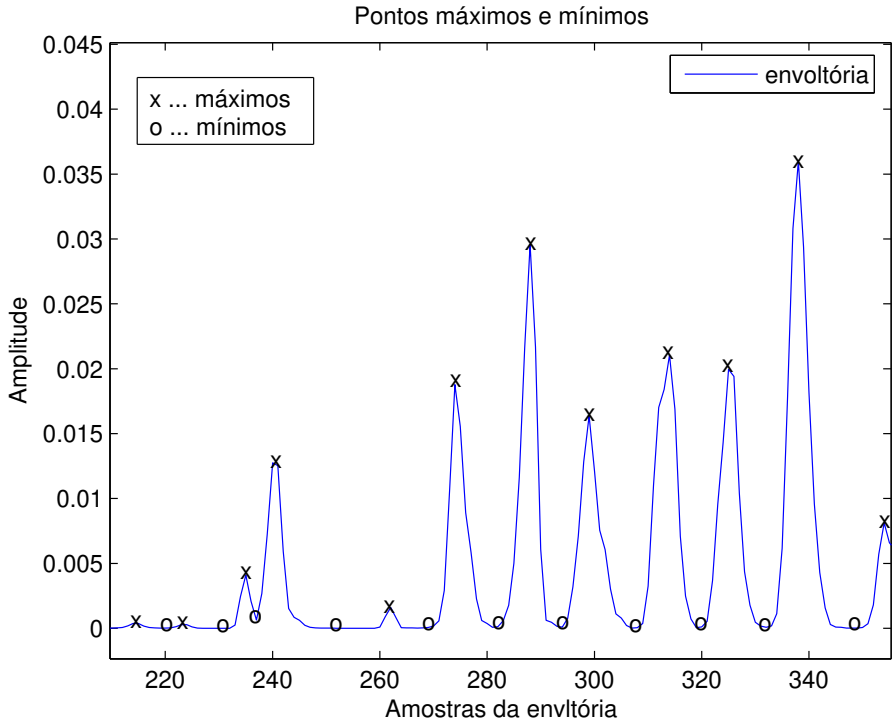


Fig. 4.24: Pontos máximos e mínimos da envoltória

onde lp é o limiar, n_e o número de pontos da envoltória normalizada Env.

Se a diferença de amplitude entre um par de pontos locais adjacentes é menor do que um predeterminado limiar lp, os dois pontos são considerados características irrelevantes do sinal e não são usados como características discriminantes. Somente as características discriminantes do sinal são usadas para formar o vetor de características e, assim, obter um sistema de reconhecimento mais robusto. Na figura 4.25 pode-se observar um diagrama que resume o cálculo dos pontos máximos e mínimos.

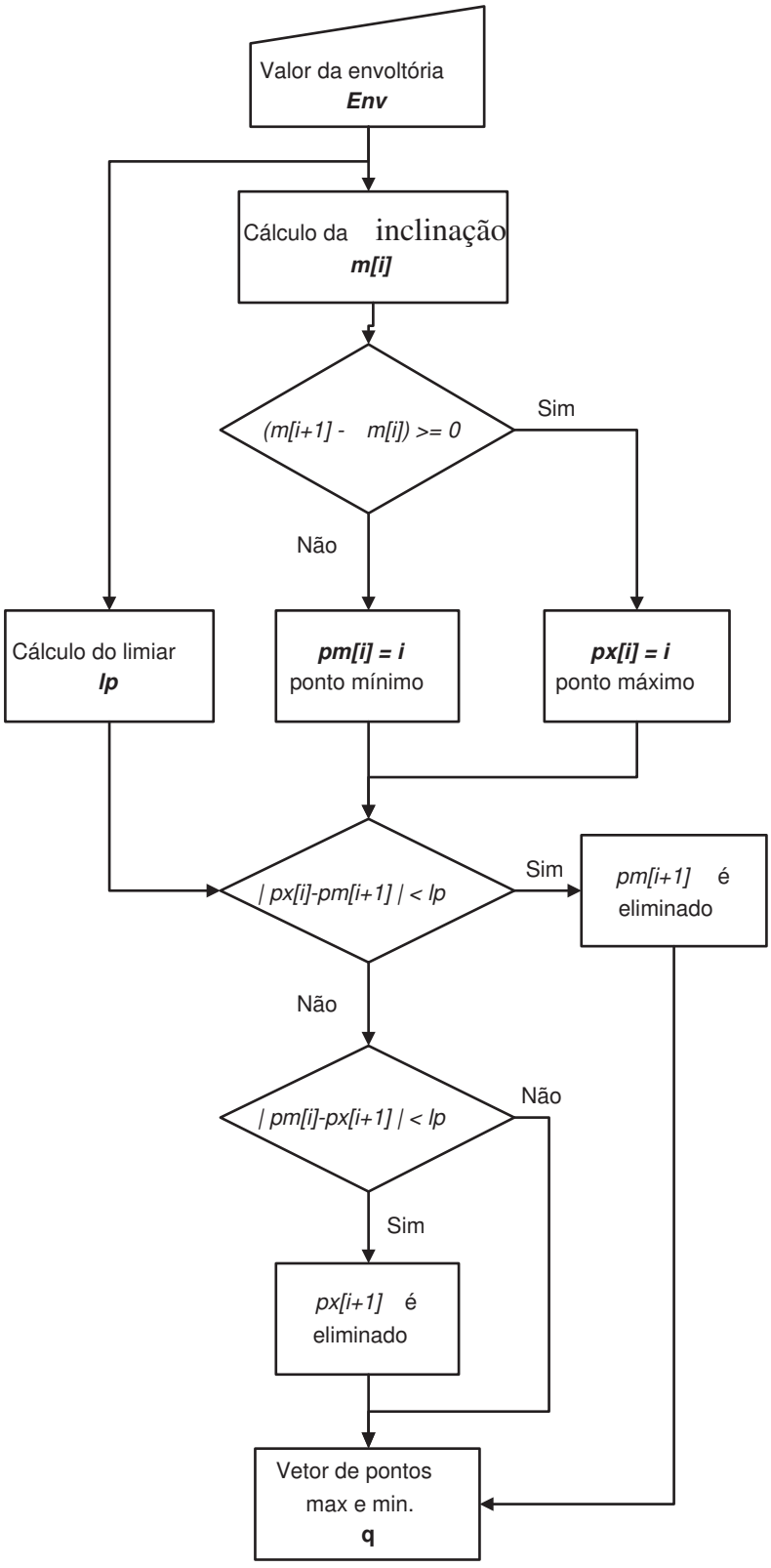


Fig. 4.25: Diagrama das etapas para calcular os pontos máximos e mínimos da envoltória.

Uma vez que os pontos máximos e mínimos foram calculados e eliminados os pontos locais que tem uma diferença de amplitude menor que lp, procede-se a construir o vetor de características da assinatura da seguinte maneira:

Os pontos da envoltória que estão entre um ponto mínimo e um ponto máximo serão representados por um valor de "1" e os pontos que estejam entre um ponto máximo e um mínimo por um valor de "0". Isto faz uma representação binária dos traços que formam a assinatura.

$$\begin{array}{rl} q = & [pm_1 \ldots px_1 \ldots pm_1 \ldots . px_2 \quad \ldots \quad pm_{k-1} \ldots px_k \ldots pm_k] \\ & \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \qquad \downarrow \quad \downarrow \quad \downarrow \\ Vcp = & [1\,1\,1\,1\,1\,0\,0\,0\,0\,1\,1\,1\,1\,1\,0\,0 \ldots \; 0\,0\,1\,1\,1\,1\,1\,0\,0\,0\,0\,0] \end{array} \tag{4.17}$$

Onde pm_i é um ponto máximo da envoltória, px_i é um ponto mínimo,q é o vetor que contem este pontos e Vcp é o vetor de características calculado pelo método I. A figura 4.26 apresenta um diagrama das etapas a seguir para calcular o vetor Vcp.

Na figura 4.27, pode-se observar uma parte da envoltória do som de uma assinatura e seu vetor de características Vcp que codifica os picos da envoltória com uns ou zeros.

De cada uma das assinaturas que formam a base de dados de um assinante é calculado o seu vetor de características.

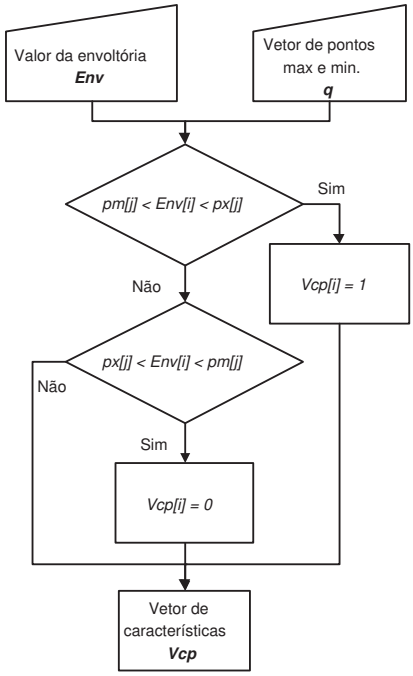


Fig. 4.26: Diagrama das etapas para calcular o vetor de características pelo método I.

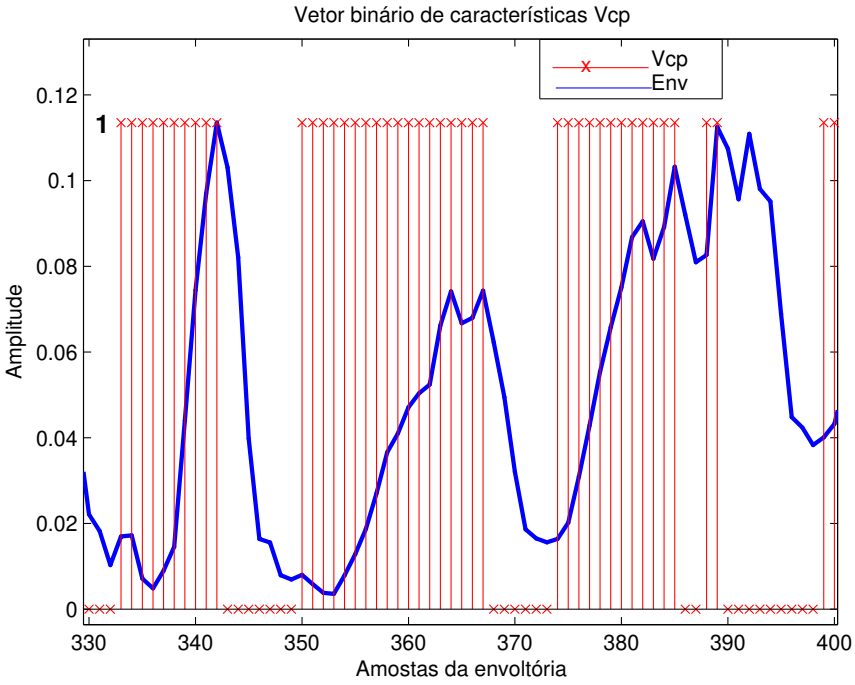


Fig. 4.27: Vetor binário de características da envoltória.

Método II:

Os traços que compõem a assinatura de uma pessoa estão representados pelos picos que formam a envoltória de seu sinal de som. Este método para calcular o vetor de características, está baseado em utilizar três limiares que permitem criar um vetor binário que está relacionado com os traços que formam a assinatura, isto é, os valores binários do vetor vão depender da amplitude e da posição dos picos da envoltória.

O vetor de características será formado por valores tais que os pontos da envoltória maiores ou iguais ao limiar são representados como um e os valores menores como zero.

