

VERIFICAÇÃO OFF-LINE DE ASSINATURAS UTILIZANDO FILTROS ORIENTADOS E COMPARAÇÕES ELÁSTICAS

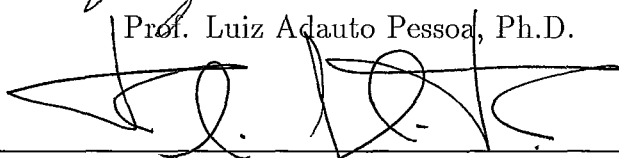
André Rodrigues Pereira

TESE SUBMETIDA AO CORPO DOCENTE DE COORDENAÇÃO DOS
PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE
EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

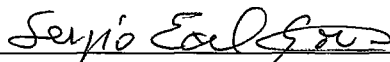
Aprovada por:



Prof. Luiz Adauto Pessoa, Ph.D.



Prof. Felipe Maia Galvão França, Ph.D.



Prof. Sergio Exel Gonçalves, D.Sc.



Prof. Jonas de Miranda Gomes, D.Sc.

RIO DE JANEIRO, RJ - BRASIL
AGOSTO DE 1999

PEREIRA, ANDRÉ RODRIGUES

Verificação Off-line de Assinaturas Utilizando Filtros Orientados e Comparações Elásticas. [Rio de Janeiro] 1999

IX, 131 pp., 29.7 cm, (COPPE/UFRJ, M. Sc., Engenharia de Sistemas e Computação, 1999)

Tese – Universidade Federal do Rio de Janeiro, COPPE

1 – Verificação de Assinaturas

2 – Filtros Orientados

3 – Comparações Elásticas

I. COPPE/UFRJ II. Título (série)

Resumo da tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M. Sc.)

VERIFICAÇÃO OFF-LINE DE ASSINATURAS UTILIZANDO FILTROS ORIENTADOS E COMPARAÇÕES ELÁSTICAS.

André Rodrigues Pereira

Agosto/1999

Orientador: Luiz Aduino F. C. Pessoa

Programa: Engenharia de Sistemas e Computação

Neste trabalho estamos propondo um estudo sobre a aplicação de filtros orientados na verificação de assinaturas manuscritas. Duas formas de comparações foram implementadas, a primeira utilizando uma comparação global e a outra com gráficos elásticos. A aplicação destes processamentos, revela ser um caminho promissor para a verificação encontrando-se erros relativamente baixos, mas um maior número de testes deve ser realizado em bases de dados maiores.

Abstract of Thesis presented to COPPE/UFRJ as partial fulfillment of the requirements for the degree of Masters of Science (M.Sc.)

HANDWRITTEN OFF-LINE SIGNATURE VERIFICATION USING
ORIENTED FILTERS AND ELASTIC MATCHING

André Rodrigues Pereira

August/1999

Advisor: Luiz Adauto F. C. Pessoa

Department: Computing and Systems Engineering

In this work we propose a new approach about applicability of oriented filters for handwritten signature verification. Two methods of match have been implemented, the first uses a global match and the other utility elastic graph. The approaches showed in this thesis seem promising, but further experimentation should be carried out to evaluate it on a more large signature database.

AGRADECIMENTOS

Como sempre acontece, há pessoas a quem agradecer.

Aos meus familiares por todo o apoio;

Aos professores por tudo que me ensinaram, (os erros que porventura cometo são meus, não deles);

Aos meus caros colegas de mestrado e doutorado, por suas percepções, e assinaturas, é claro;

Ao Luiz por sua análise e auxílio;

Ao pessoal de Terê, por sua amizade;

E a muitos outros que ouviram minhas intermináveis perguntas e idéias - até as idiotas;

E a todos os homens de boa vontade que esperam, como eu, um mundo melhor

...

*Não te dei face,
nem lugar que te seja próprio,
nem dom algum que te faça particular,
ó Adão,
a fim de que tua face,
teu lugar e teus dons,
tu os desveles,
conquistes e possuas por ti mesmo.
A Natureza definida de outras espécies
foram estabelecidas, em leis, por mim.
Mas tu, a quem nenhum confim delimita,
por teu próprio arbítrio,
entre as mãos daquele que te colocou,
tu te defines a ti mesmo.
Te pus no mundo,
a fim de que possas
melhor contemplar o que contém o mundo.
Não te fiz celeste nem terrestre,
mortal ou imortal,
a fim de que tu mesmo,
livremente,
à maneira de um bom pintor
ou de um hábil escultor,
descubras tua própria forma . . .
Picco della Mirandola,
Oratio de homminis dignitate;*

*extraído do livro:
“A Obra em Negro”
de Marguerite Yourcenar,
tradução de Ivan Junqueira*

Índice

1	Introdução	1
1.1	O Problema	1
1.2	Uma Visão Geral da Área de Manuscritos	4
1.3	Uma Proposta	8
2	Histórico de Reconhecimento de Assinaturas	11
2.1	Alguns Aspectos Básicos	11
2.1.1	Dificuldades	12
2.2	Modelando o Sistema	16
2.2.1	Saídas de Aceitação (χ_1^i)	22
2.2.2	Saídas de Rejeição (χ_2^i)	22
2.2.3	Relações das Saídas	23
2.3	Desempenho Sobre Falsificações	24
2.4	Desenvolvendo os Passos	28
2.4.1	Aquisição de Dados	30
2.4.2	Pré-processamento	30
2.4.3	Extração de Características	32
2.4.4	Comparação	36
2.4.5	Avaliação de Desempenho	37
2.5	Alguns Aspectos Finais dos Processos	38
3	Sistemas	40
3.1	Técnicas Estáticas (Off-line)	41
3.1.1	Características Globais	41

3.1.2	Características Estatísticas	43
3.1.3	Características Geométricas	45
3.2	Técnicas Dinâmicas (On-line)	48
4	Filtros	50
4.1	Modelo de Imagem	52
4.2	Filtros Lineares	54
4.2.1	Convolução Digital e Máscara	55
4.3	Modelos de Filtros	56
4.3.1	Filtros Detectores de Bordas	57
4.3.2	Filtros e a Visão Humana	59
4.3.3	Formalização	63
4.3.4	Jets	66
5	Comparações Elásticas	69
5.1	Uma Revisão de Comparações Elásticas	71
5.1.1	EGM	74
5.2	Estrutura das Comparações Elásticas	75
5.2.1	Funções Globais e Locais	78
5.3	Formalização do EGM	80
5.3.1	Primitiva: Jet	82
5.3.2	Grafo do Modelo	82
5.3.3	Movimento Global	88
5.3.4	Movimento Local	89
5.3.5	Considerações de Complexidade	91
5.3.6	Comparação de Retorno	92
6	Resultados e Discussões	95
6.1	Bases de Dados	95
6.1.1	Primeira Base	95
6.1.2	Segunda Base	96
6.2	Pré-processamento	97

6.2.1	Representação Canônica	97
6.3	Fórmulas e Limiar de Decisão:	100
6.4	Comparações Globais	105
6.4.1	Comparação AND	108
6.4.2	Comparação OU	112
6.4.3	Comparação TEMP	113
6.4.4	Comparação	114
6.5	Comparação com EGM	115
6.5.1	EGM com Peso	117
6.5.2	Global X EGM	118
7	Conclusões e Trabalhos Futuros	120
7.1	Trabalhos Futuros	123

Capítulo 1

Introdução

*“Quando se consegue ler algo sem esforço,
foi necessário muito esforço para o escrever.”*

Enrique J. Porcela

1.1 O Problema

O grande avanço tecnológico dos últimos anos na área de computação possibilitou que as máquinas fossem cada vez mais utilizadas na substituição de tarefas que antes eram realizadas por seres humanos ou em novos processos que surgiram com esta evolução tendo como objetivo usar máquinas para trabalhar mais rapidamente, melhor e com menos esforço. Neste sentido, criou-se um grande interesse na utilização dos computadores em tarefas de autenticação mediante verificação de assinaturas manuscritas, devido ao fato do método de autenticar documentos com uma firma ser um aspecto comum e largamente utilizado no campo financeiro, pela sua aceitação e confiança. Uma outra possibilidade aventurada é a utilização nos casos onde necessita-se identificar uma pessoa por questões de segurança, como permitir o acesso desta a um setor de segurança. Aqui a assinatura pode ser superior a várias outras técnicas de identificação biométricas (como verificação da retina), que, embora sendo mais confiáveis, são mais intrusivas.

A maioria de nós escolhe, em determinado momento, um estilo de assinatura e passa a segui-lo na hora de assinar documentos, embora seja um fato notório que duas assinaturas genuínas de uma pessoa não são precisamente iguais. Ou melhor, não há duas assinaturas exatamente idênticas, nem mesmo feitas pela mesma pessoa, no mesmo papel, com a mesma caneta, no mesmo lugar, e praticamente à mesma hora. Especialistas indicam que se duas assinaturas de uma pessoa apresentadas num papel são idênticas, elas podem ser consideradas falsificações por traçado ¹. Apesar disto, todas as vezes que uma pessoa assina seu nome, permanecem características que identificam a sua assinatura específica, sendo estas quase sempre os detalhes artificiais² que o autor escolheu ao “criar” a assinatura, ainda que inconscientemente.

Geralmente, também, no processo de elaboração da assinatura estamos com o pensamento voltado para o tipo de imagem que queremos transmitir, o que é mais um dado artificial que influi em sua composição. Em relação à este ponto, costuma-se dizer que a simples assinatura de uma pessoa revela muito sobre a sua personalidade, possuindo em si um valor simbólico, uma vez que muitos poucos assinam seu nome com absoluta naturalidade; tomando-a em vez disso com uma espécie de logotipo pessoal, e portanto, sempre como um toque de artificialidade. A “imagem” pública se reveste de grande importância para pessoas que ocupam cargos de responsabilidade refletindo numa assinatura “singular”. Um exemplo que revela claramente a alta posição hierárquica de quem a possui, é a assinatura da Rainha Elizabeth I (Figura 1.1) onde pode-se notar vários aspectos grafológicos como floreios das letras e retoques bem elaborados.

Assim, assinaturas de diversas pessoas, apresentam grandes variações não só de formato, como também nas características que apresentam, de acordo com a forma

¹ Estas falsificações são as produzidas, por exemplo, pela sobreposição do papel sobre a verdadeira, onde o falsificador tenta copiar exatamente a assinatura.

² Descrevo como “artificiais” porque a pessoa, em geral, determina a forma de sua assinatura com características particulares, e não da mesma forma que escreve um texto.

From Hapsberigs scribled this 27th of October.
Dont lovinige hylar Elizabeth

Figura 1.1: Assinatura da Rainha Elizabeth I. Note os floreios e retoques bem elaborados, marcas da artificialidade de uma assinatura.

que cada uma escreve e da personalidade que deseja representar. Ainda que existam homônimos que escrevam seus nomes na hora de assinar, suas firmas serão completamente diferentes, embora representem o mesmo nome. As sucessivas assinaturas de uma mesma pessoa diferem também localmente (na disposição dos traços) e globalmente (escala e ângulo de escrita), além de possuírem uma alta diversidade de escala e orientação. Mesmo com este enorme universo de possibilidades, especialistas humanos são muito bons na identificação de falsificações. Automatizar este procedimento é o grande objetivo da Verificação Automática de Assinaturas [1].

Convém notar a importância da Verificação de Assinaturas devido a sua presença em várias situações da nossa vida; por exemplo, quando se verifica as assinaturas de cartão de crédito. No caso bancário, ter-se-ia a vantagem de conseguir armazenar a informação da assinatura eletronicamente, permitindo uma busca mais rápida que a atual, assim como o acesso remoto de qualquer uma de suas agências conectadas ao banco de dados das assinaturas dos clientes, a qualquer momento; sendo que a verificação poderia ser feita da mesma forma digital e remota.

Um alto nível de segurança contra falsificações pode não ser necessário, e uma faixa de aceitação de 10% de assinaturas falsas já reduz as fraudes e pode ser um método melhor que o de verificação atual [2], estabelecendo um acesso mais rápido e liberando funcionários para outras obrigações. Mesmo um sistema de verificação

múltiplo, onde o processo automático só indicasse quais as assinaturas com problemas para serem posteriormente conferidas por especialistas humanos, já seria útil, na forma que reduziria o número total de assinaturas verificadas manualmente.

Como a segurança dos sistemas de informação está se tornando cada vez mais necessária, há também um grande interesse no desenvolvimento de técnicas de segurança não intrusivas e invasoras. Métodos baseados em assinaturas são uma boa opção em vários casos, em substituição às senhas. Eles vão apresentar uma dificuldade maior para o invasor do sistema, na medida que, em vez de tentar conjecturar (ou obter com alguma observação) uma senha, terá pela frente um método mais evoluído de segurança. Um ponto favorável sobre outros sistemas é ser um método já aceito pela maioria da população e não ser tão intimidador (como verificação da retina) para a grande maioria da população, sendo facilmente aceito pelas pessoas comuns.

1.2 Uma Visão Geral da Área de Manuscritos

Para que o leitor compreenda a necessidade da Verificação demonstramos a importância do reconhecimento automático de assinaturas na Seção anterior. Porém, é útil também examinar o campo de processamento de manuscritos para visualizar as várias áreas de automatização relacionada aos processos de escrita, com diferentes aplicações e requerendo diferentes estudos. Isto servirá não só para demonstrar a importância total da área, mas também entender a que o termo verificação se refere.

Uma das maiores aspirações no campo da Inteligência Artificial, sem esquecer-nos do objetivo primordial de analisar e emular a inteligência humana, é simplesmente permitir ao computador executar tarefas humanas intelectuais, ou seja realização, avaliação e interpretação de impressões sensoriais. Com isto, estas máquinas pode-

riam ser capazes de interagir com as pessoas e atuar no mundo de uma forma menos limitada. Este objetivo pode ser visto na indústria, quando esta produz computadores com uma interface “amigável”. Seguindo este pensamento, os computadores tem saído dos laboratórios e entrado em escritórios e casas; e a interface, antes feita com cartões e chaves, tornou-se mais simples, com teclados, mouses e ícones. No caminho desta evolução para encontrar interfaces mais humanas encontra-se a escrita, que é uma das mais naturais formas de comunicação entre as pessoas, sendo utilizada a milênios na vida de todos. Por isto prover uma interação melhor entre computadores e textos manuscritos, permitindo a manipulação por parte destes da informação textual, é um dos campos de estudo da atualidade.

Outras abordagens neste campo podem ser vistas no reconhecimento de código postal e endereços, compensações de cheques (não só verificando a assinatura, como conferindo valores expressos) e armazenagem de documentos escritos. Sendo que as automatizações executadas nestes exemplos foram aplicados por vários motivos como reduzir custos, aumentar a qualidade ou ampliar o volume de trabalho executado.

O problema de Reconhecimento de Padrões é geralmente apresentado como classificação ou discriminação do conjunto de processos ou eventos. O número de classes é determinado pela aplicação, podendo ou não, ser pré-definido. Para reconhecimento de caracteres em um texto precisamos de um número de classes igual ao número de letras do alfabeto, mas para descobrir se o mesmo está em inglês ou em português, só é necessário duas classes. Ou seja, precisamos saber qual o conhecimento que está sendo avaliado para iniciar um estudo sobre o mesmo.

No contexto de escrita, podemos dividir em duas castas: a parte *semântica*, que está relacionada com a mensagem do que está escrito, e a parte *singular*, correspondente às características do escritor. Assim, vários tipos de sistemas podem ser apresentados, sendo definidos pela informação a ser extraída do texto.

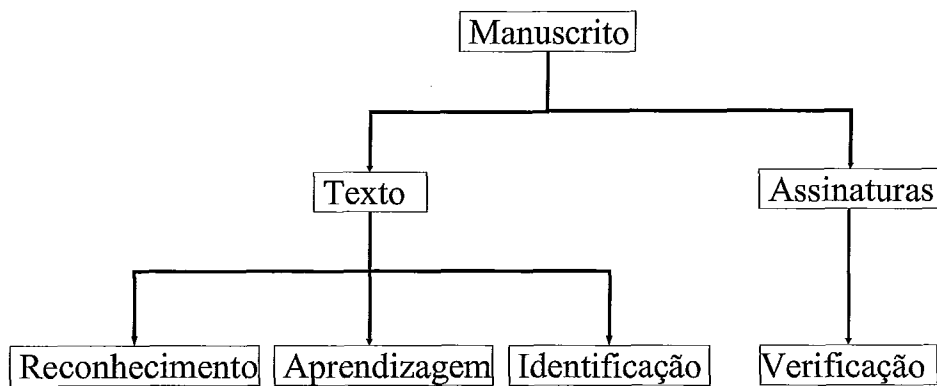


Figura 1.2: Apresentam-se as várias espécies de sistemas que trabalham com manuscritos.

Um sumário dos sistemas [3, 4] é ilustrado na Figura 1.2 para uma melhor visualização das definições acima. Eles podem ser definidos como:

- a *Aprendizagem*, onde busca-se a parte semântica, sem conhecermos a classe a qual o texto pertence, a priori;
- o *Reconhecimento*, onde também objetiva-se a parte semântica, mas aqui se conhece a classe de classificação desde o início, preocupando-se em descobrir se esta idéia é correta;
- a *Identificação*, onde deseja-se reconhecer as características particulares da escrita de cada indivíduo, embora se desconheça quem é o escritor;
- a *Verificação*, onde, partindo do princípio que o autor é conhecido inicialmente, precisa-se reconhecer particularidades da assinatura, a fim de certificarmos que a identificação é verdadeira.

A *Aprendizagem* e o *Reconhecimento* referem-se a informação codificada da mensagem a ser determinada, independente do escritor. Como ilustração, pode-se indicar o exemplo acima de descobrir se o texto está em português ou em inglês para a *Aprendizagem*. O *Reconhecimento* seria, partindo do princípio que o texto está

em inglês, certificar-se que isto é verdade. Note que ele difere da *Verificação* na medida que são utilizadas informações semânticas (o que as palavras significam) no método, sendo independente do autor do texto. Estes dois processos estão ligados à comparação de padrões, e nestes casos, um conjunto de protótipos (ou modelos) é armazenado na máquina. A entrada³ é comparada com cada classe, no caso da *Aprendizagem*, ou com um único modelo, no caso do *Reconhecimento*; a classificação é baseada num critério de similaridade. Em outras palavras, no primeiro caso, se a entrada se aproxima do modelo i com um valor de distância (limiar) mais baixo que todos os demais ela é aceita como sendo do escritor de i , e no segundo, se a entrada possui um valor de similaridade maior que um determinado nível, um limiar de decisão, é aceita como verdadeira.

A *Identificação* e *Verificação* descrevem um processo de informação relativos a buscas das propriedades mais “singulares” (individuais ou particulares) de um texto, para se determinar a identidade do escritor. Esta identidade pode ser descoberta pela comparação dos atributos específicos num conjunto de N escritores de uma base de referência R , no caso de *identificação*. Um sistema de verificação, por outro lado, decide se a identidade de um escritor é realmente i (confirma a hipótese) por processos de comparação com modelos armazenados de i . Ou, nas palavras de Miller [5] a *Verificação* requer que a escrita seja identificada ao reclamar uma identidade, tendo o sistema uma escolha binária de aceitar ou rejeitar a personalidade reclamada. A questão para máquina seria “Você é quem diz que é?”. Já a *Identificação* requer que o sistema busque entre vários conjuntos de características armazenados e selecione a melhor comparação de características de um indivíduo desconhecido presente, onde a questão é: “Eu sei quem é você?”.

³Com classificação desconhecida para a *Aprendizagem* e conhecida para o *Reconhecimento*.

1.3 Uma Proposta

Há um grande número de sistemas de verificação automática de assinaturas manuscritas (*VAM*). Como prova disto, existem mais de 100 patentes de sistemas na área [2]. Nesta tese buscou-se estudar a performance de um método de verificação estático a partir da aplicação de filtros orientados sobre a imagem da assinatura. Também foi verificado a utilização do algoritmo de Elastic Graph Matching (*EGM*) [6] sobre as imagens como instrumento da autenticação.

Um primeiro processo a ser executado é colocar a base de assinaturas num padrão que facilite o processo de verificação. Com este intuito, as imagens das assinaturas são pré-processadas eliminando “bordas” (partes da imagem que não contém traços da assinatura) a fim de que dispuséssemos só do espaço útil das assinaturas em si. Com isto, deste conjunto de assinaturas retirou-se um tamanho mínimo, para o qual todas as outras assinaturas foram dispostas, com o objetivo de que as mesmas sejam representadas num formato canônico, diminuindo as deformações globais.

Foram aplicados filtros sensíveis a contornos orientados, ou seja, que revelem o sentido dos traços da assinatura em uma região, à forma canônica da assinatura. Estes filtros, utilizados em reconhecimento de objetos e faces [7, 8], são motivados na psicofísica do sistema visual humano e tem como objetivo de processamento permitir que o computador perceba a distribuição dos mesmos e utilize esta informação na verificação. Com isto obtiveram-se oito imagens que são o resultado das oito orientações dos filtros. Estas, apresentam novas vistas da assinatura, onde cada uma enfatiza uma orientação de traço em particular. Uma nona imagem formada é o valor da soma das imagens anteriores, onde se vê em todas as bordas dos traços independente do sentido do mesmo.

Comparações globais foram utilizadas a fim de conseguirmos demonstrar o de-

sempenho das imagens pré-processadas com os filtros orientados. Como extensão, também verificamos se é mais interessante termos uma assinatura como modelo ou possuímos um conjunto com alguns modelos pode representar um ganho de informação. Na aplicação do *EGM*, temos uma verificação que permite comparações elásticas, onde pequenas distorções de assinaturas são aceitas. Este algoritmo permite algumas variações particulares que serão estudadas.

No Capítulo 2, apresenta-se uma introdução sobre vários aspectos da verificação de assinaturas, tendo como objetivo familiarizar o leitor com os termos da área, bem como introduzir conceitos necessários à compreensão dos métodos utilizados nesta tese, bem como a forma de avaliação dos resultados. Apresenta-se um formato “modelo” de como os sistemas são implementados, para fornecer uma idéia de como o sistema pode ser disposto e que aspectos deve possuir.

O Capítulo 3 faz uma revisão dos métodos utilizados na área, apresentando alguns dos vários caminhos que os pesquisadores têm buscado para encontrar um bom método de verificação. Demonstra também o interesse na área, medido no número de trabalhos publicados.

O Capítulo 4 fornece um estudo sobre o processamento de imagens utilizando filtros, demonstrando características dos mesmos e suas qualidades. Seu objetivo é ambientar o leitor com o processamento de imagens, permitindo que este compreenda posteriormente o motivo do pré-processamento executado. Introduce também a origem biológica do filtro aplicado neste trabalho.

No Capítulo 5, temos uma apresentação do método Elastic Matching, um histórico sobre o mesmo e um estudo de suas características. Aqui se introduz uma bibliografia sobre o tema, além de seus conceitos e utilização. O capítulo permite que o leitor perceba a riqueza do método.

Temos o Capítulo 6 com os resultados de testes implementados sobre o banco de dados de assinaturas. Os valores e métodos utilizados são definidos e os resultados se apresentam no formato definido no Capítulo 3.

Por fim, mas não menos importante, temos o capítulo de conclusões e perspectivas futuras, onde são discutidos os resultados e propostos novos caminhos para pesquisas futuras.

Capítulo 2

Histórico de Reconhecimento de Assinaturas

*“Pesquisa: Processo de entrar em
becos para ver se têm saída.”*

Barstow Bates

2.1 Alguns Aspectos Básicos

Como o sistema é formado, e quais características usa é uma questão que cada pesquisador que desenvolve sistemas de *VAM* deve decidir. Contudo, ele deve ter sempre em mente o fato que um sistema de verificação deve ser construído com base no fato de que as pessoas não escrevem sempre num padrão exato, e que há variações de acordo com cada autor.

Basicamente, pode-se dividir as espécies de sistemas pela sua forma de aquisição da assinatura [4]. Assim, têm-se os processos *Dinâmicos* que utilizam sinais digitalizados variando no tempo ($U(t)$) para representar a assinatura (por exemplo, aceleração) extraídos com uma caneta especial e os processos *Estáticos* que, por outro lado, trabalham com a imagem da assinatura digitalizada em duas dimensões, capturadas tipicamente por um scanner ou câmera. Nesta, a imagem do texto no papel fornece os dados que serão analisados de acordo com a distribuição de pixels. Apresentaremos mais detalhes sobre cada tipo de sistema no Capítulo 3, preferindo

discutir agora alguns aspectos básicos necessários à compreensão dos mesmos.

Com o intuito de executar as comparações necessárias, os sistemas de verificação geram um conjunto de referência M_i para cada usuário i , obtendo os espécimes que formam a base de dados com a qual as assinaturas devem ser testadas. Armazena-se um conjunto de uma ou algumas assinaturas e relaciona-se o autor delas através de um número de identificação (ID). Assume-se, a partir deste ponto, que tem-se como assinatura de referência o conjunto (que pode ser unitário) de assinaturas genuínas do indivíduo i (M_i) armazenado no banco de dados, a fim de ser comparado com outras assinaturas, que acredita-se serem de autoria de i . Uma assinatura A_j (assinatura de teste) é apresentada ao sistema para se verificar se esta é mesmo uma assinatura do indivíduo i , que tem seu conjunto de referência individual armazenado. Coleta-se as características da assinatura A_j , que podem ser funções de tempo ($U(t)$), como a aceleração da caneta durante a escrita da assinatura, ou parâmetros (P), como o tamanho da assinatura final, e as compara com as características deste conjunto de referência. A diferença entre A_j e M_i é computada usando uma ou várias medidas de distância que são aplicadas para verificar se esta é inferior a um determinado limiar, quando é aceita, caso contrário é rejeitada, como veremos adiante.

É útil indicar que em um sistema de identificação (ao contrário de verificação), não teríamos um identificador indireto (ID) para ser usado, e a A_j seria comparada com todas referências (ou modelos) da base de dados $M = \{M_1, M_2, M_3, \dots, M_n\}$, para se encontrar o autor.

2.1.1 Dificuldades

A verificação automática de assinaturas não pode ser considerada um problema simples de reconhecimento de padrões, devido à várias perguntas, como: Quan-

tas características são suficientes? A característica pede alguma transformação dos dados (redimensionamento, por exemplo)? Quantos exemplos são necessários para produzir a referência? A assinatura de referência poderá ser atualizada? Qual o limiar a usar para teste e referência? Estas são algumas das inúmeras perguntas que os pesquisadores devem enfrentar. Ao longo desta tese desenvolveremos alguns dados necessários ao estudo destas questões e apresentaremos nossa proposta.

Os problemas não só se restringem aos aspectos da implementação:

- Como o processo envolve pessoas, é difícil gerenciá-lo. Durante o tempo no qual se desenvolvem os testes, o estado psicológico pode afetar o processo, resultando em valores altos de divergências de saída. Isto quer dizer que, ao efetuar-se a captura dos modelos em um dia, nos dias seguintes, ao compararmos esta com as novas verificamos divergências oriundas do estado emocional. O algoritmo deve ser capaz de minimizar esta variação provocada pelo estado mental, embora eliminá-la completamente seja um fator quase impossível, devido à grande variação que este pode causar e à necessidade de não deixar o sistema muito “livre” para que falsificações não sejam aceitas.
- Não é óbvio descobrir como representar todos os potenciais usuários (homens e mulheres de diferentes idades e atividades, manidestros¹, canhotos e ambidestros) de uma população, sabendo-se que isto pode influenciar no processo no caso real. Tome-se como exemplo comparar um autor chinês com um brasileiro, o primeiro teria um sentido de escrita inverso ao segundo.
- As características que uma assinatura apresenta são também determinadas pelo tipo de educação que a pessoa recebeu. Assinaturas de pessoas com baixo nível de instrução possuem muitas distorções, pela falta de habilidade

¹Pessoas mais hábeis com a mão direita, ou que dela se servem com preferência.

de escrita. Aqui se percebe um fato interessante de, por vezes, a assinatura de uma pessoa torna-se mais “complicada” (com floreios, e por vezes, ilegível) com o desenvolvimento intelectual da mesma, sendo que se o indivíduo que possui um nível escolar baixo, retoma seus estudos e evolui intelectualmente, sua assinatura evolui também, e de tal forma que, a uma certa altura de seu desenvolvimento, não será mais reconhecida.

- A boa saúde também é importante. As características naturais de uma pessoa só transparecerão na escrita se a caneta puder se movimentar dentro de um ritmo uniforme, em traços ascendentes e descendentes; no caso de uma doença que prejudique a coordenação muscular, isto se torna impossível e a distorção fatalmente ocorrerá, tornando a assinatura irreconhecível.
- As condições não usuais de escrita podem afetar a assinatura. Por exemplo o apoio do papel, ângulo em relação ao papel, caneta, etc., produzindo uma assinatura que não é natural.
- A habilidade dos falsificadores pode chegar a um tal ponto de especialização na forma da assinatura que as dissimilaridades entre a falsa e as autênticas podem ser tomadas como dados espúrios de uma assinatura verdadeira, ou seja as pequenas diferenças serão consideradas como as variações do autor *i* que existem entre as assinaturas verdadeiras. Deve-se salientar, que isto não ocorre em todos os casos, pois a maioria dos falsários não chega a tal tipo de perfeição, produzindo muitas vezes falsificações bem grosseiras. O processo de falsificação de uma assinatura, ou a simulação da caligrafia de outra pessoa, se tem sucesso, envolve um processo duplo onde o falsário não só copia características do autor imitado *i*, mas também esconde as suas características pessoais de escrita. Se a escrita é livre e rápida, ela certamente irá mostrar, quando cuidadosamente analisada, várias características naturais do escritor

de forma que uma distinção pode ser feita.

Note que, se alguma assinatura for tomada como padrão durante uma doença ou simplesmente com um apoio do papel de uma forma incômoda, muitos problemas podem surgir, tornando o modelo distoante das assinaturas posteriores genuínas testadas. Isto é apresentado para mostrar como aspectos simples podem se tornar magnos neste processamento.

Como já foi dito, sucessivas assinaturas humanas variam de muitas formas e por isso, especialistas humanos, mesmo treinados para identificar falsificações, não são tão hábeis em aprovar todas as assinaturas legítimas. Herbst e Liu [9] citam referências que indicam que 25% de assinaturas genuínas foram rejeitadas ou não foram classificadas (seria o caso da resposta ser : não sei se é verdadeira ou falsa), por examinadores treinados em documentos, enquanto nenhuma falsificação era aceita; enquanto um grupo não treinado possuía um índice de 50 % de aceitação de falsificações. Houve também casos de especialistas humanos ficarem na dúvida para aceitar algumas assinaturas legítimas [10]. Estes dados demonstram bem as dificuldades encontradas. Se mesmo humanos possuem estas faixas de erros para boas falsificações, desejar altos índices de verificação correta sob as mesmas circunstâncias para as máquinas torna-se um objetivo difícil de alcançar.

Embora existam estas variações, Hilton [11] notou que a assinatura envolve padrões particulares, e a grande maioria dos usuários, uma vez que tenham o estilo de assinatura estabilizado, não a variam muito. Há casos de variações grandes em relação à última versão, devendo a assinatura ser sempre atualizada. Em alguns poucos casos, uma pessoa tem uma versão especial para assinar documentos especiais. O sistema deve ser capaz de absorver estes casos particulares, ao formar o conjunto de modelos, devendo manter esta lista de referência no decurso do processo.

Quando a assinatura varia no tempo, como então encontrar os fatores invariantes? Esta é uma questão difícil de solucionar. Na vasta maioria dos casos a assinatura corrente é suficiente para a proposta de verificação e a atualização do banco de dados irá excluir este problema.

2.2 Modelando o Sistema

A verificação automática de assinaturas pode ser representada como um problema de particionamento em duas classes. Seja α o universo de assinaturas, existe neste um subconjunto α^i com todas assinaturas que são indicadas como relativas ao autor i . Idealmente, uma decisão binária deve particionar α^i em [4]:

- α_1^i : classe de assinaturas genuínas e
- α_2^i : classe de falsificações²

onde $\alpha_1^i \cup \alpha_2^i = \alpha^i$, o subconjunto das genuínas de α^i abrange todas as variações de assinaturas que o indivíduo i produziu ou produzirá; e a classe de falsificações representa todas as assinaturas que não foram produzidas pelo mesmo, sendo formado por assinaturas normais de outras pessoas e tentativas de falsificações envolvendo a assinatura de i .

O processo de decisão avalia a resposta do algoritmo de comparação com respeito a um limiar (T_0) para aceitar ou não a assinatura. Por exemplo, usando uma *medida de similaridade* $d(A_j, M_i)$ entre uma assinatura teste A_j e uma assinatura de referência M_i : se $d(A_j, M_i) \geq T_0$ então A_j é aceita. ou se $d(A_j, M_i) < T_0$ então A_j não é aceita.

² Note que as duas classes estão intimamente relacionadas com o indivíduo i , e só existem para o mesmo. Sem a indicação do autor, não pode-se estabelecer estes grupos.

Estas inequações se invertem no caso de se utilizar medidas de diversidade e serão amplamente utilizadas neste trabalho.

O sistema de reconhecimento de padrões atua mapeando o objeto estudado em um conjunto de características, o que resulta numa nova forma de representação da entrada (Figura 2.1). O padrão, no caso uma assinatura, é representado como um ponto em um espaço Euclidiano de dimensão n , o qual é obtido ao se apresentar a assinatura de acordo com n “dimensões” (características ou atributos) selecionadas que definem suas coordenadas neste novo universo, que supostamente, deve ser invariável ou mesmo insensível com respeito a variações ou distorções, além de conter menos redundâncias [12, 13].

Neste contexto, o classificador de padrões, que no nosso caso é o algoritmo do processo de decisão, mapeia os pontos do espaço de características no espaço de categorias. Mediante os valores das coordenadas destes define-se o conjunto ao qual as assinaturas, que são representadas pelos mesmos, pertencem³, a fim de conseguir segmentar o espaço de categorias em regiões de decisão que limitam as áreas de cada categoria. Em qualquer caso, a ação resposta é executada pelo sistema de saída de acordo com o resultado desta separação, na medida que para cada região deve existir uma informação acerca do que indica seu mapeamento para a saída.

Na Figura 2.2 está representado um espaço que foi segmentado por uma reta de decisão que define bem as duas regiões dos Padrões 1 e 2. A entrada é mapeada por duas características que iremos definir neste ponto, a título de exemplo, como altura e comprimento da assinatura. Estes dois parâmetros são representados pelas retas ortogonais da figura. Uma assinatura teste ao ser mapeada neste espaço a partir destas características, é definida como um ponto no mapa. Esta localização

³Nesta tese, só há duas categorias, definidas para todos sistemas de verificação, a verdadeira e a falsa.

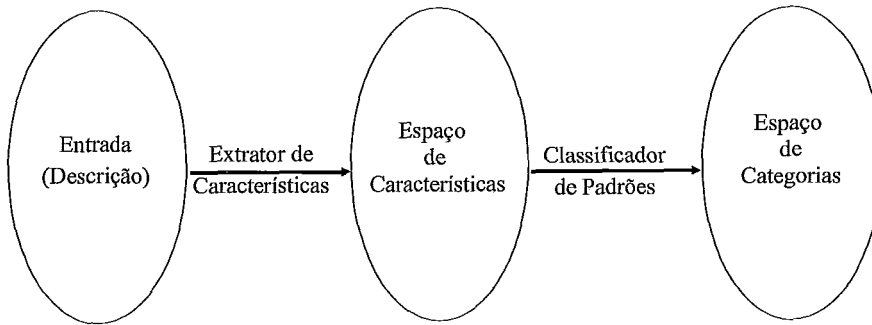


Figura 2.1: Mapeamento do espaço de entrada em um espaço de categorias.

tem como objetivo estipular a que Padrão a assinatura pertence, a partir da reta de decisão. Ao considerarmos o Padrão 1 como a classe de assinaturas verdadeiras, teremos o Padrão 2 como as falsas. Assim quando o ponto gerado pelos dados da assinatura estiver a esquerda da reta de decisão, teremos a assinatura como verdadeira e a direita como falsa. Para definir esta área, utiliza-se um conjunto de funções que devem ser capazes de associar valores mais altos para pontos dentro de uma região do que os que estão fora, discriminando-os para esta classe, sendo denominadas funções de decisão [14]. O objetivo da função discriminante é desenvolver os limites dos espaços de decisão. É claro que como o espaço aqui descrito tem baixa dimensionalidade, possivelmente terá um resultado ruim, embora um alto número de dimensões possa prejudicar o processo de outra forma.

Ainda que exista certa estabilidade nas assinaturas, ou seja, as assinaturas de uma mesma pessoa mantenham certas características entre si, o processo de seleção não é preciso, repetível e idêntico para várias tentativas, devido aos problemas apresentados na Seção 2.1.1. Isto quer dizer que existirá uma variação entre os valores de respostas das assinaturas aos sistemas de verificação, mesmo para assinaturas de um mesmo indivíduo. Em outras palavras, podemos dizer que o conjunto de assinaturas genuínas possui uma grande dissimilaridade e estas diferenças provocarão

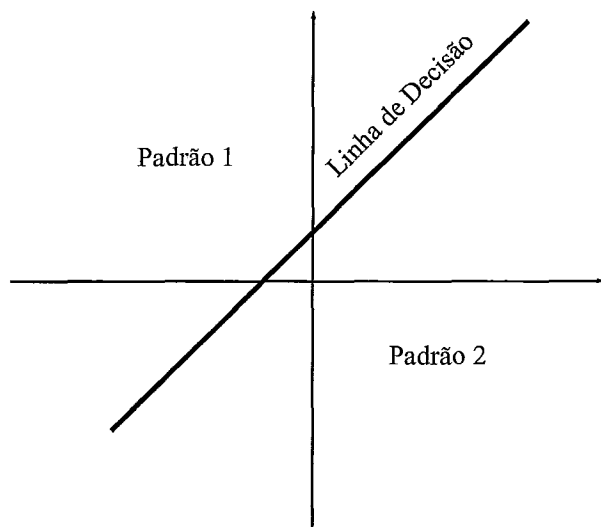


Figura 2.2: Espaço de categorias apresentando um linha de decisão que o segmenta em áreas de decisão.

uma resposta diferente em qualquer algoritmo de reconhecimento que as busque, de acordo com o grau de divergência entre a assinatura testada e o exemplo tido como padrão. Por exemplo, se considerarmos o tamanho da assinatura como uma característica de comparação, quando compararmos o modelo M_i com a assinatura A_j , mesmo que esta assinatura testada seja do indivíduo autor do modelo, existirá uma variação no tamanho dela em relação ao modelo, e a proporção desta diferença do valor esperado, dado pelo modelo, é que fornece o índice de variação na saída. Como nenhuma pessoa irá medir sua assinatura para que esta fique sempre exatamente do mesmo tamanho, pode existir uma diferença grande nos valores de similaridade apresentados no sistema, para assinaturas de uma mesma pessoa. Vale a pena lembrar que quando ocorre o fato de duas assinaturas serem exatamente idênticas, uma delas é uma falsificação (veja na página 2 a nota do rodapé). De fato, uma grande variação pode ser observada nas assinaturas de acordo com o país, idade, hábitos, estado psicológico e mental, condições práticas e físicas. Para fins de verificação, faz-se necessário estudar as variações de assinatura em dois grupos

distintos [4]:

- $V_w i$: diferença intraclasse ou intrapessoal do indivíduo i , que diferencia as espécies de assinaturas genuínas dentro da classe α_1^i de um indivíduo. É a variação que o indivíduo i produz ao assinar várias vezes, e faz com que nenhuma assinatura do mesmo seja igual.
- $V_b ij$: dissimilaridade interclasse ou interpessoal entre os indivíduos i e j , que diferencia as espécies de assinaturas entre as classes α_1^i e α_1^j dos indivíduos i e j . Ela distingue uma assinatura entre pessoas diferentes, indicando pela caligrafia e forma de assinatura de cada um dos indivíduos.

Em teoria, V_w deve ter o menor valor possível e V_b ter o valor alto o suficiente para ser usada na separação de classes. Desta forma, um objetivo dos sistemas de verificação é capturar o melhor conjunto de características, que quando utilizadas na comparação, eliminem $V_w i$ e acentuem $V_b ij$, produzindo a situação ideal vista na Figura 2.3, com a separação total entre o conjunto de assinaturas genuínas (α_1^i) e falsas (α_2^i). Na prática estas classes não serão bem definidas, e temos dois tipos de erros da verificação: a *imitação* e a *degeneração*, ilustrados na Figura 2.4.

Para melhor explicar os diferentes erros que podem existir estudaremos as saídas resultantes do sistema de verificação. A saída possui uma forma binária que pode ser classificada como *aceitação* (χ_1^i) onde a assinatura de teste é aceita como uma assinatura de i , ou *rejeição* (χ_2^i), quando o sistema assume que houve uma falsificação. Um terceiro estado, o de indefinição pode ser implementado; ele seria o resultado do sistema quando este não conseguir definir se é uma falsificação ou não, indicando a assinatura para uma investigação posterior.

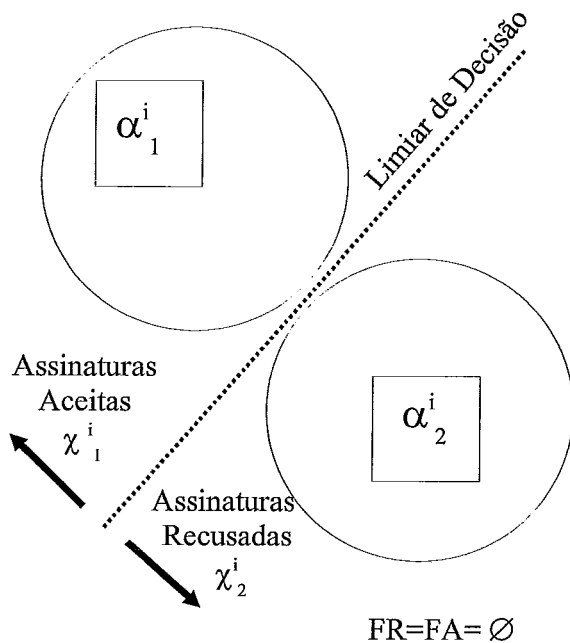


Figura 2.3: Ilustra-se a separação perfeita, com total separação entre assinaturas verdadeiras α_1^i e falsas α_2^i . A situação ideal. FA e FR denotam Falsa Admissão e Falsa Rejeição.