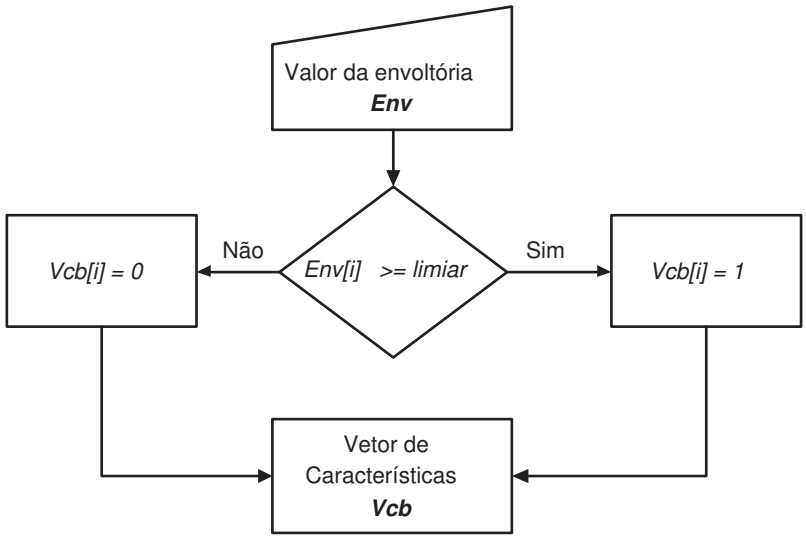


Fig. 4.28: Cálculo dos elementos do vetor de características utilizando o método II

Devido ao fato da assinatura estar formada por diversos traços, os picos que formam a envoltória tem amplitudes diferentes. Cada valor de limiar utilizado permite obter um vetor de características diferentes, visto que os valores destes vetores vão depender do limiar. Com um valor de limar baixo o vetor representará a maioria dos picos e com um valor alto só serão representados os picos dos traços mais fortes.

Para obter um vetor de características que represente de melhor forma a envoltória do sinal, neste método, utiliza-se três limiares lb_1, lb_2 e lb_3. O limiar lb_2 permite representar a maioria dos picos do sinal. Com lb_1 leva-se em conta os picos de amplitude média ou maior e finalmente com lb_3 leva-se em conta só os picos cuja amplitude seja maior que a média.

Utilizando estes limiares calcula-se três vetores auxiliares: Vcb_1, Vcb_2 e Vcb_3. A união destes vetores forma o vetor de características do sinal, chamado Vcb.

Os limares são calculados utilizando as seguintes equações:

$$lb_1 = \frac{1}{n_e} \sum_{i=1}^{n_e} Env[i] \tag{4.18}$$

$$lb_2 = b_1 \cdot lb_1 \tag{4.19}$$

$$lb_3 = b_2 \cdot lb_1 \tag{4.20}$$

Os valores utilizados neste trabalho para as constantes são: $b_1 = 0,5$ e $b_2 = 2$.

A geração dos três vetores auxiliares de características da assinatura ocorre da seguinte maneira :

Os pontos da envoltória que sejam maiores em amplitude do que lb_1 são representados como um e os pontos cuja amplitude seja menor do que lb_1 são representados como zero. Este procedimento repete-se para os limiares lb_2 e lb_3 para determinar os vetores auxiliares de características Vcb_2 e Vcb_3, conforme são mostrados a seguir,

$$Vcb_1[i] = \begin{cases} 1 & se\, Env[i] \geq lb_1, \\ 0 & se\, Env[i] < lb_1. \end{cases} \quad i = 1, ..., n_e \tag{4.21}$$

$$Vcb_2[i] = \begin{cases} 1 & se\, Env[i] \geq lb_2, \\ 0 & se\, Env[i] < lb_2. \end{cases} \quad i = 1, ..., n_e \tag{4.22}$$

$$Vcb_3[i] = \begin{cases} 1 & se\, Env[i] \geq lb_3, \\ 0 & se\, Env[i] < lb_3. \end{cases} \quad i = 1, ..., n_e. \tag{4.23}$$

O vetor de características da assinatura resulta da união dos vetores auxiliares como se indica na equação 4.24.

$$Vcb = [\,Vcb_1 \cup Vcb_2 \cup Vcb_3\,] \tag{4.24}$$

Nas figuras 4.29, 4.30, e 4.31 pode-se observar uma parte da envoltória do som de uma assinatura e seus vetores de características Vcb_1, Vcb_2 e Vcb_3 que codifica os picos da envoltória como uns ou zeros.

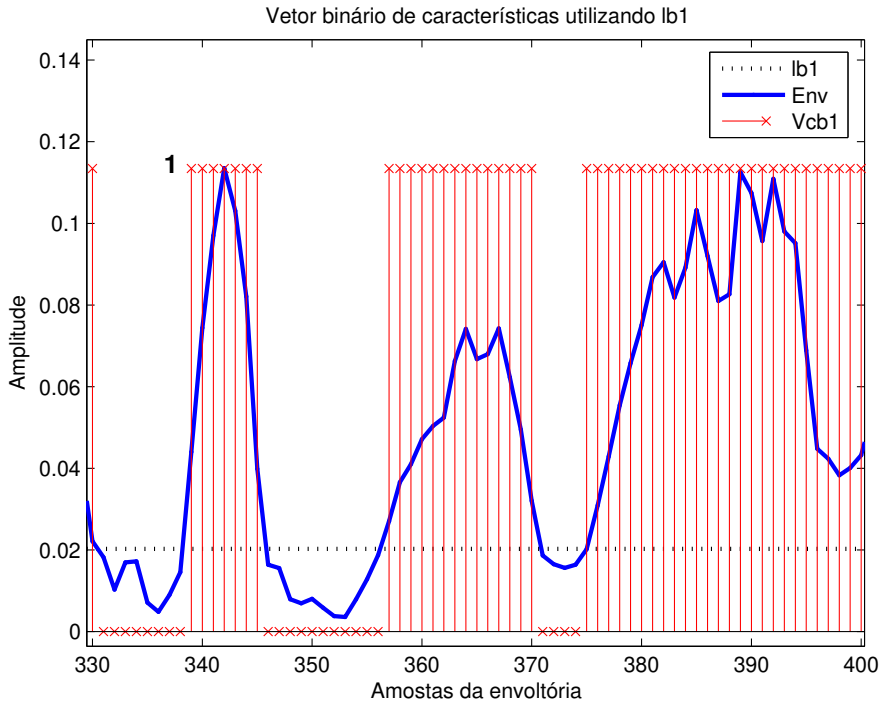


Fig. 4.29: Vetor binário de características calculado utilizando lb_1.

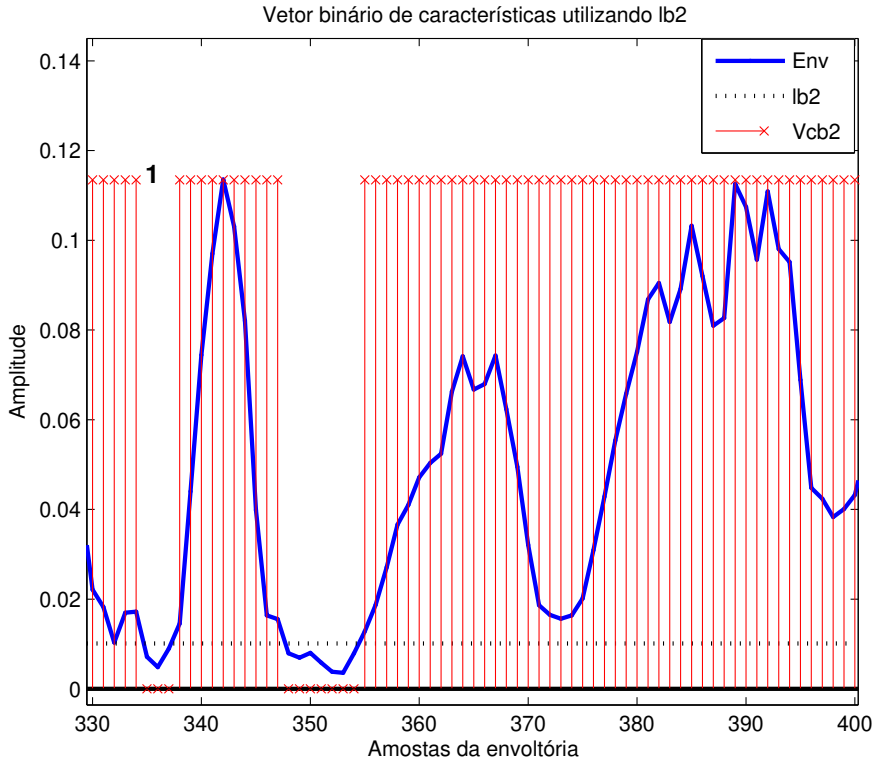


Fig. 4.30: Vetor binário de características calculado utilizando lb_2.

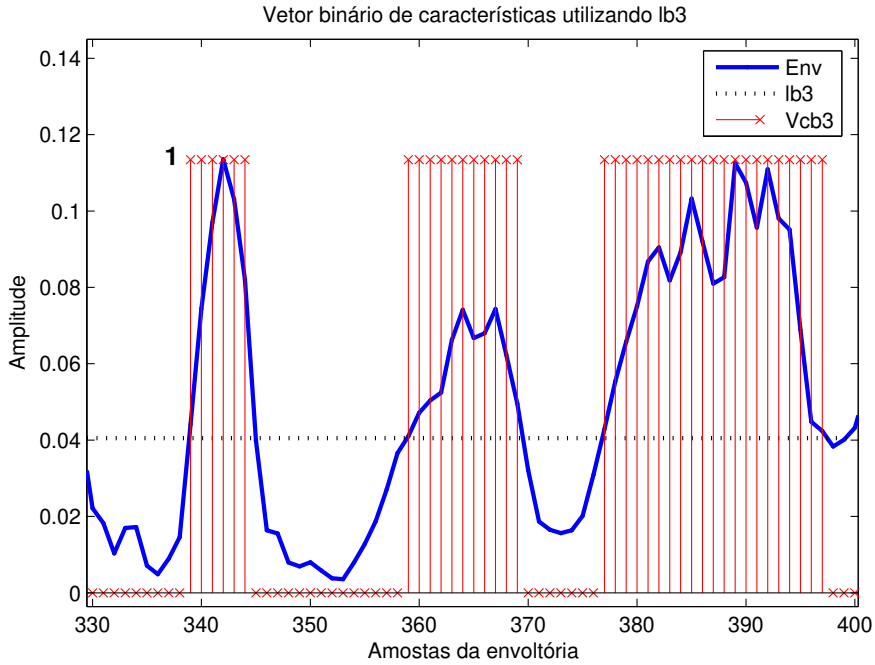


Fig. 4.31: Vetor binário de características calculado utilizando lb_3.

4.4 Comparação:

Conclue-se que duas assinaturas pertencem ao mesmo assinante ou provêm de uma mesma classe comparando-se as similaridades entre os seus sinais característicos correspondentes. Esta similaridade poderia ser expressa como a distância entre um par desses sinais característicos mas esse procedimento tem mostrado que não é conveniente. Uma abordagem de duas etapas é empregada nesse trabalho para realizar a comparação [41]. Estas etapas são:

- Um vetor de características é extraído do sinal original e é transformado em um vetor binário (vetor de características binário).
- A similaridade entre um par de vetores de características binários é calculada usando uma operação lógica OR exclusivo XOR.

Um procedimento para transformar as características originais provenientes de um sinal em amplitude de uma envoltória para uma seqüência de 0's e 1's é apresentado na equação 4.17. O resultado dessa transformação é um vetor de características binário Vc [42]. Esta seqüência binária apresenta o mesmo comprimento do sinal da envoltória e sofre uma transição de 1 para 0 ou de 0 para 1 nos pontos máximos ou mínimos, respectivamente, do sinal de características original.

A função de similaridade entre dois vetores de características é definida como uma distância Dp calculada pela seguinte equação:

$$Dp = \frac{1}{L}\left(Vcb^1 \oplus Vcb^2\right), \tag{4.25}$$

onde Vcb^1 e Vcb^2 são dois vetores de características binários distintos, L é o comprimento dos vetores de características, $\oplus$ é um operador OR exclusivo (XOR). Como a operação XOR é utilizada para uma medida de similaridade e entre dois vetores de características, esta distância vai apresentar um valor menor quando os vetores comparados forem mais semelhantes e vice-versa.

Para que esta medida de similaridade seja representativa e mais eficiente, é importante que os vetores de características binários sejam invariantes a translação, escala, e rotação.

No presente trabalho, o sinal da envoltória que é transformado num vetor binário é normalizado tanto em amplitude quanto em número de amostras de modo que o vetor de características binário é invariante em relação à escala. O vetor de características também é invariante à rotação, já que o sinal capturado é o som da assinatura. Por outro lado, a medida de similaridade entre dois vetores de característica é afetada pela translação. Ainda que dois vetores sejam semelhantes, se no momento de compará-los, um deles estiver deslocado em relação ao outro, o resultado dessa comparação será um valor muito maior do que aquele, quando os dois vetores estiverem alinhados.

Uma forma de diminuir o efeito da translação, é encontrar os pontos de início-fim do sinal de som para eliminar as amostras não necessárias do sinal. Apesar disso, ocorre uma pequena translação de um vetor de características com respeito ao outro. Para corrigir este problema, dividi-se o vetor de características binário em blocos de K_c amostras. Por este procedimento, calcula-se a distância de similaridade de pequenos blocos do vetor de características denominados vc. A soma total destes valores é o valor da distância Di [41].

$$Di = \frac{1}{L}\sum_{j=1}^{N} vc_j^1 \oplus vc_j^2. \tag{4.26}$$

Para comparar dois blocos vc_1 e vc_2 primeiro deslocamos o bloco vc_2 uma amostra e computa-se a semelhança entre eles. Depois de vários deslocamentos tanto à direita quanto à esquerda, a menor das distancias de semelhança é considerada como o valor de distancia. O número de deslocamentos é igual a 1% do valor de L com um valor mínimo de 4 amostras.

Processo de Decisão

A medida de similaridade obtida através da comparação dos vetores de características de duas assinaturas, é utilizada na tomada de decisão para definir se um sinal que está sendo analisado pertence a uma assinatura feita pelo assinante legitimo ou um impostor. No processo de decisão, deve-se definir um valor limite à distância de semelhança, o qual represente a fronteira de decisão que estabeleça quando uma assinatura é verdadeira ou falsa.

Para alcançar este objetivo, segue-se os seguintes passos:

Primeiro: Utilizando o conjunto de assinaturas de treinamento que é formado por assinaturas verdadeiras e assinaturas de impostor, calcula-se a distância inter-classe e um usuário, ou seja, encontra-se as distâncias de semelhanças entre os vetores de características binários das assinaturas verdadeiras.

Sejam S_i e S_j dois sinais verdadeiros de som de um mesmo assinante. Determina-se as envoltórias normalizadas $Env's$ e seus respectivos vetores de características,

$$\begin{aligned} &S_i \Rightarrow Env_i \Rightarrow Vc_i = [vc_1^i, vc_2^i, ..., vc_N^i] \\ &S_j \Rightarrow Env_j \Rightarrow Vc_j = [vc_1^j, vc_2^j, ..., vc_N^j] \\ &Dv_{i,j} = Vc_i \oplus Vc_j \;\; com \;\; i \neq j \end{aligned} \tag{4.27}$$

onde, $Dv_{i,j}$ é a distância entre as assinaturas verdadeiras i e j.

Segundo: Calcula-se a distância de cada vetor de características das assinaturas verdadeiras em relação ao vetor de características das assinaturas do mesmo assinante feitas por um impostor.

Sejam S_i o sinal verdadeiro de som de um assinante e I_j o sinal da assinatura do mesmo assinante feito por um impostor. Determina-se as envoltórias normalizada $Env's$ e seus respectivos vetores de características,

$$\begin{aligned} &S_l \Rightarrow Env_l \Rightarrow Vc_l = [vc_1^l, vc_2^l, ..., vc_N^l] \\ &I_k \Rightarrow Env_k \Rightarrow Ic_k = [ic_1^k, ic_2^k, ..., ic_N^k] \\ &Di_{l,k} = Vc_l \oplus Ic_k \end{aligned} \tag{4.28}$$

onde, $Di_{l,k}$ é a distância entre a assinatura verdadeira l e a assinatura falsa k.

Terceiro: Em função dessas distâncias entre os vetores de características de assinaturas verdadeiras e falsas é obtido o diagrama da distribuição destas distâncias e mostrado na figura 4.32.

A partir dos valores do histograma são calculadas a média (m) e a variância (σ) das distâncias verdadeiras e impostoras para que sejam ajustados a uma distribuição gaussiana utilizando a equação 2.20. Esta gráfica representa a distribuição dos valores das distâncias

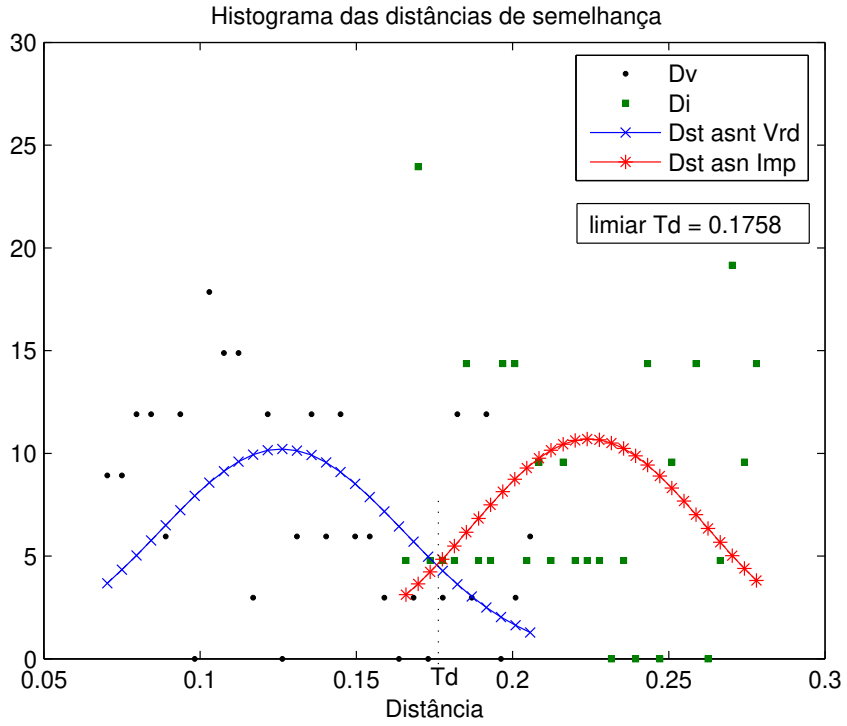


Fig. 4.32: Limite de Td

entre assinaturas verdadeiras e a distribuição das distâncias entre as assinaturas verdadeiras e falsas.

O ponto onde as duas curvas gaussianas se cruzam é o valor limite Td que divide os valores que podem ser considerados como a distâncias de semelhança entre assinaturas verdadeiras e entre assinaturas verdadeiras e falsas.

Uma forma de calcular o ponto Td é utilizar a fórmula que define as gaussianas. Se y_1 é a representação matemática da distribuição gaussiana das distâncias Dv com sua respectiva média m_1 e variância σ_1, e se y_2 a representação das distâncias Di com média m_2 e variância σ_2; então o ponto Td ocorre quando $y_1 = y_2$ e $x_1 = x_2$,

$$y_1 = \frac{1}{2\pi\sigma_1}\, e^{\left(-\frac{(x_1-m_1)^2}{2\sigma_1^2}\right)} \tag{4.29}$$

$$y_2 = \frac{1}{2\pi\sigma_2}\, e^{\left(-\frac{(x_2-m_2)^2}{2\sigma_2^2}\right)} \tag{4.30}$$

se $y_1 = y_2$,

$$\begin{aligned}
\frac{1}{2\pi\sigma_1}\, e^{\left(-\frac{(x_1-m_1)^2}{2\sigma_1^2}\right)} &= \frac{1}{2\pi\sigma_2}\, e^{\left(-\frac{(x_2-m_2)^2}{2\sigma_2^2}\right)} \\
\sigma_2\; e^{\left(-\frac{(x_1-m_1)^2}{2\sigma_1^2}\right)} &= \sigma_1\; e^{\left(-\frac{(x_2-m_2)^2}{2\sigma_2^2}\right)} \\
\ln(\sigma_2) - \frac{(x_1-m_1)^2}{2\sigma_1^2} &= \ln(\sigma_1) - \frac{(x_2-m_2)^2}{2\sigma_2^2} \\
\frac{1}{\sigma_1^2}\, x_1^2 - \frac{2m_1}{\sigma_1^2}x_1 + \frac{m_1^2}{\sigma_1^2} - 2\ln(\sigma_2) &= \frac{1}{\sigma_2^2}\, x_2^2 - \frac{2m_2}{\sigma_2^2}x_2 + \frac{m_2^2}{\sigma_2^2} - 2\ln(\sigma_1)
\end{aligned} \tag{4.31}$$

define-se as variáveis,

$$a_1 = \frac{1}{\sigma_1^2},\;\; b_1 = -\frac{2m_1}{\sigma_1^2},\;\; c_1 = \frac{m_1^2}{\sigma_1^2} - 2\ln(\sigma_2) \tag{4.32}$$

$$a_2 = \frac{1}{\sigma_2^2},\;\; b_2 = -\frac{2m_2}{\sigma_2^2},\;\; c_2 = \frac{m_2^2}{\sigma_2^2} - 2\ln(\sigma_1) \tag{4.33}$$

se $x_1 = x_2 = x$,

$$\begin{aligned}
a_1x_1^2 + b_1x_1 + c_1 &= a_2x_2^2 + b_2x_2 + c_2 \\
(a_1 - a_2)x^2 + (b_1 - b_2)x + (c_1 - c_2) &= 0
\end{aligned} \tag{4.34}$$

a solução desta equação determina o ponto do limiar Td.

Quarto: Quando uma assinatura de teste St é analisada pelo sistema, calcula-se as distâncias desta assinatura com respeito as assinaturas verdadeiras Dt.

$$\begin{aligned} &S_l \Rightarrow Env_l \Rightarrow Vc_l = [vc_1^l, vc_2^l, ..., vc_N^l] \\ &St_k \Rightarrow Env_k \Rightarrow Vct_k = [vct_1^k, vct_2^k, ..., vct_N^k] \\ &Dt_{l,k} = Vc_l \oplus Vct_k \end{aligned} \tag{4.35}$$

A média aritmética de $Dt_{l,k}$ é calculada e comparada com o valor limiar Td. Se este valor é menor que o valor do limiar Td, a assinatura pertence a um usuário legitimo mas se é maior, a assinatura foi feita por um usuário impostor.

Este procedimento para determinar se a assinatura pertence a um usuário verdadeiro ou impostor é realizado utilizando os dois métodos de cálculo dos vetores de características. Os resultados obtidos são apresentados e avaliados no próximo capítulo.

Capítulo 5

Resultados e Conclusões

Neste capítulo são apresentados os experimentos com os seus respectivos resultados e as conclusões deste trabalho em função dos resultados obtidos e as sugestões para futuros trabalhos nessa linha de pesquisa.

5.1 Experimentos

Os experimentos realizados neste capítulo estão destinados a avaliar os resultados obtidos pelos dois métodos empregados para calcular os vetores de características que foram descritos no capítulo anterior. Estes métodos fazem uma representação binária da envoltória das assinaturas verdadeiras e falsas que estão armazenadas na base de dados.

Quando um experimento é realizado, um sinal de teste é tomado da base de dados para ser avaliado e, como este sinal pode pertencer a um assinante verdadeiro ou a um impostor, pode ocorrer uma das quatro seguintes situações de classificação :

1. O sinal de entrada pertence a um usuário legítimo e o classificador considera-o legítimo;

2. O sinal de entrada pertence a um usuário impostor e o classificador considera-o impostor;

3. O sinal de entrada pertence a um usuário legítimo e o classificador considera-o impostor;

4. O sinal de entrada pertence a um usuário impostor e o classificador considera-o legítimo.

Pode-se observar que, nas situações 3 e 4, o classificador cometeu um erro de decisão. Nestes casos, o desempenho da metodologia é medido pelas taxas de erro FRR e FAR, que foram descritas no capítulo 2 (seção 2.1.3) e podem ser estimadas pelas equações 2.2 e 2.3.