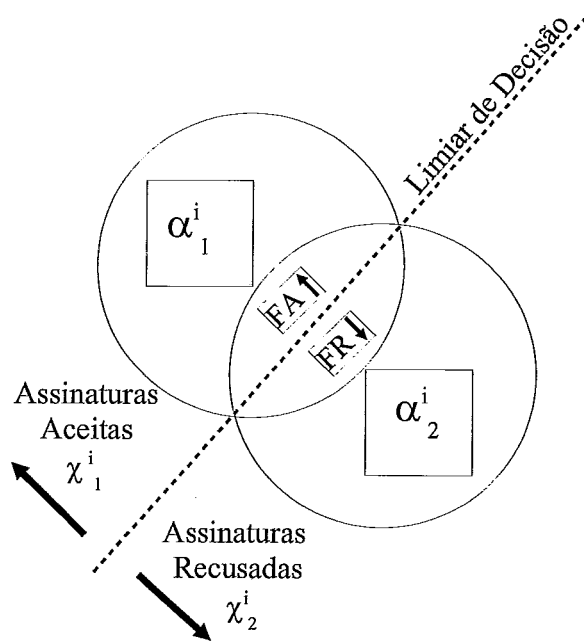


Figura 2.4: A separação imperfeita que ocorre no sistema real. Os conjuntos não são disjuntos, e na área de intersecção gera-se as FA e FR, discriminadas por onde o limiar de decisão corta os conjuntos.

2.2.1 Saídas de Aceitação (χ_1^i)

No caso do sistema aceitar a assinatura como verdadeira ($\in \chi_1^i$), teremos a possibilidade de ocorrência de dois casos:

- *Autêntica*(α_1^i): Ocorre quando a assinatura testada (A_j) é do autor i da referência e A_j tem boa similaridade com o modelo armazenado de i , M_i ; sendo este valor de similaridade superior (ou igual) ao limiar de decisão T_0 que define-se como o menor índice de similaridade que a assinatura deve ter para ser aceita. Na forma matemática:

$$d(A_j, M_i) \geq T_0$$

com $j = i$, $A_j \in \alpha_1^i$ e $A_j \in \chi_1^i$. A assinatura é genuína e é aceita como tal, e o sistema acerta.

- *Imitação*: Quando a assinatura testada (A_j) é de autoria de j que não é o indivíduo i da referência e se A_j tem similaridade superior (ou igual) ao limiar de decisão T_0 com o modelo armazenado de i , M_i . Na forma matemática:

$$d(A_j, M_i) \geq T_0$$

com $j \neq i$, $A_j \notin \alpha_1^i$ e $A_j \in \chi_1^i$. Ou, mais claramente, é uma falsificação que o sistema não foi capaz de recusar. Um erro do sistema que será definido a partir daqui como *Falsa Admissão (FA)*.

2.2.2 Saídas de Rejeição (χ_2^i)

Já no caso do sistema recusar a assinatura como verdadeira ($\in \chi_2^i$), teremos:

- *Degeneração*: Ocorre quando a assinatura testada (A_j) é do autor i da referência e A_j não tem um índice de similaridade superior ou igual ao T_0 com

o modelo armazenado de i (M_i). Matematicamente:

$$d(A_j, M_i) < T_0$$

com $j = i$, $A_j \in \alpha_1^i$ e $A_j \notin \chi_1^i$. Este outro erro do sistema que será definido a partir daqui como *Falsa Rejeição (FR)*.

- *Falsa* (α_2^i): Quando a assinatura testada de j (A_j), que não é o autor da referência i , não é semelhante com o modelo armazenado de i (M_i). A falsificação é reconhecida como tal, ou seja:

$$d(A_j, M_i) < T_0$$

com $j \neq i$, $A_j \notin \bar{\alpha}_1^i$ e $A_j \notin \chi_1^i$. O sistema acerta.

2.2.3 Relações das Saídas

É válido dizer que as causas básicas dos erros podem ser duas:

- o sistema não foi capaz de absorver todas as variabilidades pessoais de i e quando i produziu uma assinatura que diferenciou-se nos aspectos que o sistema utiliza para a verificação, a mesma foi recusada;
- pode ter sido gerada uma assinatura que fugiu completamente dos padrões conhecidos armazenados no sistema, pois como uma assinatura é gerada por uma pessoa, e esta é sujeita as mais variadas situações em sua vida o que pode provocar uma modificação para uma forma não usual da assinatura.

Desta forma, temos a relação entre a saída do sistema (χ_1^i e χ_2^i) e o conjunto real de assinaturas relativas ao indivíduo i (α_1^i e α_2^i) dado pela Equação 2.1:

$$\chi_1^i = \alpha_1^i + FA - FR \tag{2.1}$$

O conjunto de assinaturas aceitas pelo sistema (χ_1^i) deve ser formado, principalmente, pelo conjunto de assinaturas verdadeiras (α_1^i) para que o sistema tenha um bom desempenho. Como o sistema pode errar e não aceitar todas as assinaturas do indivíduo i , deve ser subtraído o conjunto de assinaturas verdadeiras que estão contidas em α_1^i e são recusadas (FR), e acrescentado o conjunto de assinaturas que não pertencem a α_1^i , e portanto não são de i , mas que também são aceitas (FA).

$$\chi_2^i = \alpha_2^i + FR - FA \quad (2.2)$$

O conjunto de assinaturas recusadas pelo sistema (χ_2^i), ver Equação 2.2, deve ser formado pelas assinaturas falsas (α_2^i) das quais subtraímos o subconjunto de assinaturas falsas aceitas (FA), e somamos o subconjunto de assinaturas verdadeiras, que não pertencem ao α_2^i , mas também são recusadas (FR).

2.3 Desempenho Sobre Falsificações

Sob as definições da Seção 2.2, o desempenho do sistema é avaliada de acordo com a representação de um problema de reconhecimento entre duas classes: a autêntica e a falsa; com base nas faixas de erros de dois tipos:

- *TIPO I* (FRR -false rejection rate): taxa de erro de falsa rejeição
- *TIPO II* (FAR -false acceptance rate): taxa de erro de falsa admissão

Neste contexto, pesquisadores produzem testes utilizando um conjunto de assinaturas genuínas e falsas, para descobrir o quanto seus sistemas são robustos e válidos para a verificação automática de assinaturas. Ao determinarem a faixa de *falsa rejeição* e a faixa de *falsa admissão* da técnica estudada em relação a um banco de assinaturas podem comparar estes valores com outros métodos e tirar conclusões

quanto a qualidade da abordagem escolhida. Obter uma boa estimativa de FAR é difícil por causa da necessidade de “falsificações genuínas”, ou seja, falsificações que poderiam ocorrer numa situação real, que podem ser discriminadas nas possibilidades:

- Simples: falsificador não simula os traços da assinatura genuína (simplesmente “escreve” o nome de i com sua caligrafia)
- Aleatórias: aqui, testa-se o uso da própria assinatura (ou uma qualquer) ao invés da testada. Neste caso, não existe a preocupação nem de saber o nome do autor da referência sob a qual a assinatura é testada; é também denominada falsificação de menor esforço, ou esforço zero.
- Especializadas: um falsário pratica, imitando a informação da assinatura por meio de treinamento e teste.

A avaliação de desempenho para os sistemas de verificação fia-se sobre dois tipos principais de falsificações: uma feita por um indivíduo que não tenha acesso as assinaturas genuínas e outra que é feita por uma pessoa que tem acesso às verdadeiras. Utiliza-se para o primeiro teste as aleatórias e para o segundo as especializadas. Testes com as primeiras falsificações “tendem” a ter menor FAR que as com peritos.

Com já foi dito, devemos escolher excelentes características para termos uma boa separação entre os conjuntos (2.3). Esta separação completa de verdadeiros e falsos não admite imitação ou degeneração que são os geradores de FAR e FRR , respectivamente. Porém em situações práticas, falsificações não são avaliadas (ou melhor, todas suas possibilidades não o são) e a seleção de características não é satisfatória, ocorrendo um conjunto de características que não é ideal, aparecendo as quatro espécies de resposta descritas na Seção 2.2, resultando na nova distribuição que pode ser vista na Figura 2.4.

Iniciaremos um estudo gráfico necessário para visualizar uma característica presente em todos os sistemas de verificação: a relação dinâmica entre FAR e FRR . Os gráficos são demonstrados nas Figuras 2.5 e 2.6. Eles possuem duas curvas. A que decresce no sentido do incremento do limiar é a taxa de erro do tipo II e representa a taxa de assinaturas falsas aceitas que decresce exponencialmente de acordo com o maior índice de similaridade exigido. A curva que cresce representa a taxa de erro do tipo I de assinaturas verdadeiras que são rejeitadas pelo alto índice de semelhança que exige-se para limiares de valores altos. O limiar T_0 representa o valor escolhido pelo sistema, que distingue à sua esquerda as assinaturas que são rejeitadas e à direita as aceitas, com a reta tracejada disposta para indicar que valores o limiar apresenta para FAR e FRR . No gráfico da Figura 2.5, representando o sistema ideal, a escolha do limiar não é problemática. Porém, no gráfico do sistema real da 2.6, isto não ocorre e, dependendo do comportamento desejado, deve-se estabelecer o limiar.

Se o objetivo do sistema de verificação for eliminar o erro do tipo II, adotar como limiar de decisão o valor T_0'' seria mais correto que T_0 , na medida que o valor de FAR ali é bem menor que em T_0 . Assim, desloca-se a linha tracejada do limiar para a direita, ela cortará a curva de FAR em um ponto mais baixo que T_0 . Porém, verifica-se o fato que a FRR nesta região será bem mais alta, o que significa que está conseguindo-se dirimir falsificações pela eliminação de várias assinaturas verdadeiras, o que pode ser um problema, dependendo da aplicação. Imagine, por exemplo, que em um banco o indivíduo tenha quase todas as suas próprias assinaturas rejeitadas, e os seus cheques, assinados pelo próprio, seriam recusados pelo sistema.

Utilizar T_0' que elimina o problema de FRR terá como consequência um crescimento da FAR , e em decorrência disto, não combaterá com afincos assinaturas

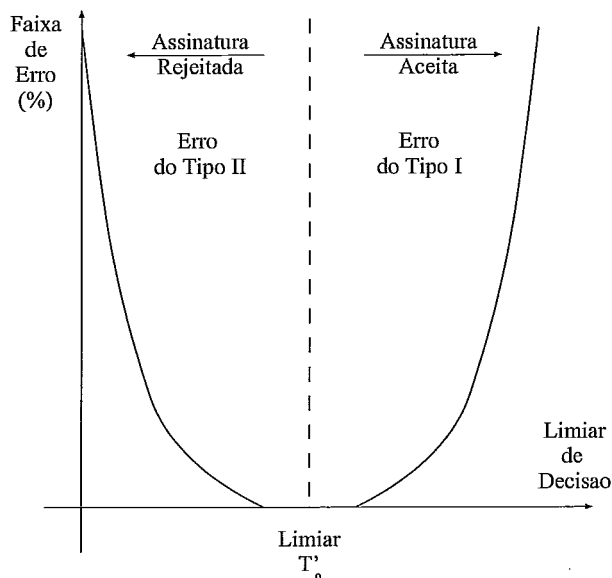


Figura 2.5: Gráfico demonstrando como o limiar afeta os valores da faixa de erro em uma situação ideal. Com a separação perfeita entre classes o limiar T'_0 torna-se o valor perfeito para indicar o grau de similaridade que a assinatura testada deve possuir

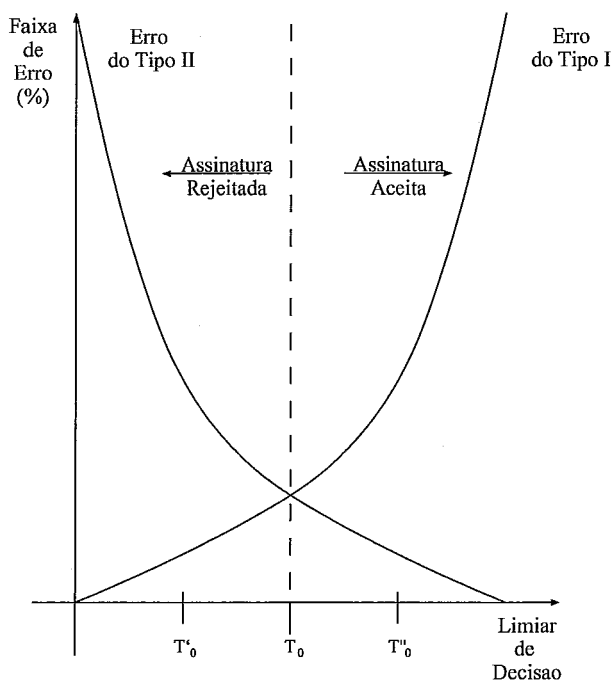


Figura 2.6: Gráfico de Erros do sistema real, ele apresenta a dificuldade de estabelecer um correto limiar de decisão. Ao decidirmos por T''_0 , uma faixa de erro do tipo I alta, onde várias assinaturas verdadeiras serão rejeitadas. Ao optar por T'_0 , tem-se uma faixa de erro do tipo II alta, e aceita-se muitas falsificações. O limiar deve ser escolhido de acordo com as características da implementação.

ilegítimas, que pode ser uma dificuldade em algumas aplicações. Uma forma de representar a composição destes erros é computar a taxa de erro total que é a média das duas [15].

Na realidade, a aplicação é que definirá qual o índice de escolha T, T_0'' ou T_0' . Para alguns [2], FRR deve ser quase nula e FAR pode ter um valor médio (entre 10% e 20%). Pode-se também estabelecer que o valor de FRR deve ser ajustado para zero aceitando valores de até 25% para FAR [16]. Outros sistemas podem estar mais interessados em recusar falsificações e aceitar índices de FRR altos para isto.

Se um conjunto bom de exemplos de falsificações não for avaliado (ou não o for em número suficiente), α_2^i nunca é conhecido completamente e técnicas estatísticas não serão apropriadas. Mas dependendo das condições de teste e projeto, os autores podem indicar diferentes tipos e falsificações para validar seus sistemas. É claro que, nos casos onde só falsificações aleatórias são usadas na avaliação de desempenho do sistema, os erros tipo I e II são artificialmente reduzidos.

2.4 Desenvolvendo os Passos

Em seu levantamento, Plamondon e Lorette [4] sugerem que a maioria das técnicas de verificação envolvem as seguintes fases:

- Aquisição de Dados
- Pré-processamento
- Extração de Características
- Processos de Comparação
- Avaliação de Desempenho

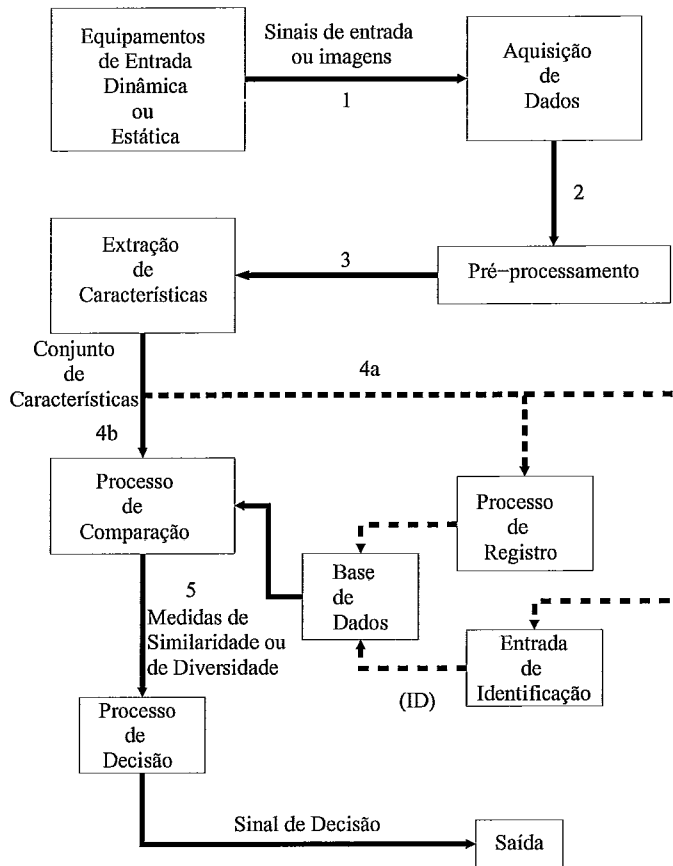


Figura 2.7: Metodologia Básica dos Processos de Reconhecimento

Na Figura 2.7, pode-se observar o fluxo do processamento das assinaturas nas quatro primeiras fases, que são geralmente implementadas em todos os sistemas. A última fase, a *Avaliação de desempenho*, não é incorporada diretamente ao sistema de verificação, servindo para indicar o desempenho de um sistema perante os outros, e por isso não aparece nesta figura. As assinaturas são adquiridas e pré-processadas para eliminar informação espúria e filtrar sinais ou imagens importantes (1 e 2). O próximo passo é referente ao processo de extração de características (3), onde funções específicas e/ou parâmetros P , são usados para representar a assinatura. Com este conjunto de características, tem-se a informação da assinatura que podem seguir dois caminhos diversos, o primeiro (4a) é seguido durante a formação do banco de dados, que são armazenadas junto com o *número de identificação (ID)*,

que indica o autor, a fim de serem buscados na verificação. O processo de verificação em si mesmo, percorre o curso de (4b) indo para a comparação e, via *ID*, recebendo o modelo armazenado (ou as características do mesmo). O *ID* é importante para identificar quem vamos ter de comprovar como sendo o autor. Após a comparação, possuiremos os valores de saída (5) para o processo de decisão, que gera o sinal final de decisão. Estes passos serão descritos mais detalhadamente à frente.

Muitos métodos, mas não todos, durante as três primeiras fases (aquisição, pré-processamento e extração de características), podem gerar uma assinatura de referência para cada indivíduo, sendo necessário um número mínimo de assinaturas do usuário para capturar a forma geral da assinatura.

2.4.1 Aquisição de Dados

Sua importância no processo advém do fato de que a qualidade dos sinais apresentados é crítica para a comparação, tempo de execução e menor pré-processamento. Altas distorções irregulares podem ser causadas por perspectivas e ângulos de aquisição, além de movimento durante a aquisição da imagem provocando as distorções que serão mais um problema a ser tratado durante o processamento. O equipamento utilizado aqui irá definir o estilo do processamento, se será *estático*, quando utiliza-se um “scanner” ou câmera; ou *dinâmico*, utilizando uma mesa digitalizadora e/ou canetas especiais. Estes estilos são melhor detalhados no Capítulo 3.

2.4.2 Pré-processamento

Este passo é a preparação da informação relevante a partir de vários processos: localização da assinatura, correção da não uniformidade desta, extração da assinatura das características que compõem o fundo da imagem, partição, aplicação de

limiares e filtragem e redução de dados. Pode-se aplicar também transformações geométricas, com o objetivo de eliminar rotação translação e mudanças de escala, normalizando-as.

O primeiro passo de encontrar a assinatura em um documento, em geral, não é um trabalho muito difícil, devido ao fato da maioria dos experimentos trabalhar com cheques, os quais têm a posição da assinatura padronizada, permitindo a utilização de um operador *janela*. Este operador é um algoritmo que, dado o formato da imagem, a partir de uma referência desta, seleciona a área. No caso de ter-se uma grande variação de posições no documento, pode ser um problema extrair a assinatura do mesmo, sem nenhum artifício, como uma caneta com tinta especial que seja facilmente identificável. De posse desta, o objeto ocupa uma faixa distinta do nível cinza, maior que o fundo. Então pode ser extraído por um limiar:

$$O(x, y) = \begin{cases} 1 & \text{se } f(x, y) \geq \theta \\ 0 & \text{se não} \end{cases}$$

onde $O(x, y)$ é a imagem com o objeto extraído, $f(x, y)$ é a imagem original, e θ é o limiar. Caso isto não ocorra, é necessário implementar processos mais elaborados de segmentação para distinguir a assinatura do fundo. Pode-se usar polígonos de ajuste para esta extração, pois estes avaliam a forma da assinatura e podem servir como uma característica básica para a próxima etapa, uma vez normalizados.

Dependendo de quais características serão utilizadas na próxima fase, o pré-processamento pode fazer alterações mais específicas na imagem extraída com o intuito de transformar a imagem de tal modo que a informação seja mais facilmente discernível ou melhor aplicada. Por exemplo, ao descobrir-se quais aspectos são mais variáveis na imagem, pode-se eliminá-los da comparação por possuírem pouca

informação, além de descobrir os que são mais importantes e discriminar um peso maior para a comparação executada com eles.

Outro exemplo de alteração específica é vista no algoritmo que avalia a faixa de alta pressão na assinatura, baseado na hipótese que um alto nível de cinza na imagem revela uma região de alta pressão, para utilizá-la na próxima etapa do processo: a extração de características, que foi proposto por Sobel [4].

No final deste passo é comum as assinaturas se encontrarem num formato canônico pelo ajuste de parâmetros das características para valores pré-especificados. Com isto, a parte da sua variabilidade intrapessoal derivada de distorções geométricas, por exemplo, é atenuada, o que pode ser visto nesta tese, no Capítulo de Resultados.

2.4.3 Extração de Características

A divisão de características em sensíveis ao texto e insensíveis, como é proposto (e utilizado) por especialistas humanos, é de difícil aplicação. Embora todo humano possa, uma vez treinado, executar este método de comparação, é difícil saber como ele penaliza todo o processo desde a percepção visual da assinatura até as etapas neuronais de reconhecimento e processamento para possibilitar implementá-lo num algoritmo. Estes métodos são ligados aos vários fatos particulares à estrutura da assinatura em si, como exemplo:

- as características de abertura e fecho⁴ de escrita, relativas ao floreio que muitas assinaturas apresentam nestes lugares
- os retoques (o fechamento posterior dos laços dos “l”, “e”, “g” e “o”, por exemplo) porque revelam uma preocupação com o formato, o que, em geral,

⁴ Abertura e fecho são denominações do início e fim da assinatura, respectivamente.

não é feito pelos autores verdadeiros.

- uma constância em toda a assinatura, com uniformidade de estilo, sem partes entrecortadas e irregulares e indicações de escrita lenta, como traços curvos, interrompidos ou tremidos.

Estes itens podem ser utilizados por especialistas humanos na verificação. Porém, no caso da automatização, determinam uma relação incômoda, pois o número de caracteres ou símbolos num espécime de assinatura é em geral irrelevante (no contexto alfabético) para os sistemas, que não buscam estas particularidades. Ou melhor, o sistema teria de ser dependente da espécie de assinatura que a pessoa faz, sendo que para cada uma teria de existir um processo de decisão, onde se buscasse os itens acima os que melhor se adaptariam ao processo, o que é um problema difícil de resolver para a automatização.

Desta forma, pesquisadores não se preocupam em imitar o processo utilizado por especialistas humanos e sim em utilizar características que possam ser extraídas pelo computador de qualquer assinatura, não importando detalhes muito específicos. Um vetor de características que representam a assinatura será gerado. Nem todas as informações (ou características que podem ser extraídas da assinatura) precisam ser usadas, ao contrário, um índice muito grande de características irá fazer o processo ficar lento na etapa de comparação, além de torná-lo difícil de ser administrado no sentido de determinar-se os limiares do processo e que características estão influenciando para os resultados. Uma coleção de valores são suficientes para comparação (como o comprimento total da assinatura, centroídes, ângulo, entre outras) e o autor do processo é que deve determinar qual o número destas e quais são as mais importantes para o mesmo. Uma boa forma de escolher as características é verificar os pontos abaixo no momento de extraí-las:

- O sistema de informação tem bom desempenho para reconhecimento com uma das classes desconhecidas?
- O espaço de características consegue absorver as variações de escrita?
- O custo para extração do conjunto de características é alto?
- As características, além de assinalar o padrão de uma classe, descrevem aspectos ilegíveis para as outras?

Embora a respostas a estas perguntas seja ainda uma grande controvérsia no campo de *Verificação*, pois as respostas não são simples.

Intuitivamente, as variações intra e inter-pessoais são dependentes do nível de características utilizadas. Para um nível baixo, todas são similares. Por exemplo, qualquer uma tem um formato retangular consistindo de traços de escrita. Em um nível alto de detalhes, todas são diferentes. Um sistema de verificação pode não ter sucesso se “tender” a um destes extremos, visto que, se o nível for baixo, falsificações são aceitas, e alto, rejeita assinaturas verdadeiras com pequenas variações [15].

A assinatura tem de ser comparada por uma correlação entre valores da assinatura de teste e o modelo correspondente. A princípio, comparações simples ponto a ponto são ambíguas porque dada uma característica, ela pode vir a variar em vários lugares, e duas assinaturas genuínas de uma mesma pessoa, podem variar significativamente. A correlação também pode ser afetada por translação, rotação e escala e em sua comparação, sendo necessário usar algum tipo de alinhamento e redimensionamento.

As características que são utilizadas serão discutidas no Capítulo 3. Mas, uma vez selecionadas, temos duas sequências:

- apresentando uma assinatura modelo na formação e atualização da base de dados, é necessário somente armazenar (ou combinar) seus valores para formar a assinatura de referência.
- uma assinatura de teste precisamos também retirar a suas características, só que o objetivo agora é compará-las com as do modelo (o que pode representar uma compressão de informações).

Neste ponto, se o sistema está sendo inicializado, deve-se decidir como gerar o conjunto de referência, escolhendo o número de assinaturas usadas. Uma espécie $M_i = \{M_1\}$ por autor facilita a comparação, mas não parece ser representativa. Usar o valor médio de algumas espécimes não é necessariamente melhor, visto que seria como indicar uma assinatura M_i . Vejamos isto com um exemplo: considere que foram extraídas de um conjunto de assinaturas o vetor de características \vec{v}_x , sendo este formado pelas seguintes características: tamanho da assinatura, ângulo de escrita e centro de massa da assinatura. Os vetores de 3 assinaturas modelo (M_1 , M_2 e M_3) são: $\vec{v}_{x1} = \{10.5, 5.6, 6.4\}$, $\vec{v}_{x2} = \{9.0, 4.4, 7.3\}$ e $\vec{v}_{x3} = \{11.0, 4.4, 8.5\}$ são obtidos, o vetor médio \vec{v}_{xm} seria dado por $\vec{v}_x = \{10.17, 4.8, 7.4\}$, ele pode ser tão diferente de qualquer uma assinatura A_j como qualquer um dos M_i . Manter um conjunto grande de assinaturas $M = \{M_1, M_2, \dots, M_n\}$ afeta a memória necessária e o algoritmo de decisão. Similarmente, os processos de atualização podem ser executados com a evolução da assinatura de referência a partir de exemplos do usuário e tentativas de falsificação. Osborn [2] verificou que variações na escrita habituais são fáceis de ser notadas em qualquer coleção de assinaturas genuínas produzidas em datas diferentes sob uma variedade de condições. Esta investigação levou-o a afirmar que 5 assinaturas formariam uma base satisfatória, sendo que 10 seria melhor que 5.

2.4.4 Comparação

O processo de comparação é baseado na idéia que o conjunto de características extraídas de assinaturas verdadeiras é mais estável que o conjunto das características de falsificações (variações intrapessoais menores que interpessoais). Então, uma assinatura desconhecida pode ser classificada de acordo com a similaridade com o conjunto de referência M_i especificado pela entrada de identificação (ID).

A comparação pode ser feita usando as diferenças dos valores das características extraídas. Estes métodos computam a medida da distância $d(A_j, M_j)$ entre o modelo M_j e a assinatura testada A_j ; tomando a decisão a partir de um limiar de decisão (T_0) pré-estabelecido por um conjunto de treinamento, utilizando, em geral, uma função linear discriminante.

Um sistema que possua vários níveis de decisão que faça comparações simples primeiro e depois passe as mais elaboradas a fim de eliminar rapidamente falsificações grosseiras pode ser implementado, melhorando o desempenho do sistema. Ele é particularmente útil se notarmos que falsificações aleatórias e simples representam 95% dos casos geralmente encontrados [17].

A comparação não precisa também ser linear, comparações elásticas tem sido aplicadas também, assim como árvore de decisão e correlação regional.

Mais recentemente, surgiram trabalhos com Redes Neurais que trazem a vantagem de poder retreinar a rede com assinaturas recentes (assimilando mudanças que ocorreriam com o tempo), mas que requerem um grande número de espécimes, o que comercialmente se torna difícil [18].

Pode-se usar também utilizar duas ou mais técnicas em paralelo criando sistemas híbridos, onde a saída seria dada pelo resultado destas em combinado, sendo possível definir até mesmo pesos diferenciados para as técnicas mais confiáveis).

2.4.5 Avaliação de Desempenho

A mais comum medida de desempenho é o poder de identificação. Esta medida é definida pelo “escorregadio” par de índices estatísticos de erros do sistema (conhecidos como FAR e FRR) já apresentados, que são balanceados por limiares. E, embora seja essencial verificar estes índices, eles podem não ser um indicador verdadeiro, se não representarem uma população grande. Além disso, como os métodos de avaliação variam muito, é difícil compará-los. Inclusive, as bases de dados usadas diferem, assim como tamanho e população dos dados. Na prática, a escolha do método reflete a natureza da aplicação, número de usuários, memória e faixas de erros aceitáveis.

Muitas pessoas não gostam de ter suas assinaturas armazenadas num computador. Assim, a maioria das assinaturas são de voluntários do laboratório de pesquisa, com uma porcentagem baixa de pobres, velhos, incapazes, doentes (de artrite, por exemplo), que são uma porcentagem grande da população e a quem o sistema também deveria atender. Isto é visto nas faixas de erros obtidas acima para grupos mais representativos e heterogêneos, onde, a maioria dos sistemas tiveram faixas de erro de 10-20% [4], levados a populações não familiares tecnicamente (definiremos este conjunto heterogêneo, daqui para frente como conjunto de *campo*). O alto índice de FRR com o uso de um grupo mais representativo não é surpresa porque o processo de recolher assinaturas no *campo* não é tão consistente como quando as assinaturas da base de dados são coletadas, visto que haverá uma variabilidade maior entre os fornecedores de assinaturas, que não precisam estar na mesma faixa de idade dos professores ou estudantes do laboratório de onde foi produzida a base de dados de teste.

Poucas tentativas de avaliação foram executadas entre os vários estudos feitos na

literatura [4] por serem críticas, devido ao fato de muitos pesquisadores terem suas bases de dados de teste com números diferentes de assinaturas genuínas (alguns com 10, outros com 100); alguns com falsificações habilidosas, outros não. Alguns filtram a base de dados, eliminando assinaturas não aceitáveis, o que torna a verificação mais simples. O número de assinaturas para produzir a de referência varia, assim como diferentes testes e limiares têm sido utilizados, além de definições de FAR e FRR. O sistema pode ser implementado para possuir FAR zero, por exemplo [16]. Há também respostas com a combinações destas como a FT, que seria a média das duas [15, 19]. Alguns usam limiares diferentes para cada teste, e outros não.

Os pontos mais referidos como falhas comuns nos sistemas relacionam-se a problemas com a aquisição do sinal de entrada, problemas no scanner ou erros de quantização do sinal. Aquisição incompleta de espécimes, afinação do traço ou com a tinta usada (além do papel de fundo). Ocorreram também problemas com a excessiva variação nas assinaturas de alguns usuários [4].

2.5 Alguns Aspectos Finais dos Processos

As assinaturas têm formas diferentes, e há uma variabilidade, mesmo em assinaturas de pessoas que tem características similares (mesmo idioma, idade, condição social, etc). Alguns escrevem o nome, enquanto outros fazem assinaturas que lembram vagamente o seu nome, ou até não se relacionam com este. Algumas assinaturas podem ser mais simples que outras, o que determina uma facilidade maior de falsificação. É interessante notar que alguns estilos de assinaturas demonstram o ambiente que o autor vive. Por exemplo [2], um grande número de pessoas usam seus nomes como assinaturas, “escrevendo-os” com alguns floreios. Enquanto isso outros assinam traços, sem a preocupação de escrever algo. Assim um sistema que teve sucesso com o primeiro grupo, pode não ter um desempenho tão boa para o

segundo [20].

O algoritmo utilizado para comparação é só um dos aspectos que devem ser observados para uma utilização comercial. Algumas outras questões [2] têm certa relevância, devendo ser observados por um sistema comercial para detectar falsificações:

- A posição da assinatura é natural?
- A assinatura foi escrita após o documento?
- A assinatura foi escrita antes do papel ser dobrado?
- A idade aparente da tinta combina com a data do documento?
- O documento contém abrasivo químico, traços de borracha, alterações, ou substâncias de qualquer tipo?

Estes pontos mostram que o reconhecimento passa por muitos problemas, mas nem todos podem ser aplicados no reconhecimento computadorizado.

Capítulo 3

Sistemas

*“Para aprender algo novo,
volte pelo caminho que você percorreu ontem.”*

Anônimo

Embora a importância da verificação de assinaturas seja grande, a quantidade de trabalhos publicados não é tão numerosa. Isto pode ocorrer pelo alto valor comercial das inovações no campo e com isto, a relutância industrial em tornar público os resultados das pesquisas e descobertas na área. A maioria das companhias que desenvolvem estudos neste campo protegem suas invenções (PenOp, Sign-On, Signer Confidence) mantendo os resultados em segredo dentro da empresa ou patenteando suas inovações. Nestes casos muitos trabalhos deixam de ser publicados nos *journals* [2].

Como já foi comentado neste trabalho, a principal divisão na área dos sistemas de reconhecimento é entre *on-line* e *off-line*. Embora exista um número de estágios intermediários, indicados no Capítulo 2, que podem ser distinguidos, a verificação de assinaturas (e também todos os processos que envolvem escrita) são geralmente polarizados entre receber seus dados *on-line* de uma caneta e/ou mesa especial ligada ao computador (aquisição durante o processo de escrita) e o que verifica a assinatura em *off-line* a partir da imagem apresentada no pedaço de papel onde está escrita. Os

dois processos têm vantagens e desvantagens. No caso de *on-line*, há dificuldades em segmentar as assinaturas relacionadas com os períodos onde a caneta está escrevendo ou não e sobreposições podem ocorrer. Por outro lado, sistemas *off-line* podem ter dificuldades em extrair a assinatura do fundo e contam com um menor número de informações.

3.1 Técnicas Estáticas (Off-line)

Aqui, a partir de um “scanner”, digitaliza-se a imagem da assinatura produzindo uma cópia desta no sistema. Nesta técnica são utilizadas só informações ópticas, o que a torna similar a forma atual de verificação humana (forma canônica) executada no mercado financeiro, onde só a imagem da assinatura é verificada. Como informações visuais são suficientes para executar a tarefa de verificação no mercado acredita-se que os sistemas digitais baseados em estudos da mesma imagem podem produzir resultados similares aos de especialistas humanos. O estudo destes métodos pode ser realizado ao dividí-los pelas espécies de características utilizadas [21].

3.1.1 Características Globais

Os sistemas que utilizam estas características, trabalham com todos os pixels que compõem os traços da assinatura na margem. Elas, como pode-se ver pelo nome, não refletem qualquer propriedade local, geométrica ou topológica. Em seu lugar, utilizam transformações, expansões seriais e análise do gradiente da imagem. Embora sejam facilmente extraíveis e insensíveis ao ruído, são dependentes do alinhamento e influenciadas pela distorção e variação de estilo. Os pesquisadores utilizam estas características baseados em duas idéias [22]:

- a orientação e proporção total de assinaturas genuínas escritas numa área bidimensional são relativamente estáveis para cada escritor;
- a variabilidade local do traço é uma característica do autor e pode ser tomada como tal. Este fenômeno é caracterizado por deslocamentos locais dos traços seguindo o eixo principal da assinatura.

A título de exemplo, Lee e Pan [23] utilizaram a Transformada de Hadamard para apresentar a imagem em um novo domínio onde pode-se proceder a análise do seu espectro; os autores utilizaram esta transformada por preservar a energia e entropia da imagem original. Eles utilizam também um pré-processamento necessário para normalizar a imagem para o processo, consistindo em eliminar os pontos irrelevantes (pontos com altas faixas de variâncias dentro da classe) para produzir conglomerados de assinaturas.

Como outro exemplo, pode-se falar do trabalho de Gomes e Carvalho [1], que utilizam técnicas com “momentum” estatístico para representar as assinaturas como um vetor contendo seis valores de “momentum” de várias ordens, derivados da imagem da assinatura. Como estavam interessados nas taxas de reconhecimento, utilizavam um classificador de distância euclidiana mínima, sem comparar assinaturas falsificadas ou taxas de rejeição.

Outro aspecto desta técnica é capturar um conjunto de atributos localizados na imagem, como tamanho global da assinatura com a inclinação de escrita medida, valores de concavidades e etc. O trabalho de Ammar et al. [10] utilizava como características globais geométricas a altura e largura da assinatura, ângulo de inclinação, centro de gravidade, projeção máxima horizontal e a linha base da assinatura (linha média). Características extraídas da segmentação em zonas verticais e horizontais, também podem ser usadas. Assumindo que a assinatura é conhecida, o que é ver-

dade, pode-se separar o nome do sobrenome e tirar relações entre eles. Histogramas também são válidos, pois podem representar de que forma se distribuem os pixels na imagem.

Pode-se [20] ter um verificador de assinaturas randômicas (definidas como “free-hand”) que aplica uma combinação de parâmetros: razão de dimensão, ângulo de inclinação, relações entre letras e assinatura total, letras grandes e pequenas, etc; assim como construir a base de dados com assinaturas bancárias, implementando técnicas de limiares para segmentação da imagem. Além de localizar as zonas (alta, media e baixa) da assinatura, e a partir delas formar heurísticas. Há uma conglomeração dos valores de características das verdadeiras para se formar um grupo homogêneo. Este tipo de sistema pode encontrar dificuldades com pessoas que não “escrevem” seu nome, preferindo traçados, pois toda ela está elaborada na heurística de zonas e toma algumas considerações particulares. Sua grande informação se refere ao fato de indicar que é possível criar um sistema que detecte todas as falsificações randômicas.

3.1.2 Características Estatísticas

Características Estatísticas são derivadas da distribuição de pixels na assinatura; inicia-se uma verificação com a divisão da imagem em zonas (alta, média e baixa), início e fim de segmentos de traços (similar a grafotécnica) e trabalha cada subimagem extraíndo características da zona e sua proporcionalidade. Esta segmentação é demonstrada na Figura 3.1.

A técnica estatística inclui a extração de fatores de alta pressão com respeito à zonas de segmentação (por exemplo, superior, média e inferior) e a razão da assinatura com a altura. Elas utilizam informações topológicas, dinâmicas e variações de estilo. Entretanto, devido ao grande número de variações associadas com a assi-

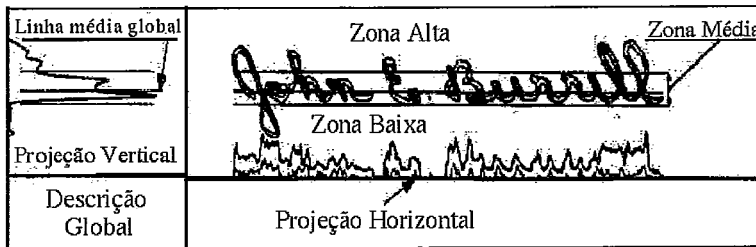


Figura 3.1: Ilustração de como é feita a divisão da assinatura em zonas. Primeiro encontra-se a linha média utilizando a projeção vertical. A partir desta, extrai-se a zona média. As zonas alta e baixa são definidas em consequência desta.

natura, uma contínua investigação a fim de se obter uma extração de características mais poderosas é requerida.

Um exemplo disto pode ser visto no trabalho de Rigoll e Kosmala [24], onde utilizou-se *Hidden Markov Models* (HMMs), alcançando taxas de 92,2% de acerto. O interessante neste artigo, é que o autor notou, com surpresa, que a utilização de uma característica off-line, como a distribuição de pixels, teve resultados superiores a vários outros dados on-line sozinhos, como velocidade e aceleração. Seu sistema final consiste de um método híbrido de características on-line e off-line, aproveitando as características off-line que são mais fáceis de implementar e necessitam de um equipamento menos sensível.

Outro exemplo, é o trabalho de Drouhard *et al.* [17], onde aplicou-se a *Probability Density Function* (PDF) e utilizou uma rede treinada em backpropagation como classificador. Seus resultados chegaram a 5% de FAR, porém utilizaram-se falsificações aleatórias para o teste. Um ponto que vale a pena realçar é que devido à necessidade de um grupo de treinamento com um número bom de exemplos de uma assinatura verdadeira, foram criadas novas assinaturas do grupo inicial digitalizado com um algoritmo que rotacionava as mesmas numa faixa de 2 em 2 graus numa faixa de ± 6 graus de rotação, sintetizando novos modelos.

3.1.3 Características Geométricas

Alguns sistemas utilizam métodos que descrevem características geométricas preservando as propriedades intrínsecas (que não variam entre assinaturas genuínas) e extrínsecas (que variam entre assinaturas genuínas). Os algoritmos que utilizam estas técnicas incluem aproximação polinomial, extração características por convolução e contorno de traço. Esta tem alta tolerância a distorção e variação de estilo, podendo tolerar um certo grau de rotação e translação.