Os resultados são apresentados em duas partes, de acordo com o método utilizado para o cálculo dos vetores de características.

5.1.1 Resultados Obtidos pelo Método I.

Vários experimentos são realizados para verificar o impacto nas taxas de FAR e FRR do diferente número de amostras da envoltória (N) e do diferente número de blocos (K) por envoltória. Os resultados destes experimentos se podem observar nas tabelas 5.1 a 5.8, que mostram as taxas FAR e FRR obtidas em função de N e K. Para cada tabela, escolheu-se um valor de N e testou-se para diferente número de blocos por envoltória.

Método I **Envoltória com $N = 200$ amostras**		
Número de blocos (K)	**FAR(%)**	**FRR(%)**
4	16.73	11.27
5	16.55	11.45
6	15.64	11.82
8	16.18	11.45
10	16.73	12.00
15	18.18	14.00

Tab. 5.1: Valores das taxas de erro para $N = 200$.

Método I **Envoltória com $N = 400$ amostras**		
Número de blocos (K)	**FAR(%)**	**FRR(%)**
4	14.91	18.36
5	13.64	17.82
6	16.00	16.18
8	13.64	17.27
10	15.64	16.00
15	15.82	14.00

Tab. 5.2: Valores das taxas de erro para $N = 400$.

Método I **Envoltória com $N = 500$ amostras**		
Número de blocos (K)	**FAR(%)**	**FRR(%)**
4	14.55	22.55
5	14.55	22.00
6	14.91	19.64
8	13.64	17.27
10	15.09	18.55
15	14.91	16.55

Tab. 5.3: Valores das taxas de erro para $N = 500$.

Método I **Envoltória com $N = 600$ amostras**		
Número de blocos (K)	**FAR(%)**	**FRR(%)**
4	14.73	20.00
5	14.55	19.09
6	14.36	16.91
8	14.18	18.36
10	16.00	15.82
15	15.27	14.73

Tab. 5.4: Valores das taxas de erro para $N = 600$.

Método I **Envoltória com $N = 800$ amostras**		
Número de blocos (K)	**FAR(%)**	**FRR(%)**
4	12.36	16.73
5	12.18	16.73
6	11.46	15.27
8	11.64	15.64
10	12.18	14.55
15	13.09	14.91

Tab. 5.5: Valores das taxas de erro para $N = 800$.

Método I **Envoltória com $N = 1000$ amostras**		
Número de blocos (K)	**FAR(%)**	**FRR(%)**
4	13.45	17.64
5	12.55	16.00
6	12.91	14.73
8	11.46	15.82
10	12.55	13.82
15	14.36	14.36

Tab. 5.6: Valores das taxas de erro para $N = 1000$.

Método I **Envoltória com $N = 1200$ amostras**		
Número de blocos (K)	**FAR(%)**	**FRR(%)**
4	14.91	14.00
5	14.00	15.09
6	13.64	14.00
8	13.09	14.64
10	13.82	12.91
15	14.55	13.46

Tab. 5.7: Valores das taxas de erro para $N = 1200$.

Método I **Envoltória com $N = 1500$ amostras**		
Número de blocos (K)	**FAR(%)**	**FRR(%)**
4	17.09	15.82
5	16.91	15.09
6	17.09	13.27
8	16.00	15.27
10	16.36	14.00
15	15.46	11.27

Tab. 5.8: Valores das taxas de erro para $N = 1500$.

Os resultados nas tabelas mostram a influência de N e K sobre os valores de FAR e FRR. Essas taxas variam mais em função do número de amostras da envoltória (N) do que em função do número de blocos (K) por envoltória. Tais variações podem ser vistas nas figuras 5.1 e 5.2.

Analisando as tabelas, nota-se que os menores valores de FRR ocorrem para N=200 e os de FAR ocorrem para N=800. Entretanto, o melhor compromisso entre as duas taxas ocorre para $N = 1000$ e $K = 10$ onde FAR = 12,55% e FRR = 13,82%.

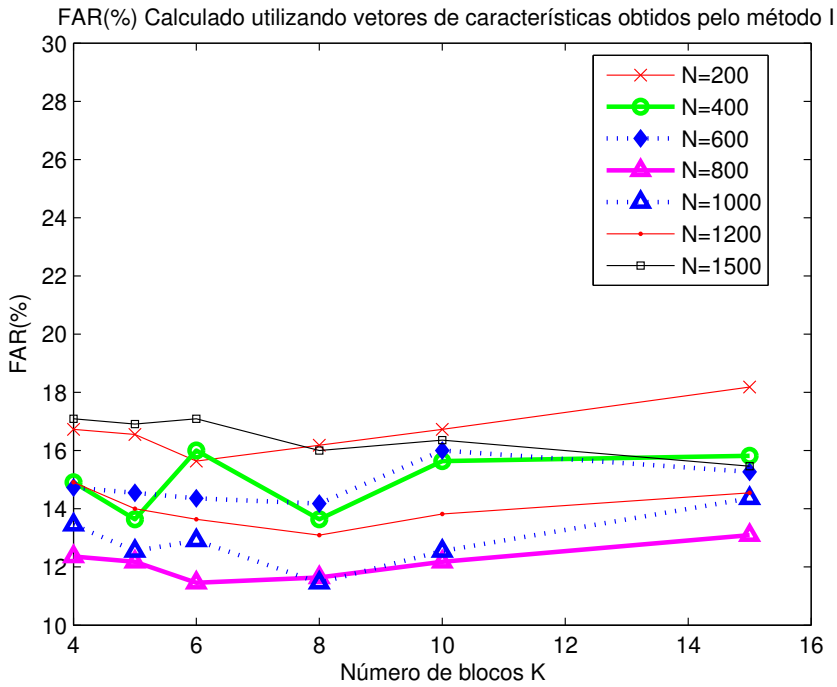


Fig. 5.1: Valores de FAR(%) obtidos pelo Método I.

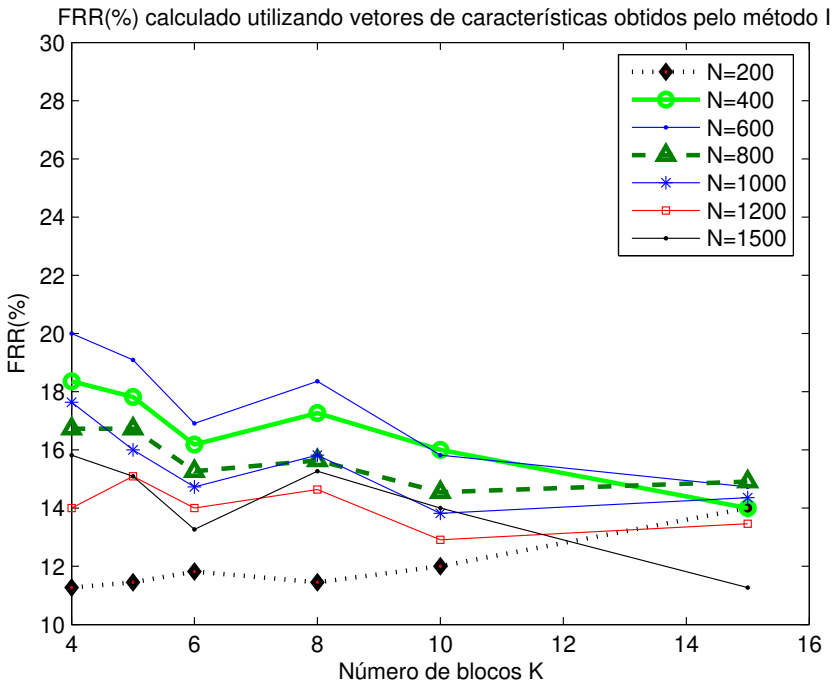


Fig. 5.2: Valores de FRR(%)obtidos pelo Método I.

5.1.2 Resultados Obtidos pelo Método II.

As tabelas 5.9 a 5.15 mostram as variações das taxas FAR e FRR em função do número de amostras (N) da envoltória e do número de blocos (K) por envoltória para os resultados obtidos utilizando o método II de cálculo dos vetores de características.

Para realizar o cálculo das taxas FAR e FRR, a envoltória do sinal de som passa por um processo de reamostragem para obter N amostras e é dividida em blocos de K amostras. Para cada combinação de valores de N e K obtem-se o número de falsas aceitações e o número de tentativas de identificação de impostor, com esses valores calcula-se a taxa FAR como mostra a equação 2.2. A seguir são calculados o número de falsas rejeições e o número de tentativas de identificação de usuário legítimo para obter o valor da taxa FRR como mostra a equação 2.3. Estes valores são apresentados nas seguintes tabelas.

Método II **Envoltória com $N = 200$ amostras**		
Número de blocos (K)	**FAR(%)**	**FRR(%)**
5	31.45	5.45
8	13.09	6.73
10	8.55	8.73
12	5.82	12.55
15	3.09	18.36
20	1.64	28.36
30	0.36	51.18

Tab. 5.9: Valores das taxas de erro para $N = 200$.

Método II **Envoltória com $N = 400$ amostras**		
Número de blocos (K)	**FAR(%)**	**FRR(%)**
5	29.82	6.18
8	12.18	8.00
10	8.55	10.36
12	5.09	12.55
15	2.73	18.55
20	1.27	25.82
30	0.36	50.91

Tab. 5.10: Valores das taxas de erro para $N = 400$.

Método II **Envoltória com $N = 500$ amostras**		
Número de blocos (K)	**FAR(%)**	**FRR(%)**
5	26.00	5.82
8	13.46	8.00
10	7.64	9.64
12	4.73	12.73
15	2.55	18.91
20	1.27	37.27
30	0.36	50.91

Tab. 5.11: Valores das taxas de erro para $N = 500$.

Método II **Envoltória com $N = 600$ amostras**		
Número de blocos (K)	**FAR(%)**	**FRR(%)**
5	26.00	5.64
8	12.18	7.82
10	7.27	9.82
12	4.73	12.73
15	2.36	19.09
20	1.09	28.18
30	0.91	45.09

Tab. 5.12: Valores das taxas de erro para $N = 600$.

Método II **Envoltória com $N = 800$ amostras**		
Número de blocos (K)	**FAR(%)**	**FRR(%)**
5	25.64	6.18
8	12.36	7.82
10	7.82	9.64
12	4.55	12.18
15	2.55	18.18
20	1.09	27.64
30	0.36	54.55

Tab. 5.13: Valores das taxas de erro para $N = 800$.

Método II **Envoltória com $N = 1000$ amostras**		
Número de blocos (K)	**FAR(%)**	**FRR(%)**
5	24.90	5.46
8	13.46	7.64
10	8.55	10.00
12	4.18	12.00
15	2.18	18.73
20	1.27	29.46
30	0.18	59.27

Tab. 5.14: Valores das taxas de erro para $N = 1000$.

Método II **Envoltória com $N = 1200$ amostras**		
Número de blocos (K)	**FAR(%)**	**FRR(%)**
5	27.46	5.46
8	13.27	7.64
10	8.73	9.82
12	4.73	12.73
15	2.36	19.73
20	1.27	30.73
30	0.18	61.46

Tab. 5.15: Valores das taxas de erro para $N = 1500$.

Analisando as tabelas obtidas para o método II, nota-se que as taxas FAR e FRR são fortemente afetadas pelo número de blocos por envoltória enquanto que essas taxas sofrem menos influência pelo número de amostras da envoltória. As figuras (5.3) e (5.4) mostram a influência desses parâmetros sobre as taxas de erros. O melhor compromisso entre as duas taxas ocorre para N=200 e K=10 onde FAR=8,55% e FRR=8,73%. A taxa de erro igual (*Equal Error Rate* - ERR) pode-se obter a partir da curva Característica Operacional do Receptor (ROC) (ver 5.5) quando a taxa FAR é igual à FRR. Para este caso o valor de ERR é 8,6%.