Pode-se utilizar relações geométricas e topológicas a partir da extração de segmentos dos riscos que compõem a assinatura [21]. Este processo tem como finalidade buscar a forma como a mesma é escrita. A proposta segue quatro estágios: o pré-processamento (que inclui a extração do “esqueleto”¹ da assinatura), traçado (transforma o esqueleto bidimensional numa sequência unidimensional de traços, fornecendo algumas características dinâmicas), segmentação de pontos críticos e normalização destes. O processo tem como desvantagem a formação do esqueleto, o que é difícil de conseguir, necessitando utilizar várias heurísticas para não produzir deformações da imagem. A formação dos segmentos da assinatura também não é trivial, e suas regras heurísticas dependem do usuário que escreve (no caso, são baseadas em características da língua inglesa e de manidestros) necessitando desenvolver novas regras para novas linguagens. O método em si, no entanto, permite a captura de informações dinâmicas além das estáticas, o que o torna muito poderoso. Sua multiresolução adiciona a capacidade de variar o nível de detalhes que podem ser extraídos, e a normalização permite uma robustez a escala, translação e rotação. Não apresenta os valores de verificação, demonstrando uma taxa de 97% de reconhecimento, que foi realizado para indicar que o algoritmo tinha uma boa

¹Indica-se como esqueleto de uma assinatura a forma “fina” da mesma, onde todos os traços tem a espessura de um pixel, este termo será melhor abordado no Capítulo 5.

discriminação.

Sabourin [22] elaborou um trabalho que representa a assinatura baseado na sua descrição de forma e características pertencentes ao formato da assinatura. A principal dificuldade do seu processamento está na variabilidade local do traço de escrita, embora utilize o fato de que as assinaturas possuem um formato geral padronizado para a forma. Neste artigo, o formalismo proposto é baseado na percepção visual. A imagem de uma assinatura é centrada num grid de retinas retangulares, e a partir de operações morfológicas, uma amostra da disposição dos pixels é usada para definir uma média da atividade que excita a retina no foco do grid; permitindo medidas de variações locais que podem identificar quem escreve. A originalidade da proposta é que não se trabalha uma área específica, mas uma em qualquer região que possua a atenção no momento, como numa forma atencional que segmenta partes importantes para fazer medidas locais. Avaliou-se o algoritmo usando 800 assinaturas genuínas de 20 indivíduos. Dois tipos de classificadores utilizados foram o *nearest-neighbor* (*NN*) e um classificador de limiar, mostram faixas de erro abaixo 0,02% e 1,0%, respectivamente, no contexto de falsificações randômicas. O autor admite que o *NN* é mais poderoso pois avalia toda a informação da memória, mas que o segundo, para uma implementação real é mais realista, embora descobrir um bom valor para o limiar não seja uma tarefa simples. Assume invariância de rotação e escala, só preocupando-se com a translação.

Uma busca por pontos importantes na assinatura [25] pode ser implementada, modelando um imitador típico de uma assinatura qualquer, pelos passos visuais que dinamicamente ela possui. Isto é obtido por uma segmentação da assinatura para reconstruí-la com um número mínimo de pontos e produzir pedaços fechados para a realidade psicomotora da sua execução (ele separa os traços que compõem a assinatura). Sua atuação é simples: para cada ponto da curva, o algoritmo experi-

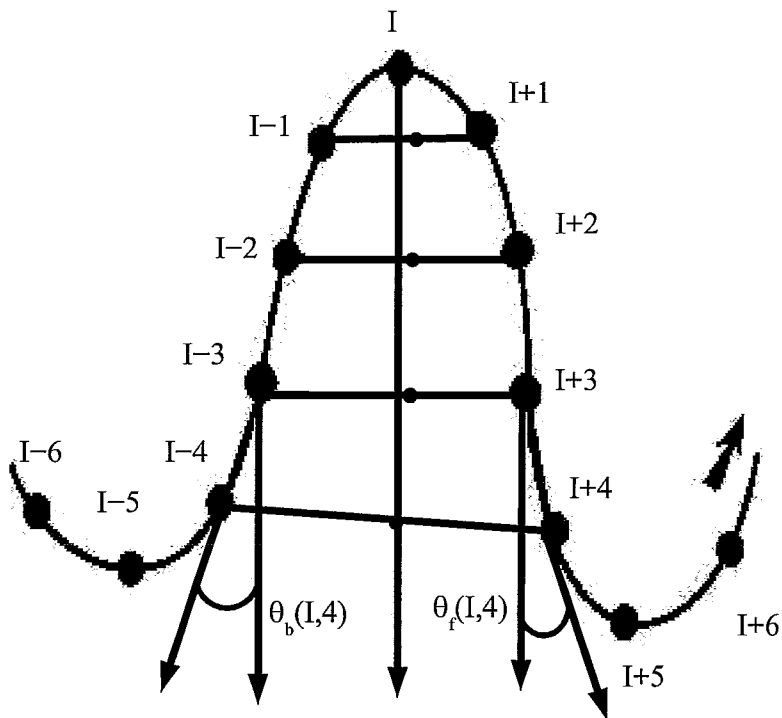


Figura 3.2: Apresenta-se como é segmentada a assinatura em pontos, no caso, a curva mostra um parâmetro, a relação entre o ponto do centróide da curva (i) e outros pontos da curva, usado para determinar o domínio e a importância de i .

menta construir um vértice centrado num ponto que ajude vizinhos de cada lado a cumprir algumas condições encontradas. Parâmetros geométricos são calculados e determina-se os pares de pontos que fazem parte do domínio, refina-se e localiza-se os pontos segmentados (Figura 3.2). O problema deste algoritmo é que o programa segue o traçado e as deformações locais irão fatalmente diferenciar os “pedaços” de traçado. Outro problema é a necessidade de “afinar” o traço da assinatura para uma linha com a espessura de um pixel (“esqueleto” da assinatura). Conseguir reconstruir o traçado não é trivial, por causa de sua grande variabilidade nas assinaturas. Sobreposições de traços podem, por exemplo, deixar o algoritmo confuso.

Propostas que utilizam redes neurais também podem ser utilizadas [18, 26]. Gomes e Filho [1] utiliza uma implementação que usa o “momentum” estatístico, uma característica global, como entrada da rede. A imagem é recomposta em um

conjunto de características de “momentum”, sem aplicação de heurísticas, pois a característica é fornecida pela função da distribuição espacial da disposição dos pixels. Seu pré-processamento envolve apenas equalização de histogramas e filtros digitais para melhorar a qualidade das imagens, sem normalização para orientação ou escala. Faz isto para provar a robustez do método em suportar variações de assinaturas genuínas, embora a normalização diminuiria os valores de discrepâncias dos momentos entre as imagens verdadeiras e falsas. O sistema é utilizado como um reconhecedor por causa da estrutura com que foi implementado, onde o banco de dados possuía 7 diferentes autores e 20 assinaturas de cada; como falsificações foram usadas as outras 6 em relação à testada, e uma saída *winner-take-all* para gerar a resposta. Porém fornece as definições que um método de verificação usando redes deve possuir.

A título de informação, dispomos na Tabela 3.1 de alguns resultados, para fornecer uma idéia dos índices dos métodos off-line.

Autor	Desempenho (%)	Características Utilizadas
Kuckuck [4]	5.0	Fourier e Funções de Autocorrelações
Sabourin [22]	1.0	Operações Morfológicas
Drouhard et al. [17]	5.0	PDF e Backpropagation, ver texto
H. Cardot et al. [18]	3.5	Redes Neurais
Rigoll & Kosmala [24]	3.9	PDF e HMMs, ver texto
Ribeiro & Vasconcelos [26]	22-10	Momentum e CCA

Tabela 3.1: Desempenho de alguns métodos de verificação off-line, fornecidos os índices totais de erros. Note que se existindo dois valores, o mais alto indica o índice de erros para as falsificações especializadas, e a menor para as aleatórias; caso contrário indica o valor de erro para as aleatórias. A abreviação CCA significa Redes “Cascade-Correlation” Auto-associativas. Os valores estão em função da média ponderada de FAR e FRR.

3.2 Técnicas Dinâmicas (On-line)

Neste processo a imagem da assinatura é capturada por uma tabela gráfica (e/ou caneta especial) que armazena a informação (coordenadas x e y relativas no tempo) e possibilita atributos mais avançados como pressão e inclinação da caneta. Geralmente, os dados resultantes do método são relativamente livres de ruído. O volume informações aqui é bem maior; além de ter toda a imagem, que pode ser utilizada “off-line”, outras funções temporais podem ser buscadas. Isto também torna mais difícil a escolha das melhores características que devemos procurar para fazer a comparação. A taxa de erros é, em geral menor, porém, força a existência de um aparelho especial (a mesa digitalizadora e/ou caneta) em todos os pontos onde a assinatura deva ser executada, ao contrário da anterior que pode ser executada nos dias atuais, sem uma grande modificação do sistema. Um outro problema refere-se ao aspecto ergonômico, a caneta deve ter tamanho e peso certos e ser confortável. Os equipamentos eletrônicos utilizados neste sistema também devem considerar a segura e intermodularidade entre a forma de aquisição e aparelhos de verificação.

Existem vários trabalhos na área [2, 18, 4, 27, 28, 29], além de estudos envolvendo os modelos de escrita [30] que tornam-se importantes para compreender melhor a biomecânica dos movimentos durante a assinatura e conseguir descobrir parâmetros melhores para estudos.

Capítulo 4

Filtros

*“As verdades que a ciência revela
superam sempre os sonhos que destrói.”*

Joseph-Ernest Reman

Uma imagem é uma função bi-dimensional $I(x, y)$ de dados definida numa certa região ou, em outras palavras, um plano onde cada ponto possui um valor representativo da sua intensidade luminosa, sendo denominado pixel a menor unidade deste. O processamento de imagens tem como objetivo analisar as mesmas para extrair informações alterando algumas características. Cada tipo de aplicação exige um processamento específico, existindo uma ligação direta entre o tipo de processamento e a aplicação desejada. O processamento visando a informação estrutural da imagem é importante não só para indicar a que classe a imagem pertence, mas também para descrever características da imagem que são recusadas pelas outras classes.

O primeiro passo para o processamento diz respeito à aquisição da cena. Trata-se de representar a informação pictórica num formato adequado para o tratamento computacional. Existem vários dispositivos associados às inúmeras aplicações para captar a imagem; como exemplo, temos o ultra-som usado para analisar o interior do corpo humano, e o “scanner” para digitalizar documentos e imagens.

Posteriormente a esta captura e representação computacional, são descritos procedimentos de transformação, capazes de realçar características relevantes. Alguns exemplos, descritos por Craizer [31], podem ser vistos como:

- **Realce:** Quando há interesse em melhorar a aparência da imagem para análise visual, ou executar uma transformação na mesma para uma representação onde sejam evidenciadas algumas características.
- **Mudança de Resolução:** Quando é necessário trabalhar com uma resolução diferente da original (em número de pixels ou cores).
- **Restauração:** Quando o método de aquisição gera uma deterioração é preciso utilizar algoritmos para restaurar as imagens.
- **Compressão:** Como as imagens requerem um número grande de bits para serem guardadas, é necessário encontrar meios de representá-las usando o mínimo de memória possível sem prejuízo de qualidade.

Estamos interessados, neste trabalho, no reconhecimento de assinaturas, por isso, utilizaremos uma forma de realce, onde mediante filtros orientados, que apresentaremos adiante, um sentido de traçado de escrita é enfatizado em relação aos outros de acordo com o filtro utilizado. Isto é realizado porque os seres humanos, para compreensão de imagens utilizam noções de textura, borda e posição, e ao segmentar a imagem, pode-se apresentar ao computador uma nova dimensão de características que indicam como estão distribuídos os traços permitindo aos computadores o reconhecimento da disposição da escrita na região.

4.1 Modelo de Imagem

Matematicamente pode-se definir imagem como uma função $I(x, y)$, bidimensional, definida numa certa região. Para a maioria das imagens, a região de definição é um subconjunto limitado do plano, e os valores assumidos pela função são números reais limitados e não-negativos [32].

Por comodidade, assume-se que a imagem é definida como um retângulo $[0, r] \times [0, s]$, e os valores tomados estão contidos no intervalo $[0, t]$. Ao valor $I(x, y)$ da imagem no ponto (x, y) dá-se o nome de nível cinza.

Uma imagem digital ou discreta é apresentada como uma função assumindo valores quaisquer definidos por duas variáveis inteiras (coordenadas) como uma grade regular de pontos da forma $(m.dx, n.dy)$; onde dx e dy são os intervalos nas direções x e y ; e os valores tomados estão contidos em $[0, M - 1]$ e $[0, N - 1]$. Aqui, cada ponto define um pixel (menor unidade da imagem) e dx e dy definem as distâncias entre cada pixel. Os valores que estes pontos assumem são também múltiplos de uma dada quantidade dz que varia em um intervalo $[0, k - 1]$, definindo o conjunto dos valores de níveis de cinza¹ que uma imagem possui. A opção por um número elevado de níveis cinza, acima do necessário, representa uma elevação do tempo de execução e aumento do volume da memória de armazenamento. Já ao optar-se por um nível reduzido pode permitir a ocorrência de falsos contornos, pelo desnível entre os elementos. Dispor a imagem em função de duas variáveis significa que mapearemos cada ponto do espaço euclidiano em coordenadas cartesianas.

A imagem digital pode ser obtida de uma imagem não digital (“contínua”) através de um processo que envolve dois passos: um, denominado “amostragem”, que discretiza o domínio de definição da imagem (o quanto da área da imagem irá

¹Consideramos que a imagens aqui tratadas não são coloridas.

se transformar em um pixel); e quantização, que consiste em escolher os valores de intensidade de cada ponto (o índice de luminância de um ponto da imagem é restringido para um intervalo discreto). Os computadores utilizam o modelo discreto para representar as imagens por seus limites físicos de armazenagem e processamento.

No caso destes pontos estarem dispostos na forma de um retângulo, com M linhas e N colunas, essa função pode ser representada por uma matriz de pixels, gerando um mapeamento das intensidades de cada ponto do espaço, expresso em coordenadas Cartesianas se cada membro desta matriz guarda uma relação espacial direta com um ponto da assinatura captada, registrando sua intensidade quantificada.

Quando ocorre de só existir dois níveis de intensidade luminosa que a imagem pode assumir, esta é dita binária. Estas têm grande importância por serem fáceis de ser obtidas e ocupar menos espaço. Embora possuam uma quantidade de informação menor, há uma grande distinção entre o objeto e o fundo.

Um conceito útil é também o de vizinhança, dado um ponto $p = (x, y)$ da imagem, os “vizinhos” de p , ou a “vizinhança de p ”, são os pontos mais próximos de p . Em particular, a “vizinhança-4” é o conjunto dos quatro pontos $(x - 1, y)$, $(x + 1, y)$, $(x, y - 1)$ e $(x, y + 1)$ e a “vizinhança-8”, o conjunto dos pontos da vizinhança-4 mais $(x - 1, y - 1)$, $(x - 1, y + 1)$, $(x + 1, y - 1)$ e $(x + 1, y + 1)$.

Na Figura 4.1 uma parte da imagem foi segmentada e representada em forma de matriz. Note pelos valores assumidos que a imagem é binária, ou seja, são atribuídos só dois valores a seus pixels. As imagens de assinaturas utilizadas nesta tese seguem um formato onde cada pixel tem seu valor variando de 255.0 (nível branco) a 0.0 (nível preto), permitindo uma excelente representação da imagem. Salienta-se aqui que no trabalho, as assinaturas foram processadas num formato

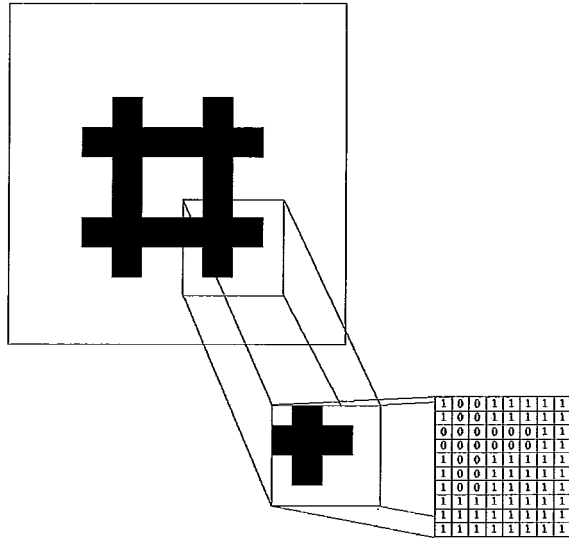


Figura 4.1: Ilustração de uma Imagem Binária, note a matriz que apresenta os valores de cada pixel da imagem

binário, estabelecendo o branco (nível mais alto) como fundo e o preto (nível mais baixo) como o objeto.

4.2 Filtros Lineares

Os Filtros de convolução são muito usados no processamento de imagens para extrair informações e permitir transformações. Utiliza-se a denominação Filtros porque derivam do campo de processamento de sinais [33]. Como os sinais tratados aqui são estáticos, é preciso tomar cuidado com estes conceitos, para não existir confusão dos filtros aqui definidos, com os filtros dinâmicos, aplicados no processamento de voz, por exemplo. O aspecto mais importante, e que define como os filtros são utilizados, é a convolução digital, que similarmente à convolução matemática, é definida como a passagem de uma função (o filtro) sobre cada valor de uma outra (a imagem) produzido como resultado uma terceira (a imagem filtrada). A convolução em si, é uma operação na qual o nível cinza de um ponto da imagem transformada é uma combinação linear dos níveis de toda a imagem de entrada e onde os coeficientes da combinação dependem das posições relativas entre o ponto e os demais pontos

da imagem, e do peso associado à esta distância, fornecido pelo filtro.

4.2.1 Convolução Digital e Máscara

Uma boa descrição de convolução digital é fornecida por Niblack [34]. Considerando $f(x)$ a entrada e $g(x)$ a saída, temos S (um filtro qualquer) como um operador que atua sobre $f(x)$ gerando uma nova função $g(x)$:

$$f(x) \circledast S \implies g(x)$$

Sendo o filtro S linear e invariante ao deslocamento, existe uma função h tal que:

$$g(x) = \int_{-\infty}^{+\infty} f(t)h(x-t)dt \quad (4.1)$$

onde $h(t)$ é uma função que caracteriza o filtro completamente, sendo a resposta ao impulso unitário do filtro. Uma convolução é uma operação na qual o nível de cinza de um ponto da imagem gerada é uma combinação dos níveis de todos os pontos da imagem de entrada. A saída para um ponto i é dada por uma soma de pesos multiplicados pelos pixels da entrada ao redor de i (pesos fornecidos por $h(t)$). Para computar $i+1$, $h(t)$ é deslocada em 1 e a soma é recalculada. No caso digital a integral se torna uma soma, pela discretização do espaço:

$$g(i) = \sum_{k=-\infty}^{+\infty} f(k)h(i-k) \quad (4.2)$$

Embora os limites sejam infinitos, a função é zero fora de uma faixa, assim, minimiza-se a computação determinando-se os limites desta como $[-w, w]$:

$$g(i) = \sum_{k=i-w}^{i+w} f(k)h(i-k) \quad (4.3)$$

Para a convolução sobre imagens, é mais fácil atuar em duas dimensões, além de ser mais claro o efeito que desejamos produzir, definindo uma soma dupla dada por:

$$g(i, j) = \sum_{k=i-w}^{i+w} \sum_{l=j-v}^{l+v} f(k, l)h(i-k, j-l) \quad (4.4)$$

aqui sobre uma imagem $f(i, j)$, a operação de um filtro gera uma nova imagem $g(i, j)$ determinada pela Fórmula 4.4, esta é a forma discreta da operação analítica de convolução. Tem-se como dimensões do filtro $[2w + 1, 2v + 1]$ e a matriz que contém os valores $h(m, n)$ é a máscara do filtro. A fim de otimizar o algoritmo, evitando que as operações não fiquem lentas, é preferível utilizar máscaras de tamanho pequeno, ou seja, deve-se evitar valores de w e v altos.

Observe que:

- a função que caracteriza o filtro é expressada na forma de um operador matricial de somatório de vizinhanças, que terão uma distribuição de coeficientes com pesos determinados pelo filtro desejado.
- O tamanho da vizinhança é determinado pelo pesquisador, onde emprega-se muito a vizinhança imediata das oito células vizinhas, mas não é necessário que sejam utilizadas somente estas pois, dependendo do tamanho do filtro, pode-se estabelecer uma área maior de atuação.

4.3 Modelos de Filtros

A fim de demonstrar como o horizonte dos filtros é vasto, apresentaremos alguns exemplos simples existentes na área. Como já dissemos, uma imagem pode ser vista como uma função. Com isto, ela possui, características que podem ser vistas como compostas por frequências sendo as altas representadas por variações abruptas do nível da imagem e as baixas pelos níveis contínuos na mesma. Desta forma, pode-se dividir os filtros que trabalham sobre estas em duas categorias, ao tomar-se como discriminador sua características de resposta:

- passa-alta, indicador de bordas

- passa-baixa, de eliminação de ruído

De acordo com seu propósito, o filtro pode ser escolhido, demonstrando uma vantagem da convolução: servir uma vasta faixa de objetivos, mediante a simples troca do operador. Um filtro simples mostrado na Equação 4.5, é o de média; é um filtro passa-baixa que tem como objetivo eliminar os sinais de alta frequência da imagem, e utiliza a “vizinhança-8”.

$$h(i, j) = \begin{array}{|c|c|c|} \hline 1/9 & 1/9 & 1/9 \\ \hline 1/9 & 1/9 & 1/9 \\ \hline 1/9 & 1/9 & 1/9 \\ \hline \end{array} \quad (4.5)$$

Aqui pode-se ver claramente que na filtragem, o nível cinza original do ponto p , após a transformação, depende do valor do nível de cinza original do ponto e de outros pontos da vizinhança de p . Esta vizinhança não só é necessariamente limitada, mas em geral os pontos mais próximos de p contribuem mais para o novo nível do que os pontos mais afastados. Na realidade, temos uma média ponderada dos pontos de I , onde os pesos são dados por $h(i, j)$.

O efeito visual de um filtro passa-baixa é o de “suavização” da imagem, uma vez que as altas frequências são atenuadas, assim como diminuem os contrastes. Para os passa-alta, o efeito é inverso, acentua-se transições entre regiões de diferentes contrastes tornando-as mais nítidas, tendo como efeito indesejado enfatizar os sinais de alta frequência que porventura exista na imagem. Passaremos a falar somente dos filtros passa-alta por que são os utilizados nesta tese.

4.3.1 Filtros Detectores de Bordas

Conhece-se várias características que uma imagem possui. Uma delas é que, em geral, ocorre uma mudança maior ou menos abrupta no nível cinza, cor ou textura na fronteira entre objetos e fundo, permitindo detectar a presença ou ausência de

limites de um objeto pela busca de contornos no mesmo (bordas). Estes limites se tornam importantes na busca do reconhecimento do objeto, pois o segmentam, facilitando a comparação.

Quando existe muito ruído, operações simples podem não conseguir detectar bordas. Porém, bordas com orientações conhecidas podem ser melhor detectados usando um “padrão de bordas”. Este, como exemplo, pode ser visto se procurarmos por retas ou bordas verticais: a partir da soma dos níveis de cinza em duas faixas paralelas adjacentes, e subtrair os resultados. A magnitude da diferença será alta se uma reta vertical estiver presente.

A busca de bordas com orientações conhecidas é bem estudada em filtros, onde vários operadores existem para extrair o gradiente $G(x, y)$ de uma superfície $Z = f(x, y)$ para qualquer ponto (x_0, y_0) , ele define o plano tangente à superfície nos pontos. A fim de extrair o gradiente de um ponto, vários operadores de filtros podem ser sugeridos e aplicados de acordo com os objetivos que se deseja alcançar. Os dois apresentados abaixo são de Prewitt (4.6) e Roberts (4.7).

$$h(i, j) = \begin{array}{|c|c|} \hline 0 & 1 \\ \hline -1 & 0 \\ \hline \end{array} \quad (4.6)$$

$$h(i, j) = \begin{array}{|c|c|c|} \hline -1 & 0 & 1 \\ \hline -1 & 0 & 1 \\ \hline -1 & 0 & 1 \\ \hline \end{array} \quad (4.7)$$

Os dois filtros podem acentuar variações da intensidade de uma imagem capturando as combinações de peso das intensidades dos pixels. Ao contrário dos operadores de passa-baixa onde todos os pesos são positivos, estes apresentam valores de pesos opostos (mas cuja soma é igual a zero). Com isto, numa região de intensidade constante, este operador de diferenças deve ter como saída zero pois não há variação de intensidade dentro de uma região com intensidade constante.

Um filtro passa baixa teria como valor uma constante, visto que a média de uma constante é a constante.

Prewitt [35] descreveu também uma alternativa para a técnica de estimar o gradiente. Ela seguiria um processo onde aplica-se um conjunto de operadores diferenciais com cada um fornecendo uma estimativa (escalar) da magnitude do gradiente e o operador correspondente fornece uma estimativa da direção do gradiente. Quatro operadores possíveis de serem utilizados são:

$$\begin{array}{|c|c|c|} \hline -1 & 1 & 1 \\ \hline -1 & -2 & 1 \\ \hline -1 & 1 & 1 \\ \hline \end{array} \quad (4.8)$$

$$\begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline -1 & -2 & 1 \\ \hline -1 & -1 & 1 \\ \hline \end{array} \quad (4.9)$$

$$\begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & -2 & 1 \\ \hline -1 & -1 & -1 \\ \hline \end{array} \quad (4.10)$$

$$\begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & -2 & -1 \\ \hline 1 & -1 & -1 \\ \hline \end{array} \quad (4.11)$$

4.3.2 Filtros e a Visão Humana

O caminho da percepção visual é longo entre a captura da imagem pelo olho até seu reconhecimento. Uma etapa deste percurso se realiza na retina, formada por várias células que transformam o estímulo luminoso em potenciais de ação. A entrada de cada uma é composta de fotorreceptores que atuam numa área específica da retina, o *campo receptor* (RF), sendo este definido como a região do espaço que uma célula monitora, na qual um estímulo visual habilita ou suprime a atividade neural.

Embora a forma precisa da função de sensibilidade das células varie com o estágio do sistema visual, no córtex visual primário o RF exibe duas bandas (ou regiões)

antagônicas paralelas de inibição e excitação (definiremos aqui a área de excitação como S^{on} e de inibição como S^{off}), sendo mais preciso dizer que a função de sensibilidade descreve a resposta da célula a entrada de cada ponto do espaço. Em vários modelos de visão o RF é definido como um filtro linear com saída dada pelo somatório das intensidades luminosas de diferentes pontos do campo visual:

$$g(i, j) = \sum_{k=i-w}^{i+w} \sum_{l=j-v}^{j+v} f(k, l)h(i - k, j - l) \quad (4.12)$$

sendo, $f(k, l)$ a distribuição de luz sobre o RF (imagem), e $h(i - k, j - l)$ a função peso (varia de acordo com a célula modelada) que descreve convenientemente a sensibilidade do RF 's no espaço (onde k e l demonstram as duas dimensões espaciais). A partir disto, transições de luminância na imagem de entrada são descobertas, enquanto regiões uniformes não produzem resposta. A saída $g(x, y)$ é a imagem gerada como resposta desta convolução.

No sistema visual, as células corticais apresentam propriedades do RF ligadas ao alongamento sobre um de seus eixos. Com já dito, estes RF alongados exibem regiões espacialmente segregadas de excitação e inibição, sendo que, o alongamento reduz o número de orientações para as quais a célula irá responder, apresentando uma seletividade de orientação (Figura 4.2). O grau desta seletividade será determinado pelo quão alongada é a área que o RF atua no eixo principal, em relação ao eixo ortogonal. Isto pode ser visto se tomarmos a razão das dimensões de largura e comprimento do RF de uma célula, se este for 1:1 temos uma célula não orientada, para a razão for 1:3 teremos uma mediantemente orientada, permitindo-se chegar a células corticais de 1:5. No nosso modelo simulamos e aplicamos estas “células” que são sensíveis a orientação do contraste num sentido definido, para extrair informação de bordas orientadas. Uma visão espacial destes filtros é demonstrada na Figura 4.4.

Devido ao arranjo espacial de suas regiões excitatórias e inibitórias, esta células

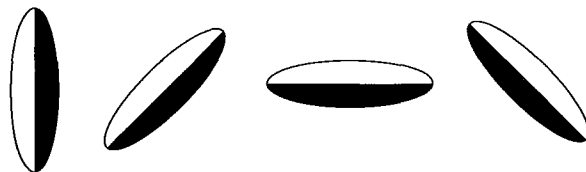


Figura 4.2: Considere esta imagem como a vista superior dos filtros, a parte escura representa o vale e a clara o pico, estes filtros irão responder a bordas nos sentidos: vertical, 45° , horizontal e 135° .

são particularmente relacionadas com as faixas orientadas de uma dada polaridade; isto quer dizer que algumas células podem responder vigorosamente para as mudanças de área clara para uma de menor luminância de fundo (transição branco-preto - S_n^{ce}), enquanto outras fortemente para uma variação da faixa escura para a clara (transições preto-branco - S_n^{ec}). Estas atribuições dependem das características da imagem processada, como a tratada neste trabalho possui valores altos para a faixa clara e baixos para a escura, sua disposição segue esta forma, uma imagem com representação de espectro inversa, terá definições inversas também com relação a estes filtros. Note que o índice n disposto serve para discriminar a orientação do filtro. De qualquer modo a célula é sensível ao preciso posicionamento do estímulo e RF, sendo a intensidade da resposta obtida proporcional a disposição do estímulo no RF. Deve-se notar que as células não são lineares, porque as respostas das células são lineares só acima de um limiar, sendo sua preformance tal que saídas negativas não são geradas.

É importante salientar que esta célula, que denominaremos *célula simples*, não pode ser vista como um simples detector de borda. Como a saída ótima só ocorre para uma determinada orientação, ela precisa estar com suas regiões de excitação e inibição sincronizadas com a disposição do contraste, sendo mais apropriado dizer que elas são impulsionadas a estimar diferentes orientações de contraste para cada localização. Esta sua especialização a um só contraste também impede que ela responda as transições de mesma direção mas de sentidos opostos e, para detectarmos

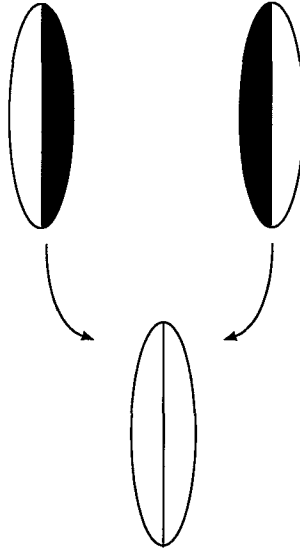


Figura 4.3: A soma das respostas de dois filtros orientados de mesmo sentido e faixas *on-off* opostas determina um filtro complexo.

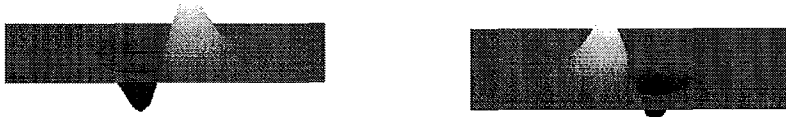


Figura 4.4: Dois Filtros Orientados, o a esquerda definido para captar transições escuro-claro S^{ec} e a direita para claro-escuro S^{ce} ; note que eles irão processar imagens que possuam um nível alto para o claro e baixo para o escuro

bordas com uma orientação particular independente do sentido de luminosidade, deve-se compor duas células simples formando uma célula composta (C_n):

$$C_n = S_n^{ce} + S_n^{ec} \quad (4.13)$$

veja a Figura 4.3.

Para construir um mapa de contornos invariante em relação a direção destes deve-se proceder a soma das respostas de todas as orientações das células complexas aplicadas sobre um ponto:

$$T = \sum_k C_k \quad (4.14)$$

4.3.3 Formalização

A organização não-isotrópica (não simétricos rotacionalmente) das células simples podem ser modeladas de várias formas, como por funções de Gabor ou diferenças de Gaussianas alongadas *ON – OFF*, sendo esta última a forma escolhida neste trabalho devido à sua facilidade de implementação.

Uma RF pode ser convenientemente descrita por uma função peso $S(x, y)$ que define a sensibilidade do RF no espaço (x e y são os eixos dimensionais) [36] demonstrando a resposta da célula à entrada de cada ponto do espaço (mais inibição, ou mais excitação, por exemplo). Esta função deverá reduzir o “peso” de resposta na razão da distância do ponto tratado, determinando que a área periférica tenha uma importância menor, permitindo-se indicar mais força a pixels mais próximos da origem que contribuem com um valor maior a soma do que os distantes, uma característica das distribuições Gaussianas. Desta maneira, respostas de células simples podem ser obtidas pela subtração de duas gaussianas alongadas deslocadas, que servem como descritores convenientes do comportamento celular do ponto de vista matemático, para gerar cada RF. Esta união de duas gaussianas iguais permite, ao se equacionar áreas similares, que quando todos os pontos do RF estiverem ativos não exista sinal de saída gerado pela célula e que as contribuições das posições *ON – OFF* são iguais.

Uma função Gaussiana alongada é definida pela equação

$$G(i, j) = \frac{1}{2\pi\sigma_i\sigma_j} \exp\left(\frac{-1}{2} \left(\frac{(i-k)^2}{\sigma_i^2} + \frac{j^2}{\sigma_j^2}\right)\right) \quad (4.15)$$

onde $G(i, j)$ é a função Gaussiana com largura σ_j para o eixo ordenado e σ_i para o eixo das abscissas, o valor de k fornece a coordenada do centro da Gaussiana $(k, 0)$. Os sigmas são importantes não só para determinar o tamanho do filtro, mas por extensão, indicar o nível de sensibilidade desejado. Isto porque cada filtro

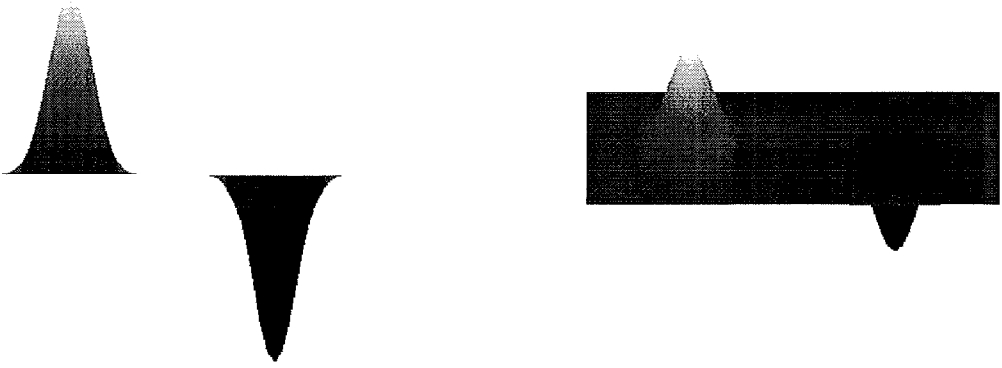


Figura 4.5: A esquerda duas Gaussianas são demonstradas, e a direita, temos uma imagem destas no espaço onde pode-se visualizar que com o deslocamento, elas irão formar as Figuras 4.4

é particularmente parametrizado pela frequência espacial definida pelos σ 's, tendo um tamanho e orientação próprio. Como nosso objetivo é formar um filtro alongado com regiões antagônicas, ao utilizarmos $k = \sigma_i$ para uma gaussiana (S^1) e $k = -\sigma_i$ para uma outra (S^2), teremos duas gaussianas idênticas, de orientação vertical (ver figura 4.5), porém deslocadas em relação ao eixo ordenado. Ao subtrair-se as duas se obtém o filtro sensível ao contraste claro-escuro:

$$F^{ce} = S^1 - S^2 \quad (4.16)$$

A equação acima torna-se

$$F^{ce} = S^{on} + S^{off} \quad (4.17)$$

ao considerarmos que $S^{on} = S^1$ e $S^{off} = -S^2$.

Na Figura 4.5 podemos ver S^{on} e S^{off} como a Gaussiana e sua inversa. A imagem é demonstrada com as duas gaussianas separadas para fornecer uma ideia do seu formato. Nestas, o valor de k é muito superior a σ_i ; para valores menores, imagine que as duas se deslocam em direção do centro, surgindo a composição que forma o filtro, visto na Figura 4.4.

Já possuímos o filtro inicial para extrair bordas perpendiculares a disposição horizontal, para outras disposições, faz-se necessário rotacionar o mesmo, orientando-o



Figura 4.6: Aqui pode-se visualizar o resultado da passagem dos filtros da Figura 4.2 sobre as assinaturas apresentadas, seguindo a mesma ordem, vertical, 45°, horizontal e 135°. A representação da resposta segue uma convenção onde o preto indica a resposta zero e o branco a resposta máxima. O filtro da direita tem maior dimensão que o da esquerda, gerando uma maior área de resposta.

como vemos na Figura 4.2. A utilização destes como operadores na imagem, produzirá os resultados vistos na Figura 4.6. Note que para a assinatura a esquerda possui um ângulo de escrita, e para o filtro que tem a mesma orientação deste (45°), existe uma maior resposta de bordas e para o de orientação oposta (135°), a resposta é bem menor. A direita verificamos melhor como é a representação gerada pelos filtros. O traço quase horizontal, na base da assinatura é quase eliminado no filtro vertical, enquanto é bem salientado no horizontal.

A sensibilidade aos contrastes foi utilizada nesta tese para permitir gerar uma informação sobre a orientação do traçado que percorre cada ponto da assinatura de uma pessoa. Isto ocorre porque o traço produz uma região de borda na imagem e pode-se aplicar o filtro para extrair esta informação, produzindo-se imagens que demonstram todos os traços da assinatura de uma orientação particular, tornando-os mais discerníveis para o computador. No caso do estudo de assinaturas, as orientações do traçado da escrita são um dos aspectos importante que os especialistas

humanos utilizam na verificação.

4.3.4 Jets

Até este ponto só nos preocupamos com a disposição geral do RF célula, mas outro aspecto importante a ser comentado é o tamanho da região sobre a qual este é aplicado. Há células com diferentes tamanhos de RF , mas com seletividade similar para uma mesma região do espaço visual, processando-a para diversas resoluções [36], o que permite uma informação visual melhor, na medida que utiliza um conjunto maior de informações que uma resposta simples.

Os operadores de múltiplos tamanhos, ou, também denominados, processadores de múltiplas escalas, são importantes para representar as várias características da imagem, especialmente aquelas relacionadas a textura, que podem ser descritas para uma variedade de dimensões. Numa escala fina (alta resolução), aspectos locais tornam-se importantes, para uma escala mais alta (baixa resolução) só se percebe formas grossas, tendo um menor nível de resolução para detalhes, embora possua uma maior uniformidade, isto é os valores de uma escala maior atribuídos a um pixel não diferem muito dos valores atribuídos aos pixels vizinhos. Esta variação das escalas é conseguida pelo valor dos sigmas das Gaussianas e seu objetivo pode ser visto como uma busca de informações estruturais. Um traço mais fino terá uma resposta diferente para dimensões de filtro de diversos tamanhos; em uma escala pequena (ou seja utilizando um filtro de dimensões pequenas), tem-se o valor do filtro resultante somente do traço, porém para escalas maiores (filtro de dimensões superiores), teremos a informação da área em torno do traço fornecendo também pesos para a saída que indicará um conjunto maior de informações. Assim, uma combinação de filtros é útil para descrever, na forma de vários valores, o espaço que cerca um pixel, apresentando vários níveis de discernibilidade de detalhes do

ponto estudado. Um algoritmo que consegue utilizar a combinação da saída de vários operadores para múltiplas escalas de uma forma coerente e sistemática possui uma estrutura mais poderosa que aqueles que se utilizam de um só filtro [35].

Uma forma de representação de um conjunto de propriedades de um ponto é o “Jet” (J). Em um grande número de trabalhos [6, 37] ele é apresentado como um vetor de valores formado a partir da passagem dos filtros descritos na Seção 4.3, com múltiplos tamanhos e orientações sobre um ponto definido da imagem de onde deseja-se extrair informação, sendo descrito como uma estrutura vetorial:

$$J_x^I = \{J_{x_\alpha}^I \mid \alpha = 1, \dots, p\} \quad (4.18)$$

onde p é o tamanho total do vetor, o índice I é utilizado para indicar que o *Jet* pertence ao modelo, e x discrimina o *Jet* entre os vários que compõem o gráfico que apresentamos mais a frente. A região em torno de um pixel é descrita por estas respostas, e o *Jet* é robusto à pequenas variações na aparência. O *Jet* pode ser adaptado para absorver mudanças nos sentidos dos traços devido a rotação o que é realizado ao permutar-se os valores das respostas dos filtros. Para cada orientação de traço existe um filtro que responderá mais fortemente e conseqüentemente, o filtro com orientação perpendicular a este terá um valor baixo. Se ocorrer uma rotação na imagem, esta disposição será modificada, sendo a nova disposição similar a antiga, com os valores deslocados. Assim, o ângulo de rotação pode ser anulado por uma transformação linear entre os componentes do mesmo [48].

A Fig 4.7 apresenta uma idéia das relações dos tamanhos e orientações dos filtros aplicados, apresentando 3 tamanhos de filtros com 3 orientações diferentes.

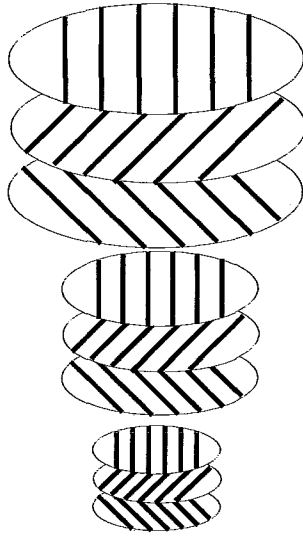


Figura 4.7: Um jet formado por respostas às três dimensões e três orientações de filtros diferentes. A variação de diâmetro do círculo indica o tamanho da área sobre a qual o filtro atua. Os traços indicam a orientação de cada filtro e, para cada uma das três dimensões de filtro representadas, existem três orientações aplicadas: vertical, 45° e 135° . Eles representam informações de frequência do nível cinza e sentido da distribuição das bordas centradas para uma posição de pixel, que são armazenados em um vetor J_x^I .