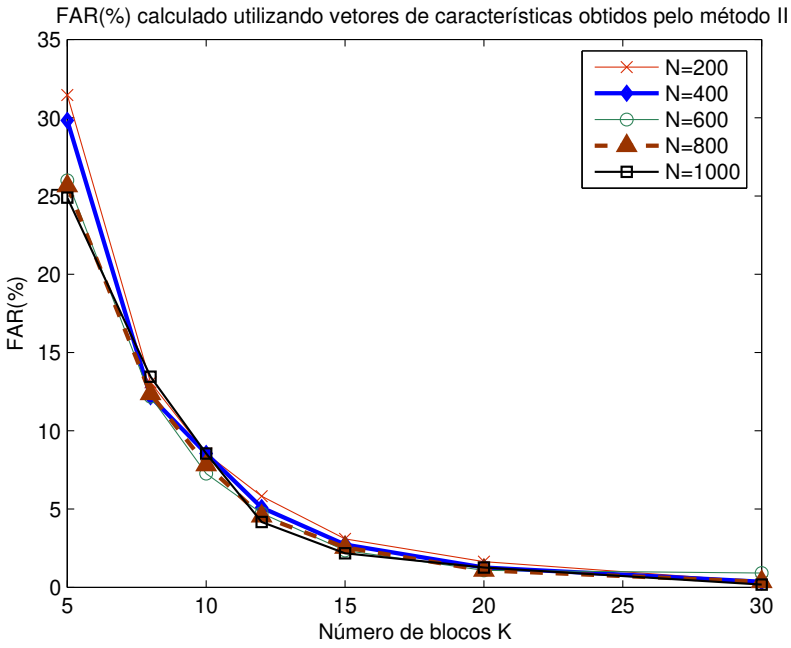


Fig. 5.3: Valores de FAR(%) obtidos pelo Método II.

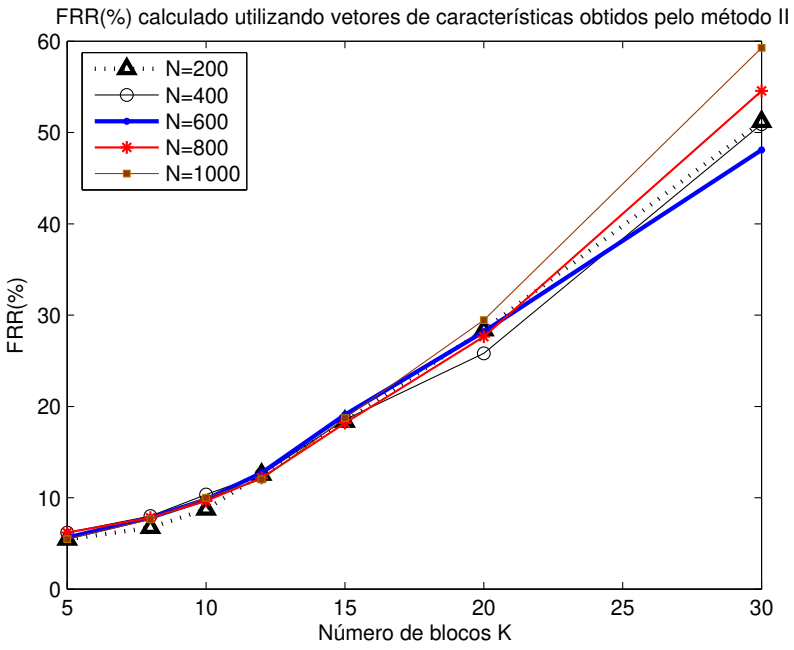


Fig. 5.4: Valores de FRR(%) obtidos pelo Método II.

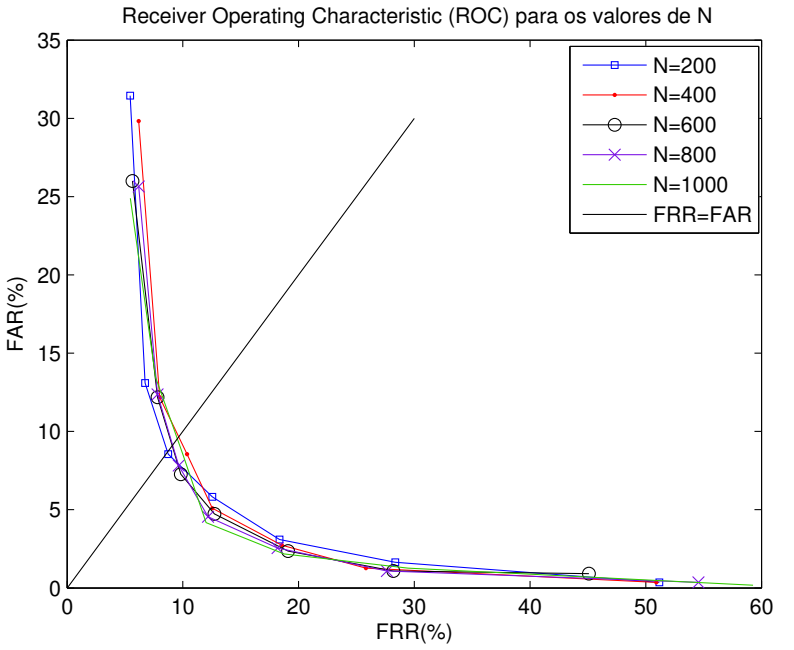


Fig. 5.5: Curva ROC para os diferentes valores de N.

5.1.3 Conjunto de Treinamento

A quantidade de amostras presentes no conjunto de treinamento é um aspecto crucial, pois à medida que esta quantidade é reduzida, os erros de classificação aumentam [43]. Nesta metodologia, a quantidade de amostras presentes no conjunto de treinamento é de 10 amostras. Para verificar o impacto da quantidade de amostras do conjunto de treinamento, um experimento é realizado reduzindo a quantidade destas amostras. Na tabela 5.16 é mostrado o comportamento da FRR e da FAR com a redução da quantidade de amostras no conjunto de treinamento para os melhores resultados dos dois métodos de cálculo do vetor de características.

Método	**Taxas**	**Número de amostras de treinamento**			
		10	8	5	3
Método I	FAR (%)	12.55	16.03	27.03	32.49
$N = 1000$, $K = 10$	FRR(%)	13.82	16.94	30.09	37.83
Método II	FAR (%)	8.55	9.73	16.58	25.77
$N = 200$, $K = 10$	FRR(%)	8.73	12.43	20.18	29.73

Tab. 5.16: Taxas de erros em função do número de amostras de treinamento.

Como se observa na tabela 5.16, a redução da quantidade de amostras do conjunto de treinamento influência em ambas as taxas. À medida que a quantidade de amostras é reduzida, a capacidade de discriminar que tem os vetores de características também é reduzida, causando o aumento de FRR e FAR.

5.2 Discussão

De acordo com os experimentos realizados, algumas observações podem ser feitas:

- Utilizando o conjunto de características formado por vetores calculados pelo método II, se obteve melhores resultados, comparados com o método I, isto se deve a que o método II utiliza um maior número de elementos no seu vetor de características, o que reflete melhor a maneira como cada usuário assina.

- A forma da assinatura é um ponto importante, pois o grau da dificuldade da assinatura

dá como resultado um vetor de características mais discriminantes. Assim, assinaturas que só contém o nome do assinante são mais fáceis de imitar do que as assinaturas que contém o nome e um conjunto de linhas e traços.

- O número de amostras presentes no conjunto de treinamento é um ponto crucial, pois quanto maior esta quantidade, a média e a variância obtidas do conjunto de distâncias entre amostras verdadeiras e falsas são mais confiáveis com relação às características da assinatura. As taxas de FRR e FAR aumentam com a redução do número de amostras que formam o conjunto de treinamento (ver tabela 5.16).

5.3 Conclusões

5.3.1 Contribuições

Nesta dissertação se apresenta uma metodologia de verificação da identidade de uma pessoa via reconhecimento de assinaturas. Esta metodologia consiste em obter e processar medidas biométricas extraídas do som que se produz no momento em que uma pessoa assina.

Entre as contribuições produzidas por este trabalho pode-se mencionar o desenvolvimento de: um dispositivo próprio para a aquisição do som de assinaturas e de um programa de computador que permite a captura deste som e o cadastramento dos assinantes verdadeiros e falsos neste processo de autenticação. Além disso, utilizando este software gerou-se um conjunto amostras de assinaturas que formam uma base de dados que poderá ser utilizada em futuros trabalhos que envolvam aplicações com esta técnica de reconhecimento.

Utilizar o som da assinatura para realizar identificação pessoal não é novo. Nos trabalhos de Li [5] e Soule [9], que foram apresentados no capítulo 2, descreve-se vários experimentos utilizando o som da assinatura e da escrita. Baseados em seus resultados, Soule e Li mencionam que é possível desenvolver um sistema de verificação pessoal que utilize como dados de entrada o som.

No trabalho de Soule realizou-se um experimento no qual se mediu a similaridade de assinaturas usando características extraídas do sinal de som. Estas características são comparadas obtendo uma medida de similaridade através de um algoritmo de *Dynamic Time Warpping*

(DTW). Esta medida de similaridade é representada por um valor no intervalo de 0 a 1, onde a maior similaridade entre duas amostras é representada com um valor de 1 e a não similaridade com 0. O experimento realizado incluiu, 3 usuários, dos quais se obtiveram 3 assinaturas de cada um e 4 falsificações para cada assinante. Os melhores resultados foram de 0.92 de similaridade para assinaturas feitas pelo mesmo assinante e 0.314 de similaridade para assinaturas feitas pelo assinante verdadeiro e impostor.

No trabalho realizado por Li, realizou-se a verificação pessoal usando como informação de entrada o som da escrita. Utilizou-se, 5 escritores com um total de 50 amostras verdadeiras, 50 forjadas, 100 sons arbitrários e 200 seqüências aleatórias de igual comprimento. Li, utiliza todas as amostras da envoltória do sinal de som como características discriminantes para o reconhecimento. O sistema de reconhecimento é realizado utilizando uma rede neural multicamada e o melhor resultado obtido é de 75% de exatidão.

Os trabalhos de Soule e Li, utilizam o som como fonte de informação. Soule utiliza a envoltória extraída do sinal de som de uma assinatura e Li utiliza o som da escrita para identificar ao escritor. No trabalho realizado nesta dissertação, emprega-se características locais extraiadas do som da assinatura para fazer a sua descrição. No lugar de usar os valores da envoltória como característica de reconhecimento, desenvolveu-se dois procedimentos que permitem realizar uma representação binária das características da envoltória, com o fim de obter um vetor que realce as características discriminantes do sinal de som.

A vantagem de obter este vetor binário é utilizar uma operação XOR para calcular uma distância de similaridade entre vetores de usuários verdadeiros e impostores. Esta distância é enviada à etapa de decisão do sistema que realiza a verificação.

Neste trabalho se realizaram vários experimentos destinados a avaliar o desempenho da metodologia de verificação desenvolvida frente a falsificações habilidosas. Para estes experimentos utilizaram-se, 55 usuários, que contribuíram com 11 amostras verdadeiras e 11 imitações para cada um. Como parâmetros de desempenho de nosso sistema de verificação frente a falsificações habilidosas, obteve-se os valores de 8.55% e 8.73% de taxas de erros FAR e FRR, respectivamente.

Na metodologia proposta neste trabalho calcula-se a distância entre vetores de carac-

terísticas binários como uma medida de similaridade. Esta distância pode ser normalizada e obter um valor entre 0 e 1; e assim poder comparar o presente trabalho com o trabalho feito por Soule. Para isto realizou-se um experimento similar ao de Soule. Os melhores resultados obtidos foram: 0.95 de similaridade entre amostras de um mesmo assinante e 0.085 para a distância de similaridade entre vetores de características de assinaturas de um usuário verdadeiro e um impostor. Como pode-se observar os resultados obtidos são melhores que no trabalho de Soule.

No trabalho feito por Li, o melhor resultado obtido é de 75% de exatidão, entendendo por isso que de cada 100 amostras verdadeiras foram reconhecidas adequadamente 75 amostras. Para o trabalho desta dissertação tem-se uma taxa de falsa rejeição de 8.73%, o qual quer dizer, que de cada 100 amostras verdadeiras 8 foram rejeitas o que pode ser representado como uma exatidão de 92%, que é um melhor resultado que o apresentado por Li.

5.3.2 Perspectivas para Trabalhos Futuros

Com relação ao método de verificação será interessante testar outras técnicas de extração de vetores de características do sinal de som, como: HMM ou utilizar lógica nebulosa para comparar seu desempenho na tarefa de classificação. Também poder-se-ia utilizar as técnicas conhecidas como *Dynamic Time Warping* (DTW) para a comparação de *templates*.

É possível desenvolver um sistema de reconhecimento pessoal que utilize uma técnica de reconhecimento baseadas na imagem da assinatura associada à técnica descrita neste trabalho para oferecer um melhor desempenho na tarefa de classificação que utilize assinaturas.

Também se poderia adaptar a metodologia para propósitos criptográficos. Desta forma, gerar-se uma chave criptográfica baseada nas características de sinal de som da assinatura.

Algumas modificações no dispositivo de captura como a redução da aérea na qual é capturado o som poderiam melhorar o desempenho do sistema.

Referências Bibliográficas

[1] Polemi D., "Biometric Techniques: Review and Evaluation of Biometric Techniques for Identification and Authentication, Including an Appraisal of the Areas Where They are Most Applicable", Final Report, Institute of Communication and Computer Systems National Technical University of Athens, 1997. [on-line] Disponível na internet via Web. URL: ftp://ftp.cordis.lu/pub/infosec/docs/biomet.doc. Arquivo consultado em junho de 2005.

[2] Fairhurst M. C., "Signature verification revisited: promoting practical exploitation of biometric technology", Electronics and Communication Engineering Journal, pp. 273 - 280, dezember, 1997.

[3] Mathyas S.M., Stapleton J., "A biometric standard for information management and security", Computers & Security, Vol.19, No 5, pp. 428 - 441, 2000.

[4] Mansfield T., Roethenbaugh G., "Glossary of Biometric Terms", International Association for Biometrics (iAfB), 1999 [on-line] Disponível na internet via Web. URL:http://www.iafb.org.uk/DOC/Glossary_Terms.htm. Arquivo consultado em outubro de 2005.

[5] Li F. F., "Handwriting authentication by envelopes of sound signatures", International Conference on Pattern Recognition ICPR'04, pp. 401-404, 2004.

[6] Miguel Gustavo Lizárraga Espinosa, Um Sistema Biométrico de Identificação Pessoal via Internet com ênfase em Assinaturas Estáticas, Tese de Doutorado, Unicamp, Agosto, 2000.

[7] Rohlík O., Matoušek V., Mautner P. and Kempf J., "A New Approach to Signature Verification Digital − Data Acquisition Pen", Neural Network World, vol. 11-5, 2001, pp. 493-501.

[8] P. Mautner, O. Rohlik, V. Matousek and J. Kempf, "Fast Signature Verification without a Special Tablet", Proceedings of IWSSIP02, World Scientific, Manchester, Nov. 2002, pp. 496-500.

[9] Soule M., Kempf J., "Handwritten Text Analysis through Sounds–A new device for handwriting analysis", Proceedings of IWSSIP03, Prague, Nov., 2003, pp. 254-257.

[10] Holanda, Aurélio B., "Dicionário Aurélio Eletrônico", Versão 3.0, 1999.

[11] Jain A., Ross A., Prabhakar S.,"An Introduction to Biometric Recognition", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 14, No. 1, January 2004.

[12] Jain A., Duin, R., Mao, J., "Statistical pattern recognition: a review", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 22, No 1, pp. 4 - 37, 2000.

[13] Duda R., Hart P. Pattern classification and scene analysis. Wiley, New York, 1973.

[14] Gonzalez, R.C., Woods R.E., Digital Image Processing. Addison-Wesley Publising Company, Inc., 1992.

[15] Abramson N., Braverman D., "Learning to Recognize Patterns in a Random Enviroment", IRE Trans. Inform. Theory, IT-8, No 5, pp. S58-S63, 1962.

[16] Duda R., Hart P., Strok D., Pattern Classification, 2nd edition, Wiley Interscience, New York, 2000.

[17] James M., Pattern Recognition, British Library, 1987.

[18] Mantas J., "Methodologies in pattern recognition and image analysis - A brief survey", Pattern Recognition, Vol. 20, No 1, pp. 1 - 6, 1987.

[19] Lee , L., Lizárraga M., "Classificação de Padrões e Extração de Parâmetros Discriminantes Utilizando Redes Neurais," Anais do II Congresso Brasileiro de Redes Neurais, Curitiba, novembro, 1995.

[20] Plamondon R., Lorette, G., "Automatic signature verification and writer identification - the state of the art", Pattern Recognition, Vol. 22, No 2, pp. 107 - 131, 1989.

[21] Eden M., "Handwriting and Pattern Recognition", IRE trans. on Inf. Theory, IT-8, February, 1962.

[22] Hilton O., "Signatures - Review and New View", Journal of Forensic Sciences, Vol. 37, No 1, pp. 125 - 129, 1992.

[23] Lindgren N., "Machine recognition of human language - Part III - Cursive script recognition", IEEE Spectrum, may, 1965.

[24] Edson José Rodrigues Justino, O Grafismo e os Modelos Escondidos de Markov na Verificaçao Automática de Assinaturas, Tese de Doutorado, Pontifícia Universidade Católica do Paraná, Curitiba, 2001.

[25] Mizukami Y., Yoshimura M., Hidetoshi M., Yoshimura I. "An off-line signature verification system using an extracted displacement function", Elsevier, Pattern Recognition Letters No 23, pp. 1569 - 1577, 2002.

[26] Lee L., Berger T., "Reliable On-line Human Signature Verification System for Point-of-Sales Applications", In Proceeding of the International Congerence on Pattern Recognition, pp. 21-23, 1994.

[27] Rohlík O., Handwritten Text Analisis, Laboratory of Intelligent Communication Systems, Department of Computer Science and Engineering, Faculty of Applied Science, University of West Bohemia in Pilsen, Pilsen, 2003.

[28] Crane H., Ostrem J., Automatic Signature Verification using a Threeaxis ForceSensitive Pen, IEEE Transactions on Systems, Man and Cybernetics, Vol. 13, No. 3, 1983.

[29] Fierrez J., Nanni L., Lopez J., Ortega J., Maltoni D., An On-Line Signature Verification System Based on Fusion of Local and Global InformationProc. of International Workshop on Biometric Recognition Systems (IWBRS), pp. 188-196, Beijing, China, October 2005.

[30] Mingfu Zou, Jianjun Tong, Changping Liu, Zhengliang Lou, "On-line Signature Verification Using Local Shape Analysis", Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR'03), IEEE, 2003.

[31] Vishvjit Nalwa, "Automatic On-Line Signature Verification",Proceedings of the IEEE, vol. 85, No. 2, febrary 1997.

[32] Martens R., Claesen L., "On-Line Signature Verification by Dynamic Time-Warping", Proceedings of Int. Conf. Pattern Recognition, pages 38–42, 1996.

[33] Fierrez J., Ortega J., Martin J., Gonzalez J., "A Novel Function-Based On-Line Signature Verification System Exploiting Statistical Signal Modeling"

[34] Lamel F., Rabiner L.,Rosemberg A., Wilpon J.,"An improved endpoint detector for isolated word recognition". IEEE Transactions on Acoustic, Speech and Signal Processing, 29 (4): 777-785, 1981.

[35] Wu D., Tanaka M. Chen R., Olorenshaw L., Amador M., Mendez Pidal X., "A Robust Speech Detection algorithm for Speech Activated Hands-Free Applications", IEEE, ITT 10.7, vol. 4, pp. 2407-2410, 1999.

[36] Zhang Y., Zhu X., Hoa Y., Luo Y., " A Robust and Fast Endpoint Detection Algorithm for Isolated word Recognition", International Conference on Intelligent Processing Systems", IEEE, China, 1997.

[37] Li Q., Zheng J., Tsai A., Zhou Q., "Robust Endpoint Detection and Energy Normalization for Real-Time Speech and Speaker Recognition", IEEE Transactions on Speech and Audio Processing", Vol. 10, NO. 3, March, 2002.

[38] Savoji M, "A robust algorithm for accurate endpointing of speech signals". Speech Communication, vol 8, pp. 45-60, 1989.

[39] Schlaikjer M., Bagge J., Sorensen O., Jensen J., "Trade Off Study on Different Envelope Detectors for B-mode Imaging", IEEE International Ultrasonic Symposium, Hawaii, 2003.

[40] Rowe H., Signals and Noise in Communication Systems, D.Van Nostrand Company, 1965.

[41] Ma L., Tan T., Wang Y., Zhang D., "Efficient Iris Recognition by Caracterizing Key Local Variations", IEEE Transactions on Image Processing, Vol. 13, No. 6, June, 2004.

[42] Boles W., Boashash B., "A Human Identification Technique Using Images of the Iris and Wavelet Transform", IEEE Transactions on Signal Processing, Vol. 46, No. 4, April, 1998.

[43] Lívia Cristina Freire Araújo, "Uma Metodologia para Autenticação Pessoal baseada em Dinâmica da Digitação", Tese de Mestrado, Unicamp, Fevereiro, 2004.

Apêndice A

Apêndice - Artigo Elaborado

Larco J., Yabu-uti J., "Un Método de Verificación Personal utilizando Emisión Acústica", ANDESCON2006, Congreso del Area Andina IEEE, noviembre, 2006, submetido para revisão.

Un Método de Verificación Personal utilizando Emisión Acústica

Julio Larco y João Baptista T. Yabu-uti

Departamento de Comunicações - Faculdade de Engenharia Elétrica e de Computação
Universidade Estadual de Campinas - UNICAMP
Campinas - SP, Brasil
Email: jlarco@decom.fee.unicamp.br
yabuuti@decom.fee.unicamp.br

Resumen— El estilo de firmar de un individuo, es una entidad biométrica que puede ser usada para diferenciar una persona de otra.

En este trabajo presentamos una metodología para realizar autenticación personal, utilizando el sonido que se produce en el momento de firmar. La fricción entre la punta rígida de un bolígrafo y el papel produce un sonido que puede ser usado para verificar la identidad de una persona.

Este método, está basado en el hecho que el sonido producido al firmar está correlacionado con la dinámica y la postura del escritor. Diferentes individuos producen diferentes trazos, los cuales resultan en diferentes señales de sonido.

Cada uno de los trazos que forman parte de la firma son caracterizados como variaciones agudas en los valores de la envolvente de la señal de sonido. Estas variaciones serán representadas como vectores binarios de características que son enviados para una etapa de reconocimiento de patrones, la cual decidirá si el sonido capturado proviene de una firma que fue realizada por un usuario legítimo o un impostor.

El método presentado en este trabajo es evaluado utilizando un conjunto de muestras de prueba y entrenamiento, pertenecientes a dos tipos de usuarios: legítimo e impostor habilidoso. El usuario legítimo es el propietario de la firma y el impostor habilidoso conoce la manera como el usuario legítimo firma. Como parámetros de desempeño de este método, se obtuvo los valores de 8.55% y 8.73% de tasas de error de falsa aceptación (FAR) y falso rechazo (FRR), respectivamente.

Índice de Términos—Verificación de firmas, reconocimiento de patrones, biometría.

I. Introduccion

La verificación de firmas es un método de autenticación personal basado en una biometría comportamental que analiza la manera como un usuario firma. Al firmar, características tales como la velocidad y la presión ejercida por la mano que prende el bolígrafo son tan importantes cuanto la forma gráfica de una firma [1]. Firmar es una acción reflejo, no influenciada por el control muscular deliberado, con determinadas características como ritmo, toques sucesivos en la superficie de escritura, velocidad y aceleración [2].