Capítulo 5

Comparações Elásticas

*“Um contrato verbal não vale
a tinta com que é assinado.”*

Samuel Goldwyn

O mundo visto pelo olho humano é rico e heterogêneo. Imagens observadas no curso do tempo variam significativamente e são raramente idênticas, ainda que exista uma considerável semelhança entre as mesmas. Perceber e determinar a similaridade entre duas formas é a chave para a interpretação de imagens, sendo uma tarefa difícil ensinar a máquina como reorganizar e classificar objetos, tendo que reconhecê-los sobre um grande número de distorções que ocorrem devido aos vários pontos de vista, articulações e deformações em domínios onde as formas não apresentam transformações rígidas.

As propriedades exatas do julgamento humano de similaridade, o qual revela um grande poder de comparar formas sob estas distorções, não são tão bem conhecidas ainda para permitir que sejam modelados, sendo uma tarefa conceitual difícil decidir o nível de detalhes necessários para se modelar os sistemas e as abstrações indispensáveis para capturar propriedades essenciais das funções da mente e por isso, cada vez mais criam-se métodos que buscam representar o processamento de reconhecimento mental. Um método eficiente de comparação pode ser a *Comparação*

Elástica (Elastic Matching - EM), que avalia a similaridade de duas formas como uma soma de deformações locais necessárias para transformar um objeto em outro [38]. Esta captura de propriedades locais e globais, mesmo com deformações não rígidas é a alma do processo de comparações elásticas para medir a similaridade. Ao permitir deformações não rígidas nos objetos comparados, ela fornece uma correspondência entre pontos (ou partes) dos mesmos que deve minimizar uma função do custo de comparação.

Um problema de reconhecimento de objetos pode ser tratado como uma busca de correspondência visual. Fornecidas duas imagens do mesmo objeto deve-se decidir quais pares de pontos são correspondentes entre si e também ser capaz de determinar quais pontos não tem correspondência entre as imagens [39]. No caso presente de verificação de assinaturas, alguns aspectos destes também se aplicam, pois deseja-se reconhecê-la, descobrindo se os pontos se correlacionam com um modelo armazenado, e esta imagem de teste pode não apresentar uma correspondência total entre seus pontos e os do modelo. Para isto, EM oferece uma ótima oportunidade de perceber que pontos da assinatura modelo podem estar representados na assinatura testada e, se este número for alto, teremos uma assinatura verdadeira, caso contrário, será falsa.

Deve-se notar que é um ponto vital para a resolução do problema é desenvolver a função que avalie a similaridade da assinatura modelo com a teste e determinar quais características devem ser codificadas por esta, e como ela irá atuar; com isto seremos capazes de controlar os aspectos do processo que são importantes para o objetivo desejado. Por isso, os dois componentes principais do método de EM são a sua função de custo e o seu mecanismo de busca.

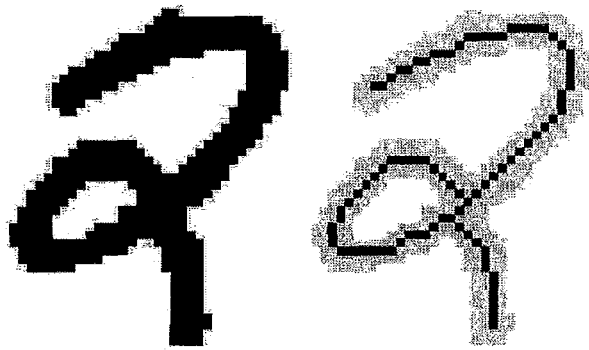


Figura 5.1: “Esqueleto” de uma assinatura. Métodos de afinamento trabalham imagens para apresentá-las de forma que cada linha só possua um pixel.

5.1 Uma Revisão de Comparações Elásticas

Comparações Elásticas (EM) é um algoritmo derivado da técnica de programação dinâmica usado para problemas de comparação de vetores unidimensionais, onde avalia-se o grau de similaridade entre duas sequências dos mesmos pelo uso de uma função de custo associada com operações elementares tais como deleção, inserção e substituição [40]. Com isso, em geral, estudos de EM sobre imagens são feitos julgando não o objeto, mas sim os seus contornos ou esqueletos, pois estes poderão ser tratados como padrões unidimensionais (a partir de conversões de dimensionalidade). Em Basri [38] temos uma boa visão geral sobre o assunto, além de ter-se várias propostas de funções de custo e diversidades de métodos.

Os primeiros trabalhos com EM em imagens bidimensionais utilizavam padrões de linha, que são imagens contendo somente linhas pretas (ou brancas) sobre um fundo branco (ou preto), devido ao fato que como o número de pixels no fundo é maior que os na linha, apresenta-se um código de representação mais compacto [41] colocando a informação num formato unidimensional, além de gerar-se um vetor que contendo o padrão da imagem poderia ser comparado com as fórmulas de comparação dinâmica [42]. Assim, estudos para reconhecimento de números, por

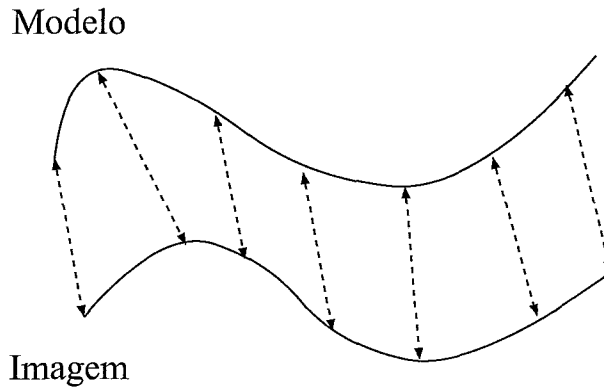


Figura 5.2: Em comparações elásticas de uma dimensão procura-se estabelecer uma hipotética correspondência entre pontos de duas curvas. A linha superior é um modelo armazenado e a inferior uma imagem testada. Um custo de deformação é definido para os pares correspondentes, de acordo com a deformação necessária para “casá-los”. As setas tracejadas indicam os pontos “espelhados”.

exemplo, primeiro utilizam algoritmos de afinamento, onde extraem o esqueleto do número manuscrito para então usar nos contornos deste o EM; a Figura 5.1 demonstra este processo de “afinar”. Este tipo de comparação permite formar uma correspondência entre pontos da imagem testada e pontos do modelo, mesmo que as curvas (ou contornos) tenham tamanhos diferentes, permitindo uma medida de similaridade entre aspectos locais. Veja na Figura 5.2 a ilustração desta correspondência, e note a tentativa de “espelhar” os pontos do modelo em pontos da imagem.

O método tem sido usado em diversos reconhecedores de letras. Tappert [42] descreve um sistema onde o EM é usado num reconhecedor de segmentos. Características de segmentos são retirados para o reconhecimento, onde são agrupados juntos em caracteres. EM é utilizado para identificar quais aspectos da letra estão presentes. Algumas hipóteses são estabelecidas e as melhores são enviadas para confirmação.

Basicamente, a comparação elástica mede a diferença entre duas formas baseada na correspondência específica entre pontos das mesmas. Uma função de custo verifica a similaridade de pontos comparados nas duas curvas com base nas propriedades

locais destes, tais como distância entre eles, diferença na tangência e curvatura do contorno destes pontos. A soma total da similaridade dos contornos é então medida como a soma destes custos locais. Tipicamente a função de custo é usada para encontrar a correspondência entre curvas; isto é, alguma busca é executada para minimizar este custo. Esta comparação avalia a relação entre vários pontos de dois contornos, mesmo entre curvas de tamanhos diferentes.

Vejam melhor isto: visualize dois vetores \vec{V}_x e \vec{V}_y , que representam traços. O vetor \vec{V}_x possui n componentes e \vec{V}_y , p . Consideremos \vec{V}_x como um modelo e \vec{V}_y como um traço qualquer que desejamos saber o quanto se assemelha ao modelo. Uma comparação simples entre os mesmos verificaria, para cada componente i de \vec{V}_x , sua diferença do i -ésimo componente de \vec{V}_y . Isto não poderia ser realizado porque, visto que as dimensões dos vetores são diferentes, existiriam pontos que não possuiriam correspondência. Para evitar isto, utiliza-se uma comparação onde o elemento i de \vec{V}_x não é diretamente associado com o elemento correspondente em \vec{V}_y , mas sim que o elemento mais semelhante a ele estará no intervalo $i \pm k$. O elemento mais similar desta faixa é encontrado e adotado como o “espelho” de i de \vec{V}_x em \vec{V}_y . Para todos os outros pontos de \vec{V}_x é realizado isto, permitindo encontrar correspondências. Por isto, os dois mecanismos de EM são, basicamente: função de custo e mecanismo de busca.

O problema da utilização dos padrões unidimensionais encontra-se no trabalho de geração dos “esqueletos”, que envolvem muitas técnicas elaboradas a fim de conseguir encontrar a forma “afinada” do objeto sem gerar ruído, ou eliminar traços do mesmo [43, 44, 45]. Como isto pode envolver um estudo considerável uma outra forma de tratar as imagens bidimensionais com comparações elásticas, além dos contornos, é estabelecer um grafo para sofrer as deformações (*Comparações com Grafos Elásticos* -Elastic Graph Matching - EGM), e determinando pontos “âncoras” no

modelo, a fim de buscá-los durante o reconhecimento, esta segunda forma possui as características elásticas da comparação unidimensional, mas trata a imagem na forma bidimensional sendo utilizada nesta tese.

5.1.1 EGM

O trabalho aqui realizado foi iniciado com uma forma de EGM similar a desenvolvida nos experimentos de Wiskott sobre reconhecimento de rostos, [6, 46, 47], só que, aplicada na verificação de assinaturas. Como o algoritmo original segue uma linha de desenvolvimento diretamente elaborada para reconhecimento de rostos, teve de sofrer algumas modificações.

Inicialmente, Wiskott desenvolveu um grid padrão, que é uma estrutura no formato de um rosto que contem nós dispostos em locais específicos da face, definidos como pontos “âncora”, que são os olhos, nariz, e boca, por exemplo. Este grid recebeu em seus nós informações locais de vários rostos, extraídas por uma filtragem baseada em “Gabor Wavelets” [48] formando um grafo “rotulado”. O algoritmo de comparação de grafos elásticos busca encontrar uma melhor posição para cada nó que maximiza as características de similaridade e minimiza os custos topográficos ao mesmo tempo. A capacidade de determinar a flexibilidade do grafo é controlada pelo custo da topografia sobre a função de custo total.

O método possui duas etapas, na primeira o grafo é disposto sobre um rosto desconhecido e toma como primeiro passo um alinhamento, para posteriormente iniciar, partindo de pontos iniciais bons, a busca de um “ajuste fino”. A partir deste método ele desenvolveu várias extensões do mesmo, onde buscava além do reconhecimento, outros aspectos como o reconhecimento de sexo [49]. Nas próximas seções estenderemos mais o conceito de EGM.

5.2 Estrutura das Comparações Elásticas

Ao utilizar parâmetros (altura, tamanho da assinatura, etc.) elimina-se a descrição das assinaturas como uma imagem bidimensional para se trabalhar com valores armazenados em um vetor de características. Um modo de evitar isto ocorre quando compara-se um padrão com uma imagem, na esperança de encontrar cópias exatas do mesmo. Para isto ocorrer elas devem ter tamanho e orientação idênticas ao modelo. Se as versões testadas possuírem deformações existirão grandes dificuldades na comparação, a menos que a deformação seja pequena.

Para usar padrões em comparações a fim de reconhecer uma característica global, um número alto de modelos será necessário a fim de representar todos os efeitos de uma deformação ou do conjunto de aspectos que pode estar sujeita a assinatura, como rotação, escala ou transformações geométricas destas.

Uma possibilidade de evitar isto é utilizar características locais simples¹. Contudo, dependendo da função de similaridade utilizada, estas são geralmente ambíguas e instáveis. Isto é, dada uma característica, ela pode vir a aparecer em vários pontos do objeto correspondente, ou até em diversos outros objetos. Veja a Figura 5.3 e imagine que buscamos a letra “V” em um texto (5.3-A). Se definirmos as primitivas como os traços inclinados 5.3-C e 5.3-D, e não houver preocupação quando a disposição espacial destas primitivas, tanto pode-se encontrar o modelo procurado, como o objeto 5.3-B que não é desejado, pois tanto ele como o padrão buscado possuem as mesmas primitivas, só diferindo na disposição relativa delas.

Para evitar isto comparações baseadas na similaridade das características locais devem ser avaliadas utilizando suas posições relativas preservando seu arranjo global. O caminho mais fácil de resolver este problema é buscar, na imagem com dis-

¹Não se deve utilizar características locais muito complexas pois estas tornam-se críticas em relação as deformações.

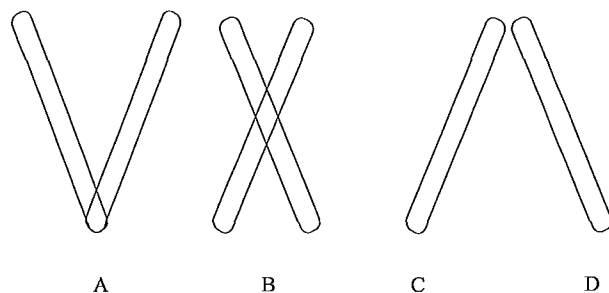


Figura 5.3: Em A temos um arquétipo e B é outro, se tomarmos como primitivas C e D, e as buscarmos as em qualquer posição, sem se preocupar na distribuição relativa dos componentes, teremos A e B iguais, ou melhor respondendo da mesma forma, pois possuem as mesmas primitivas.

torções, similaridades com um padrão composto de características locais e permitir transformações numa tentativa de otimizar o grau de comparação entre os objetos [50]. Em alguns casos, a busca pode não ser necessária, sendo possível normalizar as imagens antes da comparação, porém isto só surte efeito se as deformações forem rígidas, isto é proporcionais em toda a imagem, o que não ocorre com as assinaturas.

Para visualizar melhor o processo, assumamos que a referência, oriunda da base de dados, é uma imagem numa folha de borracha transparente. Movemos esta folha sobre a imagem testada e, para cada possível localização, comprimimos ou distorcemos a folha para gerar o melhor alinhamento possível entre a imagem referência (padrão) e a testada. Avalia-se cada comparação por uma perspectiva dupla:

- quão boa correspondência se é capaz de obter;
- quanta deformação é necessária para obtê-la.

Isto é uma comparação elástica simples. A forma de combinar estes dois efeitos será discutida mais à frente.

A fim de poder trabalhar estes pontos, é necessário utilizar alguma forma de mapear o modelo na imagem testada a partir da correlação de regiões (pontos ou subpadrões) entre as mesmas; assim, trocamos a folha de borracha por uma imagem

modelo composta por peças rígidas (componentes) unidos por “elásticos”. Uma peça rígida da imagem referência pode ser uma célula de resolução simples, ou um grande “pedaço” do modelo corresponde a uma “entidade” deste. Os elásticos ligam os componentes rígidos, e servem para controlar o movimento relativo com a finalidade de medir o *custo* (tipicamente, estes elásticos terão um comportamento não linear). Move-se o modelo sobre a imagem testada em uma busca para encontrar um bom alinhamento mas, como pode-se encontrar várias boas comparações parciais que só satisfazem alguns pontos, é necessário utilizar a característica elástica da mesma para ter todas as posições relativas corretas, descobrindo as melhores posições que faltam. Na determinação do custo embutido, mede-se a tensão de cada elástico (ela pode ser uma função da direção assim como alongamento ou mudança relativa em alguma atributo definido localmente) e também pode-se realizar uma avaliação local de o quão bom cada entidade independente se torna mais coerente em relação ao nó do modelo.

Nas comparações elásticas utilizam-se dados equiponderantes, ou seja eles possuem o mesmo peso para comparação, com objetivo de encontrar a solução. Mas numa aplicação particular podemos ter alguma informação das distorções específicas que podem ocorrer entre a imagem modelo e teste. Por exemplo, alguns subconjuntos de itens que aparecem na imagem testada podem ter suas posições mais variadas no modelo, enquanto outros são quase que fixos. Sabendo que estas variações são possíveis, e sendo capazes de saber a extensão das mudanças, é possível definir o *custo* (valor assinalado) para as mesmas, relativo a sua posição ou mudança em algum atributo não-geométrico. Este custo pode ser implementado numa reformulação de EM. Em vez de cada ponto comparado contribuir como uma soma equivalente para a distância total, enfatiza-se as características mais importantes e descartam as menos importantes, quando aplica-se a verificação. Para simular isto, acentua-se

certos pontos do modelo pela adição de pesos nestes, incrementando ou decrementando sua contribuição para a soma total [40].

O modelo descrito permite estudar as funções de similaridade como duas equações particulares, descritas por Rosenfeld e Kat [51]:

- funções globais;
- funções locais.

A separação das funções globais e locais permitem a aplicação deste algoritmo em vários problemas sem profundas modificações de implementação. A função local está somente unida a parte de encontrar uma primitiva em uma região da imagem testada, enquanto que a global se relaciona com a manutenção da topologia da imagem modelo.

5.2.1 Funções Globais e Locais

A função global é associada com a posição relativa dos “pedaços” acima descritos, e permite uma integração diretamente no algoritmo de decisão. Isto é importante porque a avaliação global envolve os piores problemas combinacionais. Ela é que determina como as primitivas se relacionam entre si e é utilizada na busca da distribuição destas. Isto significa que ela está ligada a como encontrar o grafo que indica um valor mínimo de deformação da disposição, para um máximo de similaridade entre comparações das primitivas.

Existem várias métricas para se alcançar o mínimo da função global. A princípio, temos técnicas de programação dinâmica, que colocando em termos simples, envolvem a minimização de um erro residual entre duas funções pela busca de uma função para reescalar os erros originais [40]. Elas, na prática, se revelam-se com-

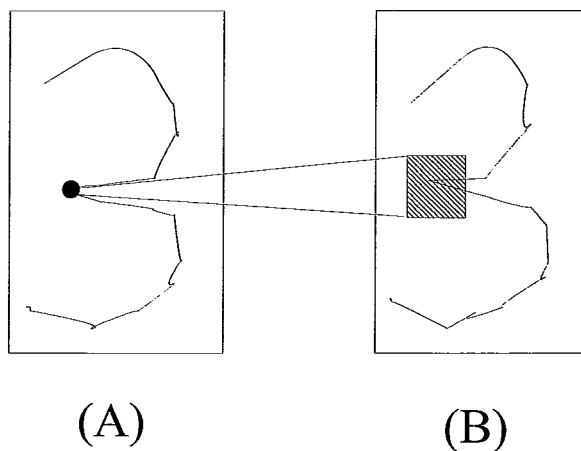


Figura 5.4: Temos em A o modelo que fornece a característica que será buscada em B na região definida pelo quadrado hachurado.

putacionalmente custosas. Muitos trabalhos aplicam uma métrica particular, desenvolvida a partir da progração dinâmica convencional que diminua o custo computacional. Um “atalho” para esta busca é encontrar um subconjunto da estrutura inicial que possua uma similaridade alta para algumas primitivas e então tentar construir uma combinação que seja uma extensão desta, usando heurísticas [50]. Outra forma, é a busca utilizada por Wiskott e nesta tese, é procurar combinar da melhor forma todos as primitivas e então permitir que as mesmas busquem uma melhor disposição em suas posições relativas. Pode-se indicar a distorção geométrica pela introdução de uma tolerância no processo, por exemplo, mapeando cada par de vetores num círculo, limitando seu espaço de deslocamento, o que seria correspondente a comparar o pixel de um com uma região do outro, como pode-se ver na Figura 5.4.

A função local é associada com cada pedaço do modelo. É independentemente codificada em relação a global e facilmente modificada de problema para problema (baseada nas considerações dependentes destes) sem requerer modificações no algo-

ritmo. Assim, a forma de avaliação local pode ser uma função (convencional) de correlação junto com um componente de referência pictorial.

O desacoplamento das funções locais avaliadoras do núcleo do algoritmo permite uma grande variação de propostas para a função de avaliação de um determinado problema, assim como “chavear” as funções para se adaptarem aos objetivos desejados. A avaliação das funções locais e globais pode ser feita separadamente de uma forma direta e intuitiva. Isto permite um melhoramento iterativo na performance por alterações seletivas nas opções dependentes do problema.

Wiskott, por exemplo, aplicou uma fórmula local para calcular a similaridade entre as suas primitivas, denominadas *Jets*, enquanto sua função global contabiliza a combinação entre similaridades locais e deformações globais, fornecendo o valor da melhor disposição encontrada até o final do processo.