Al escribir, el movimiento de la punta del bolígrafo sobre la superficie del papel genera sonidos audibles, que se propagan en el interior del materia que está bajo el papel sobre el cual se está firmando. Estos sonidos pueden ser capturados por un sensor y están correlacionados con la dinámica y la postura del escritor, por lo tanto contienen información útil para distinguir escritores diferentes [3]. Esta forma de realizar la verificación de firmas es conocida como Emisión Acústica (EA - *Acoustic Emission*) [1].

El interés por usar el sonido de la escritura para la verificación y/o autenticación de escritor, se debe a que capturar sonidos con un micrófono es más fácil que medir la dinámica de escritura usando una mesa digitalizadora. La presión, la velocidad y la aceleración de los movimientos del bolígrafo son transformadas en una única señal en el tiempo que es la señal del sonido.

En este trabajo proponemos utilizar características biométricas para realizar identificación personal utilizando verificación de firmas a través del sonido que produce un bolígrafo en el momento que una persona firma. El sistema propuesto utiliza una superficie lisa y sólida sobre la cual se coloca un papel y en la que se encuentra ubicado un micrófono para la adquisición del sonido. Este procedimiento nos lleva a un sistema de adquisición más simple y más económico que cualquier sistema que adquiera imágenes para el reconocimiento.

Es importante notar que el sistema propuesto, permite obtener la información de la firma sin cualquier registro de ella en forma visual. Esto se puede hacer firmando en la superficie citada sin la presencia del papel, lo que constituye una característica más de seguridad.

II. El Método Propuesto

En este método cada usuario realiza un proceso de registro, donde se almacena información que permitirá identificar sus muestras. El diagrama de la figura 1, presenta las etapas de um sistema de verificación personal basado en el sonido de la firma.

Esta señal de sonido es filtrada y son eliminadas las muestras que no son parte de la firma. A continuación, se obtiene la envolvente de esta señal y se extraen los vectores de características que serán enviados a la etapa decisión donde se decide si la señal de sonido pertenece a una firma verdadera o a una firma falsa.

A. Adquisición de Firmas

El sonido resultante de la fricción de un bolígrafo con el papel, depende del tipo de bolígrafo y papel utilizados como también de la posición del escritor con respecto a la ubicación del micrófono. Para minimizar tales dependencias, fue desarrollado un dispositivo que ayude en esta tarea. Tomando en cuenta que para capturar todas las muestras se utilizo el mismo tipo de bolígrafo, papel y se mantuvo la misma posición de la firma con relación al micrófono [3].

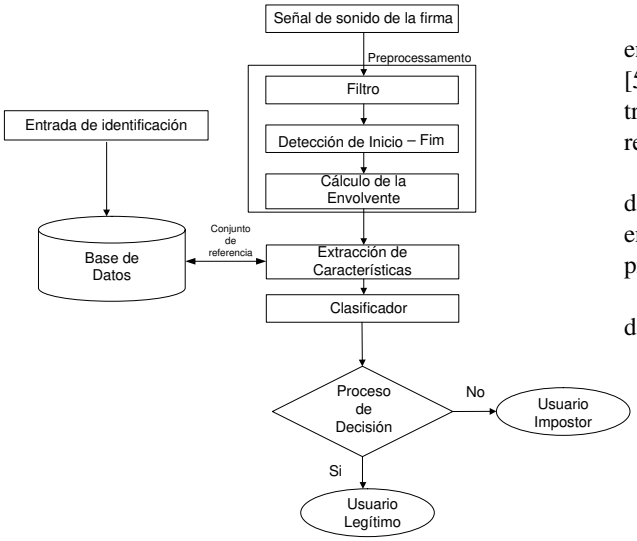


Fig. 1. Diagrama de bloques del método propuesto.

La figura 2 muestra un diagrama del dispositivo utilizado para la captura, que esta formado por una placa de madera y un micrófono. Este micrófono fue aislado para que capture apenas el sonido producido por el bolígrafo sobre el papel. El aislamiento minimiza los sonidos exteriores. La señal capturada ingresa a un computador para ser digitalizada. Esta presenta una energía cuyas componentes de frecuencia varían entre 200 Hz e 5000 Hz, por lo tanto, una frecuencia de muestreo de 16KHz se hace necesaria. La señal capturada ingresa al computador donde es digitalizada y almacenada.

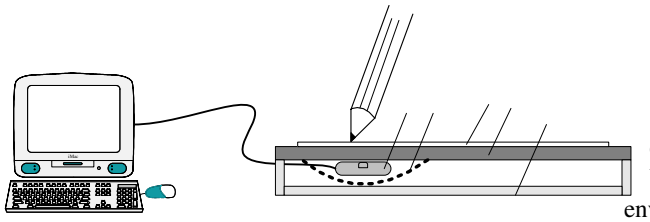


Fig. 2. Estructura interna del dispositivo de captura. 1)papel 2) tabla 3)micrófono 4)aislamiento 5)cobertura exterior 6)bolígrafo

B. Preprocesamiento.

Filtrado.

El micrófono del dispositivo de captura esta aislado para evitar que el ruido externo sea capturado, pero este aislamiento no es total. Cada vez que una muestra es capturada se produce un ruido proveniente de la fricción de la mano con la superficie del dispositivo, este ruido presenta frecuencias menores que 300Hz y por lo que puede ser filtrado.

Detección de los puntos de inicio-fin.

En lapso de tiempo antes y después de la firma, son capturadas muestras del ruido de fondo. Estas muestras no traen ninguna información relevante para el sistema de verificación, por lo que pueden ser eliminadas usando algoritmos de detección de los puntos de inicio-fin, los cuales también son conocidos como *endpoint detection.*

Estos algoritmos generalmente emplean parámetros de energía, estimación espectral, restricciones temporales [4] [5] [6]; cruzamiento por cero o por nivel [7]. Todos estos trabajos están orientados a trabajos usados en reconocimiento de voz.

En este trabajo, fueron probados los métodos de detección de puntos de inicio-fin utilizando las técnicas de energía y de cruzamiento por cero, siendo que el primero presento mejores resultados.

En la figura 2 y 3, se puede observar una firma y su señal de sonido con los puntos de inicio-fin.

Fig. 2. Firma de un usuario

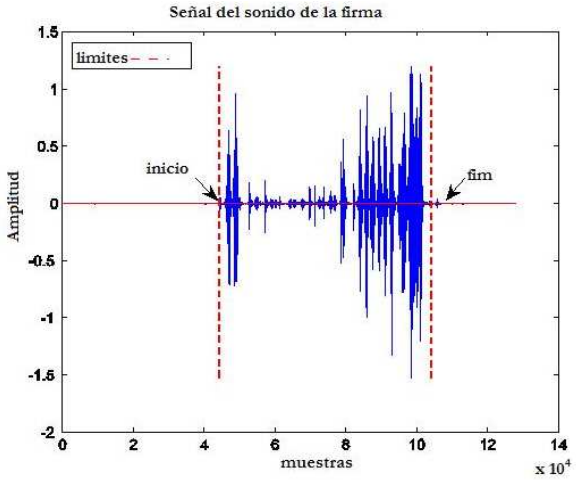


Fig. 3. Señal de sonido de la firma.

Cálculo de la envolvente.

Una técnica común y muy eficiente para a detección de la envolvente esta basada en la transformada de Hilbert [8].

En este trabajo, se utiliza la envolvente normalizada del sonido de la firma como el espacio de característica, donde los detalles de alta frecuencia presentes pueden ser vistos como señales de portadora. Diferentes papeles y bolígrafos presentan contenidos de frecuencia diferentes, esto es, portadoras diferentes. El movimiento del bolígrafo en la superficie del papel impone un efecto de modulación en estas portadoras. Se espera que el algoritmo que calcula el vector de características extraiga las características de la dinámica del movimiento de la mano, sin grande influencia del tipo de papel y bolígrafo usados. Entonces las envolventes de los sonidos de la firma son consideradas como el espacio de característica del cual se puede extraer un vector de características conveniente al reconocimiento de escritor [3][4].

La envolvente del sonido digitalizado es obtenida calculando la transformada discreta de Hilbert,

$$env[n] = |s[n] + jH_d\{s[n]\}| \qquad (1)$$

donde *env[n]* es la envolvente de la señal del sonido (*s[n]*) y H_d es la transformada discreta de Hilbert de *s[n]* [4].

La envolvente esta formada de picos y valles que están relacionados con las características de la firma. Estos picos y valles de *env* representan los trazos que componen la firma y pueden ser utilizados como una forma de representación. Para obtener la forma de la envolvente, se puede decimar *env* para reducir el número de muestras sin perder información a cerca da forma de la envolvente. En esta etapa del preprocesamiento la decimación utilizada es de 500 muestras/seg.

Se pueden encontrar envolventes de la firma de un mismo escritor con diferentes amplitudes y número de muestras. Esto justifica un proceso de normalización que permitirá tener el mismo número de muestras en cualquier firma. Para eso, se hace un submuestreo de *env[n]* tomando siempre *p* muestras igualmente espaciadas. (ver ecuación 2) [4],

$$Env[n] = \frac{env[n]}{\sqrt{\sum_{n=1}^{p} env^2[n]}} \qquad (2)$$

donde, *Env* es la envolvente normalizada, *p* es el número de muestras igualmente espaciadas obtenidas de *env[n]*. En este trabajo, se hicieron pruebas con diferentes valores de p. En la figura 4, es mostrada la envolvente normalizada de la firma presentada en la figura 2.

En la figura 4, son presentadas dos envolventes normalizadas de una misma firma hechas por el mismo escritor y en la figura 5 las envolventes de una misma firma hechas por un usuario legitimo e un impostor. Se nota que firmas hechas por la misma persona tienen pequeñas variaciones en su envolvente, pero firmas de diferentes personas tienen grandes diferencias.

*C. **Extracción de Características.***

Los picos que son parte de la envolvente representan cada letra, palabra, línea, rasgo o trazo producidos por el movimiento del bolígrafo. Cada firma presenta una forma de envolvente dependiendo de la velocidad, aceleración y presión del escritor.

Para obtener un vector de características que represente de mejor forma la envolvente, en este método, se utilizan tres límites lb_1, lb_2 y lb_3, que permiten crear un vector binario cuyos valores van a depender de la amplitud y la posición de los picos de la envolvente (ver ecuaciones 3,4,5). Los valores utilizados para las constantes son: b_1=0,5 y b_2=2 que fueron determinados empíricamente.

$$lb_1 = \frac{1}{n_e} \sum_{i=1}^{n_{ee}} Env[i] \qquad (3)$$

$$lb_2 = b_1 \cdot lb_1 \qquad (4)$$

$$lb_3 = b_2 \cdot lb_1 \qquad (5)$$

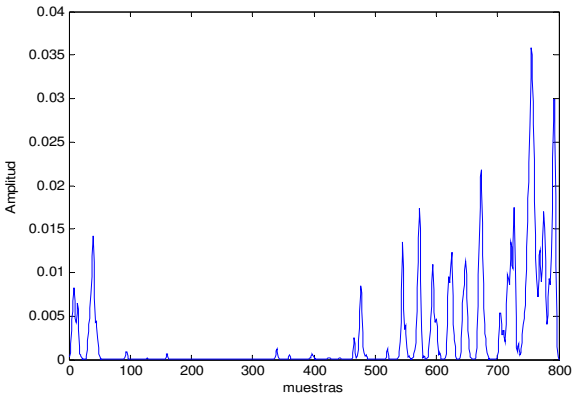


Fig. 3. Envolvente de la señal de sonido, normalizada con *p*=800 muestras.

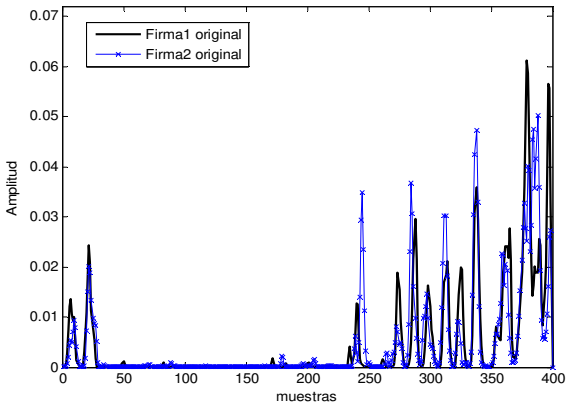


Fig. 4. Envolventes legitimas de un usuario, con *p*=400 muestras.

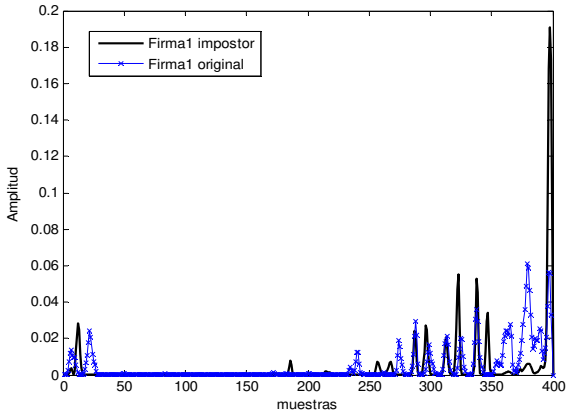


Fig. 5. Envolventes de un usuario legitimo y un impostor, con *p*=400 muestras.

Utilizando estos límites se calculan tres vectores auxiliares Vcb_1, Vcb_2 y Vcb_3, generados de la siguiente manera: Los puntos de la envolvente que sean mayores en amplitud que un limite lb_j son representados con un valor de 1 y los puntos cuya amplitud sea menor que lb_j son representados como 0 [9].

$$Vcb_j = \begin{cases} 1 & si \quad Env[i] \geq lb_j \quad i = 1,...,n_e \\ 0 & si \quad Env[i] < lb_j \quad j = 1,2,3. \end{cases} \qquad (6)$$

A unión de estos vectores forma el vector de características de la señal del sonido de la firma, denominado *Vcb*.

En la figura 8, se puede observar una parte de la envolvente y su vector de características Vcb_1 que codifica los picos da envolvente como unos o ceros.

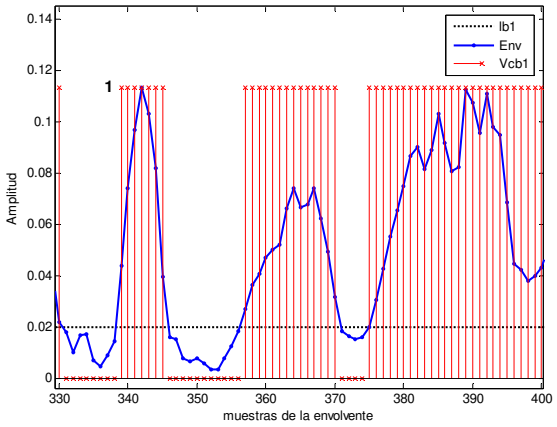


Fig. 6. Vector binario de características calculado utilizando lb_1.

D. ***Comparación.***

Dos firmas pertenecen al mismo escritor si existe una semejanza entre sus vectores de características. Esta semejanza puede ser expresada como una distancia entre estos vectores. Un procedimiento de dos etapas es empleado para realizar esta comparación [10]. Las etapas son:

1. Un vector de características es extraído de la señal original y es transformado en un vector binario.
2. La semejanza entre dos vectores de características binarios es calculada usando una operación lógica OR exclusivo (XOR).

Para que esta medida de semejanza sea representativa es importante que los vectores de características sean invariantes a la translación, escala y rotación. Dado que los vectores binarios son normalizados tanto en amplitud como en número de muestras, estos son invariantes con relación a la escala. Estos vectores también son invariantes a la rotación, ya que la señal capturada es el sonido. Por otro lado, la medida de semejanza entre dos vectores es afectada por la translación.

Aunque dos vectores sean semejantes, si el momento de compararlos, uno de ellos esta desplazado en relación al otro, el resultado será un valor mucho mayor del que cuando los dos vectores están alineados. Eliminar las muestras del ruido de fundo usando los puntos de inicio-fin minimiza el efecto de traslación. A pesar de eso, siempre existe una pequeña translación. Para solucionar esto, se dividí el vector de características e bloques de *K* muestras. A continuación, se calcula la distancia de semejanza de pequeños bloques del vector de características, denominados *vc*. La suma total de estos valores es la distancia de semejanza *D*. (ver ecuación 7) [10],

$$D = \frac{1}{L}\sum_{j=1}^{N} vc_j^1 \oplus vc_j^2 \qquad (7)$$

donde, vc_j^1 e vc_j^2 son dos vectores de características distintos, *L* es el tamaño de estos vectores y $\oplus$ es el operador XOR.

Para comparar dos bloques vc_1 y vc_2 primero desplazamos el bloque vc_2 varias muestras tanto a la derecha como a izquierda, la menor de las distancias de semejanza es considerada como el valor de distancia.

Las distancias obtenidas por la comparación de los vectores son usadas para definir si una señal pertenece a una firma hecha por un usuario legitimo o un impostor. Para ello, se debe definir un valor límite de la distancia de semejanza, que establezca la frontera de decisión entre una firma verdadera e una falsa. Para alcanzar este objetivo, se sigue los siguientes pasos:

- Se calcula la distancia entre los vectores de características de firmas verdaderas (*Dv*), y las distancia de cada vector de características de las firmas verdaderas con los vector de las firmas del mismo escritor hechas por un impostor (*Di*). utilizando la ecuación 7.
- Se obtiene el diagrama de distribución de las distancias *Dv* y *Di* (ver figura 7). A partir del histograma son calculadas la media *m* y la varianza σ de las distancias *Dv, Di* y se ajustados sus valores a una distribución gaussiana.
  Se calcula el valor del limite *Td* que divide los valores que pueden ser considerados como a distancias de semejanza *Dv* y *Di*.
- Para analizar una firma de prueba, se calcula las distancias de esta con respecto a las firmas verdaderas (*Dt*). La media aritmética de *Dt* es calculada e comparada con o valor de *Td*. Si es menor, la firma pertenece a un usuario legítimo pero si es mayor, la firma fue hecha por impostor.

III. EXPERIMENTOS

Para este trabajo, se utilizaron 55 personas, de las que se capturaron 11 muestras de cada uno, constituyendo un total de 605 sonidos de firmas verdaderas. Además se capturaron 11 falsificaciones por usuario verdadero dando un total de 605 falsificaciones habilidosas.

Una señal de prueba enviada a la etapa de comparación, puede pertenecer a un usuario verdadero o a un impostor. El clasificador puede clasificar esta señal correctamente, pero también puede ocurrir que el clasificador cometa un error de decisión y se reconozca como usuario legítimo a un impostor o viceversa.

En estos casos, el desempeño del sistema es medido por la tasa de error de falso rechazo (FRR) y tasa de error de falsa aceptación (FAR) [1].

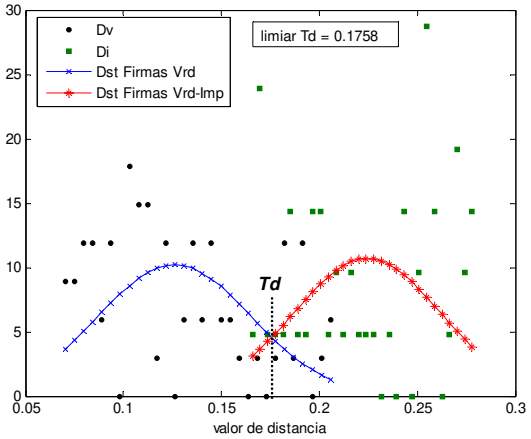


Fig. 7. Histograma de los valores das distancias *Dv* y *Di*.

Varios experimentos son realizados para verificar el impacto en las tasas FAR y FRR del número de muestras da envolvente (*N*) y el número de bloques (*K*) por envolvente. Los resultados se pueden observar en la tabla 1.

Analizando los resultados, se nota que as tasas FAR y FRR son fuertemente afectados por el número de bloques *K*, mientras que estas tasas sufren menos influencia por el número de muestras da envolvente *N*. El mejor compromiso entre las dos tasas ocurre para *N*=200, *K*=10 donde FAR=8,55% y FRR=8,73%.

Número de muestras de la Envolvente (N)	Número de bloques (K) por envolvente	FAR(%)	FRR(%)
200	5	31,45	5,45
	8	13,09	6,73
	10	8,55	8,73
	12	5,82	12,55
	15	3,09	18,36
	20	1,64	28,36
400	5	29,82	6,18
	8	12,18	8,00
	10	8,55	10,36
	12	5,09	12,55
	15	2,73	18,55
	20	1,27	25,82
800	5	25,64	6,18
	8	12,36	7,82
	10	7,82	9,64
	12	4,55	12,18
	15	2,55	18,18
	20	1,09	27,64
1000	5	29,90	5,46
	8	13,46	7,64
	10	8,55	10,00
	12	4,18	12,00
	15	2,18	18,73
	20	1,27	29,46

Tabla 1

IV. CONCLUSIONES

En este artículo se presenta un método de verificación de la identidad de una persona vía reconocimiento de firmas. Este método consiste en obtener y procesar medidas biométricas extraídas del sonido que se produce en el momento que una persona firma.

Este trabajo, se caracteriza por emplear características locales extraídas de la envolvente de esta señal de sonido, representadas por un vector binario de características, los cuales son procesados para obtener una distancia de semejanza a través de una operación XOR.

Como parámetros de desempeño de este método de verificación frente a falsificaciones habilidosas, se obtuvo los valores de 8.55% e 8.73% para tasas de error FAR e FRR, respectivamente.

V. REFERENCIAS

1. Mansfield T., Roethenbaugh T., "Glossary of Biometric Terms", International Association for Biometrics (iAfB), 1999, *http://www.iafb.org.uk/DOC/GlossaryTerms.htm*., consultada en Oct/05.

2. Polemi D., "Biometric Techniques: Review and Evaluation of Biometric Techniques for Identification and Authentication, Including an Appraisal of the Areas Where They are Most Applicable", Final Report, Institute of Communication and Computer Systems National Technical University of Athens, 1997. *ftp://ftp.cordis.lu/pub/infosec/docs/biomet.doc*. consultado en junio de 2005.

3. Soule M., Kempf J., "Handwritten Text Analysis through Sounds--A new device for handwriting analysis", Proceedings of IWSSIP\'03, Prague, Nov., 2003, pp. 254-257.

4. Li F. F., "Handwriting authentication by envelopes of sound signatures", International Conference on Pattern Recognition ICPR'04, pp. 401-404, 2004.

5. Lamel F., Rabiner L.,Rosemberg A., Wilpon J.,``An improved endpoint detector for isolated word recognition". IEEE Transactions on Acoustic, Speech and Signal Processing, 29 (4): 777-785, 1981.

6. Wu D., Tanaka M. Chen R., Olorenshaw L., Amador M., Mendez Pidal X., ``A Robust Speech Detection algorithm for Speech Activated Hands-Free Applications", IEEE, ITT 10.7, vol. 4, pp. 2407-2410, 1999.

7. Zhang Y., Zhu X., Hoa Y., Luo Y., "A Robust and Fast Endpoint Detection Algorithm for Isolated word Recognition", International Conference on Intelligent Processing Systems", IEEE, China, 1997.

8. Schlaikjer M, Bagge J., Sorensen O., Jensen J., ``Trade Off Study on Different Envelope Detectors for B-mode Imaging", IEEE International Ultrasonic Symposium, Hawaii, 2003.

9. Boles W., Boashash B., ``A Human Identification Technique Using Images of the Iris and Wavelet Transform", IEEE Transactions on Signal Processing, Vol. 46, No. 4, April, 1998.

10. Ma L., Tan T., Wang Y., Zhang D., ``Efficient Iris Recognition by Caracterizing Key Local Variations", IEEE Transactions on Image Processing, Vol. 13, No. 6, June, 2004.