5.3 Formalização do EGM

Pode-se formalizar matematicamente a proposta da seguinte forma: seja a referência composta de p componentes (primitivas) que nada mais são do que representações de pontos (ou de um ponto) do modelo em uma nova dimensão. Estes componentes são características locais extraídas de um modelo e apresentadas em um formato que permita a sua “personalização”, representando bem a distribuição do nível cinza do ponto. Para $1 \leq i \leq p$, seja x_i uma variável que pode assumir todas localizações da cena, definindo a posição do componente i^{th} . Supondo que há um mecanismo que, para a localização x_i de i^{th} , fornece um valor numérico $l_i(x_i)$ que indica o quão adequado i^{th} se encontra para a localização x_i da cena, em outras palavras, fornece o índice da similaridade local entre o componente i^{th} do modelo e uma ponto específico da imagem testada, precisaremos determinar como este ponto se relaciona com os

outros.

Até este ponto, $l_i(x_i)$ mede a presença de i^{th} para uma localização na cena independente da posição dos outros componentes. Isto é, $l_i(x_i)$ é uma medida puramente local e imprecisa da presença de i^{th} para o lugar x_i . Em adição a isto, há outras considerações:

- quão bem os diferentes componentes estão situados na busca de relações espaciais entre si;
- a relação entre os valores dos atributos dos componentes comparados com os valores de medida correspondentes na imagem teste (podemos querer especificar que o componente i^{th} seja mais grosso e claro que o j^{th}).

Como as imagens contêm conjuntos de características locais, a utilização de coordenadas ortogonais pode ser uma vantagem para discriminar posições globais, localizando os componentes i^{th} no espaço. Cada posição na imagem testada será, com isto, associada a um vetor bi-dimensional (por exemplo, os componentes do vetor podem ser uma linha e a coluna da sua localização na cena). No caso, i^{th} terá sua posição determinada por \vec{x}_i , e como consequência, para dois pontos \vec{x}_i e \vec{x}_j , do modelo, a diferença $\vec{x}_i - \vec{x}_j$ é o vetor que liga estes, e esta diferença vetorial fornece uma boa representação dos “elásticos”, permitindo identificar modificações euclidianas e angulares e fazer $g_{ij}(x_i, x_j) = \vec{x}_i - \vec{x}_j$ ser o custo associado com o custo elástico entre os componentes i^{th} e j^{th} . O custo $g_{ij}(x, y)$ é zero se não há um elástico entre eles, o que ocorre quando:

- não existem “elásticos” ligando o componente a si mesmo ($i=j$) ou,
- somente há ligações entre elementos vizinhos (o grafo não é totalmente conectado).

Seja o número de componentes na estrutura $X_i = \{x_1, x_2, \dots, x_i\}$, então o custo total da codificação de p componentes para localizações X_p é $G(X_p) = \sum_{i=1}^p \sum_{j=1}^i g_{ij}(x_i, x_j)$ que pode ser escrita como: $G(X_p) = \sum_{i=1}^p h_i(X_i)$ onde $h_i(X_i) = \sum_{j=1}^i g_{ij}(x_i, x_j)$, e $h_i(X_i)$ pode ser visto como o custo de codificar o componente i^{th} na localização x_i . Vejamos como implementar esta proposta.

5.3.1 Primitiva: Jet

Nesta tese, foram utilizados filtros orientados para apresentar as indicações dos gradientes de traçado num ponto, que indicariam uma característica utilizada na verificação humana de assinaturas, a orientação dos traços. Um conjunto de filtros com múltiplas resoluções e orientações fornece o Jet (veja seção 4.3.4) que é formado a partir da convolução da fórmula 4.15, aplicada em 3 tamanhos de filtros e oito orientações totalizando um vetor de 24 elementos que representa o nível cinza próximo do ponto. Ele é utilizado na função das primitivas descritas na subseção 5.2.1. O valor de similaridade entre dois Jets, $S_a(J, J')$, sendo o Jet do modelo J e o do ponto da assinatura teste J' , pode ser realizada pela fórmula:

$$S_a(J, J') = \frac{\sum_{j=0}^n a_j a'_j}{\sqrt{\sum_{j=0}^n a_j \sum_{j=0}^n a'_j}} \quad (5.1)$$

onde n é o número de elementos dos jets, e cada a_i é um valor do vetor jet para comparação. Esta é a fórmula utilizada no trabalho de Wiskott, mas nesta tese, procuraremos apresentar também outras propostas.

5.3.2 Grafo do Modelo

O modelo será representado por um grafo de estrutura bidimensional, veja a Figura 5.6-C. Esta técnica de definir um grid sobre o modelo e utilizá-lo para varrer a imagem buscando correspondências para estes locais de controle no objeto já é bem

conhecida e utilizada no reconhecimento de padrões [34]. No caso de EGM, um grafo rotulado G representando uma assinatura consiste de N nós conectados por E arestas, que determinam uma representação topográfica “gabarito” com o objetivo de dispor as primitivas (Jets) de acordo a posição provável que devem ser encontradas (Figura 5.6-B). Os nós deste gráfico recebem os Jets. Os nós são “ligados” entre si por elásticos, e estes serão definidos como vetores bidimensionais de distância, para permitir o cálculo da deformação dos mesmos. O elástico é definido pela fórmula abaixo:

$$\Delta \vec{x}_e = \vec{x}_n - \vec{x}_{n'} \quad (5.2)$$

$e = 1, 2, 3, \dots, E$ onde cada elástico conecta o nó n ao n' e possui, como um vetor módulo e direção. Em geral, conectamos os nós “vizinhos” entre si.

A definição do formato do grafo é relativamente simples para o reconhecimento de rostos, pois temos padrões a seguir, que é a face. Veja na Figura 5.5-B como os jets estão dispostos para capturarem pontos considerados importantes do objeto (face) que são os olhos e o nariz, por exemplo.

No caso de assinaturas não existe um padrão que possa ser estabelecido para a disposição dos nós do grafo devido a variabilidade de formatos. Por isso é necessário estabelecer um método automático que encontre um conjunto de pontos representativos para a assinatura particular de uma pessoa a fim de gerar o grafo, garantindo a sua cobertura sobre os estilos. Este conjunto de pontos deve possuir, intrinsecamente, a informação do formato a fim de utilizar o conhecimento da distribuição destes pontos do modelo no algoritmo de comparação.

Uma fórmula simples para encontrar um modelo que represente a estrutura da assinatura seria produzir um grid com o formato de um retângulo padrão, e dispô-lo sobre o modelo, como apresentado na Figura 5.5-A. Se existir um ponto sobre um dos seus nós, este é mantido, caso contrário não; com isto, gera-se uma amostra

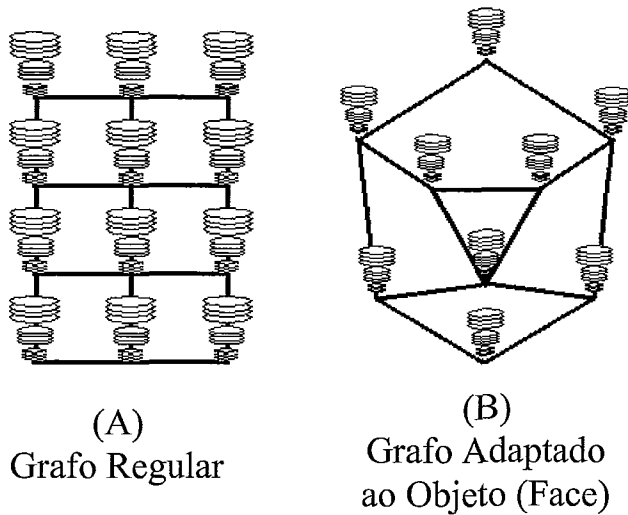


Figura 5.5: Em A temos um grafo regular na forma de um grid que pode ser utilizado para adaptar-se ao formato do objeto. Em B temos um gráfico já adaptado para rostos.

da disposição dos pontos da assinatura. Este grafo será um objeto-adaptado, devido ao fato que a estrutura geométrica bidimensional é adaptada à estrutura da assinatura produzindo um grid que seja semelhante ao objeto buscado. Ele terá uma disposição que siga um polígono que envolva a assinatura. Pode-se usar mais de um grafo, permitindo a chance de produzir modelos mais representativos, entretanto, o tamanho da base de modelos é restrita (por razões de espaço físico para armazenagem), e adicionar novos modelos também implica em ter mais modelos para comparar, diminuindo a velocidade. Uma outra possibilidade é encontrar os pontos mais informativos, sendo que este conceito do que é informativo é relativo à representação que foi adotada; por exemplo, na nossa representação, pode-se considerar que um Jet que possua a maioria dos componentes dos seus vetores diferentes de zero tenha mais informação, ou aquele que possua na soma dos seus componentes um valor acima de um dado limiar como pontos de maior importância; o problema destas duas últimas é que estes pontos podem não estar distribuídos linearmente pela assinatura e teríamos alguma perda da informação bidimensional da estrutura

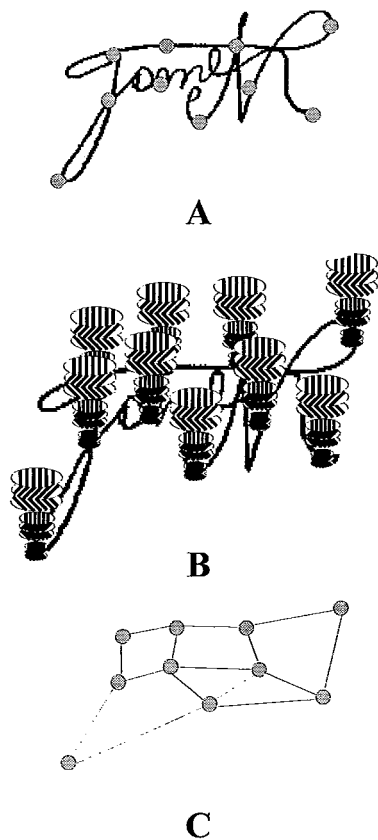


Figura 5.6: Em A temos a assinatura com a localização dos pontos tomados como primitivas, em B temos a representação que nestes pontos formou-se Jets, em C a máscara (considere cada nó contendo um Jet)

da mesma.

Descritas as características locais e a forma de sua comparação, e após indicar como representá-las numa visão global, é necessário apresentar a fórmula responsável pela comparação das posições relativas dos Jets, e seus movimentos. Para isto, definiremos a partir deste ponto como *máscara* o grafo do modelo e *imagem* como o grafo que foi estabelecido na assinatura testada. No problema de reconhecimento de faces Wiskott utilizou a descrição dos estados de disposição da máscara fornecida pela equação total da similaridade entre uma máscara G e uma imagem G' :

$$S_m(G, G') = C_s(J, J') - \lambda C_d(\vec{x}_e, \vec{x}'_e) \quad (5.3)$$

onde defini-se S_m como o índice total de semelhança fornecido pelo valor de similaridade local entre os pontos da máscara e da imagem (C_s), penalizado pelas de-

formações (C_d). O fator λ será incorporado à fórmula como uma constante que fornece o valor do peso das deformações (o quanto as deformações irão atuar sobre o valor final). Se $\lambda \rightarrow \infty$ não é permitida deformações no grafo, no caso oposto, $\lambda \rightarrow 0$ o grafo pode ser distorcido ilimitadamente. Desta forma, esta constante representa a importância relativa da semelhança topográfica.

Semelhança Local

A fórmula geral deve conter as medidas de similaridade locais e adicionar como penalização as deformações globais. Assim, o valor de C_s (Equação 5.4) é composto pela soma das similaridades locais dos *Jets* (Equação 5.1) dividida pelo número de nós que o grafo possui, N , para fornecer a média. Esta parte da Equação 5.4 representa a fórmula local definida na seção 5.2.1.

$$C_s(J, J') = \frac{1}{N} \sum_n S_n(J, J') \quad (5.4)$$

Semelhança Global

Já C_d é a deformação total composta por todas as diferenças entre elásticos. Representa-se a deformação dos elásticos pela divergência entre a forma do modelo e imagem teste, definida por:

$$\Delta \vec{x}_e - \Delta \vec{x}'_e \quad (5.5)$$

onde $\Delta \vec{x}_e$ é o elástico da máscara e $\Delta \vec{x}'_e$ da assinatura testada, sendo a diferença é utilizada para medir as deformações relativas dos elásticos. Ou seja verifica o quanto a estrutura se deformou para se adaptar a assinatura teste. Nesta tese, as diferenças entre os comprimentos dos elásticos e ângulos foram utilizados para medir a deformação [40]. Pode-se visualizar a diferença entre os elásticos no exemplo da Figura 5.7. Tendo-se o elástico do modelo indicado na forma tracejada, e na imagem

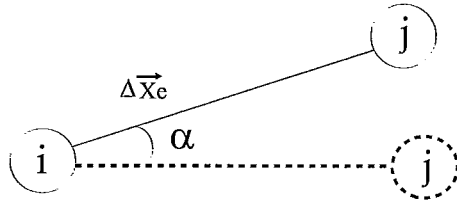


Figura 5.7: A deformação do grafo é produzida pelo deslocamento do componente j . Ele inicialmente está na posição tracejada, e após o movimento local, desloca-se para uma nova posição em relação à i , indicada com traços contínuos. O ângulo α , e a diferença entre a distância inicial e final são utilizados no cálculo da deformação. Aqui optou-se por manter o componente i fixo, para melhor visualização, mas nada impede que no caso real, o componente i mova-se também.

teste pela linha contínua, a diferença entre os dois será fornecida pelo ângulo α , assim como as distância entre i e j na disposição inicial e final.

Esta diferença é apresentada na Equação 5.6. Como na anterior, o somatório é dividido pelo número de elásticos E . Concluindo,, o valor da similaridade total fornecida por S_m é uma combinação linear de um termo de distância C_d , que mede a distorção do grafo, e o termo de vértice C_s , definido como a média da similaridade dos Jets.

$$C_d(\vec{x}_e, \vec{x}'_e) = \frac{1}{E} \sum_e \frac{(\Delta \vec{x}_e - \Delta \vec{x}'_e)^2}{(\Delta \vec{x}'_e)^2} \quad (5.6)$$

Considerações Finais

Devido as difereças entre assinaturas a saída da Equação 5.3 nunca será ótima ($=1,0$), visto que para isto ocorrer os pontos devem coincidir totalmente, tanto em características como na posição, o que dificilmente acontecerá, sendo mais provável que alguns pontos sejam até excluídos.

De posse da fórmula que calcula quão similar um grafo encontrado é diante do

modelo e que penaliza suas deformações. Resta-se demonstrar como será efetuada a busca do grafo ótimo pelo algoritmo. Há algumas formas de implementar esta busca pelo melhor grafo, como já apresentamos na seção 5.2.1. A mais simples, usada por Wiskott [6] para o reconhecimento de faces, é baseada no método do gradiente descendente, e foi utilizada nesta tese da maneira que descreveremos a seguir.

5.3.3 Movimento Global

O algoritmo de EGM é dividido em duas etapas de otimização que varrem a imagem em busca do custo total mínimo da comparação com o modelo. A primeira parte do mapeamento consiste em encontrar a parte da imagem onde o objeto está localizado, fazendo um pré-alinhamento, já que o modelo pode iniciar a comparação posicionado no centro, ou no canto direito, indiferentemente. Este é executado deslocando o gráfico enquanto mantêm-se os elásticos rígidos, e assim, movendo a máscara sobre a assinatura testada até encontrar uma disposição inicial que seja uma melhor posição de comparação sem distorcer o grafo, o que corresponde a $\lambda \rightarrow \infty$.

Esta primeira etapa não utiliza a função de custo topográfico C_d ao computar a Fórmula 5.3 pois os elásticos são movidos rigidamente e em consequência os nós são movidos uniformemente entre si. A cada deslocamento, o valor total é computado e aceito se reduz o custo total.

O grid desloca-se com um tamanho máximo, que pode ser um valor qualquer como 10 pixels. Nesta tese preferimos deslocar inicialmente com um valor máximo igual às distâncias entre os pontos que definem o grid utilizado para desenvolver o modelo. Este valor foi implementado porque a distância que pode existir de translação é pequena na medida que a assinatura já sofre um pré-processamento para eliminar o fundo, assim como uma normalização.

Como nossa imagem só contém um objeto é mais fácil executar este alinhamento;

se ela possuísse vários uma busca mais sofisticada seria requerida. Embora esta procura não revele um valor bom para todos os pontos, ela encontra uma solução mais próxima da correta que a inicial, permitindo que o algoritmo comece com uma disposição boa para um número alto de pontos e possa se deformar para chegar a um nível melhor de similaridade ao iniciar seu “ajuste fino”.

Quando executa-se este movimento global, pode-se encontrar algumas dificuldades em encontrar uma disposição relativa boa se ocorrer uma diferença entre as escalas do modelo e imagem testada. Para evitar isto, alguns pesquisadores [6, 52] permitem também a compressão ou a dilatação linear da máscara durante os mesmo, em outras palavras, comprimem ou dilatam todos os elásticos por um mesmo fator. Nesta tese isto não foi utilizado porque a normalização das assinaturas apresentada no Capítulo 6 já dispõe todas sobre um mesmo tamanho.

No caso particular de assinaturas, é possível implementar uma decisão multinível proposta na Seção 2.4.4, onde falsificações grosseiras seriam eliminadas rapidamente. Nese caso, o valor da comparação desta primeira etapa forneceria um índice que já pode ser comparado com o limiar, se o valor for abaixo da metade deste, por exemplo, a assinatura pode ser recusada neste ponto e os processos posteriores abortados.

5.3.4 Movimento Local

Após ter encontrado uma disposição boa inicial para o maior número de pontos possíveis, pode-se iniciar o processo elástico. Os nós serão movimentados relativamente entre si procurando uma melhor medida de similaridade. Como limitaremos a área de busca, os nós permanecerão com pequenas disparidades de localização, limitando, desta forma, o grau de deformação dos elásticos. Este *look-ahead* restrito é uma das características básicas do EM, e a primeira vista, pode-se utilizar como limite a área entre os Jets do modelo, o que impedirá uma deformação maior do

grafo.

A regra da busca é procurar a melhor adaptação pois, se um ponto não é uma comparação exata com seu correspondente no modelo, pode ser melhor buscar na sua vizinhança. Isto é feito de forma similar à comparação elástica de vetores descrita na Seção 5.1, só que agora busca-se em um espaço bidimensional, uma área. Ela possivelmente terá sucesso porque a maioria dos outros pontos encontraram uma boa correlação, sendo provável que, por manter a disposição do formato, o ponto buscado esteja próximo. Isto ajuda em pequenas variações da característica, mas no caso da deformação ser alta, o algoritmo não funcionará corretamente. Não estamos muito preocupados com isto, porque a permissão de uma deformação alta traria como resultado indesejado o aumento do número de falsificações aceitas.

Como possui-se uma fórmula (Equação 5.3) para determinar a qualidade de um grafo da imagem em consequência dos movimentos dos nós nestes, e a forma de buscar o melhor grafo é irrelevante, pode-se utilizar um algoritmo iterativo que atue modificando gradativamente o grafo para sua melhor adaptação [53]. A idéia utilizada é iniciar com uma configuração e fazer modificações para melhorar a qualidade. O melhor caminho para entender como desenvolve-se a solução deste problema é considerar todos os estados da disposição do grafo como uma superfície. O valor para cada ponto corresponde a avaliação da função global para aquele ponto. Deve-se mover-se ao longo desta superfície tentando encontrar os mínimos (vales) que são as melhores soluções. No nosso caso, não é necessário buscar o vale mais profundo, que corresponderia ao ótimo global, porque um ótimo local fornece uma boa idéia da disposição dos pixels.

5.3.5 Considerações de Complexidade

Alguns pesquisadores descrevem que a complexidade computacional da comparações de grafos para se encontrar um máximo global² é suficientemente alta para tornar proibitivos os algoritmos implementados com ela [54].

Note que cada ponto terá sua resposta de similaridade dependendo dos demais, e estes também dependerão do mesmo. Esta interdependência é que torna o problema difícil, pois a cada atualização do ponto, seus pontos vizinhos sofrem alguma modificação pois os elásticos tiveram de ser deformados, e isso irá contabilizar um novo valor para a fórmula de similaridade total. Mas quando atualiza-se estes temos um novo grafo, onde o ponto inicial pode ter de ser novamente movimentado para um novo mínimo. Isto direciona a solução do problema para trabalhar com Programação Dinâmica. Entretanto, este é um argumento “errado” por duas razões [54]:

- Os problemas de reconhecimento de padrões são resolvidos em quase todas as situações por boas aproximações da melhor solução que possa ser encontrada. Em particular, encontrar o máximo global torna o problema de comparação com grafos intratável. Porém, como exemplo de que isto não é necessário, podemos citar os sistemas de redes neurais que são especialistas em buscar uma solução por aproximações devido a sua natureza estocástica.
- A explosão de várias possíveis soluções é reduzida pela constatação que os pontos são distinguíveis por seus atributos, o que ocorre porque a Função 5.1 gera um ponto ótimo na forma de um atrator. Assim, uma convergência pode ser esperada pela função do campo potencial (Figura 5.8). Mesmo um mínimo local que é complicado de se obter em problemas difíceis de otimização não é

²Este máximo seria definido como a melhor disposição do “casamento” entre todos os pontos do modelo e os pontos da imagem testada.

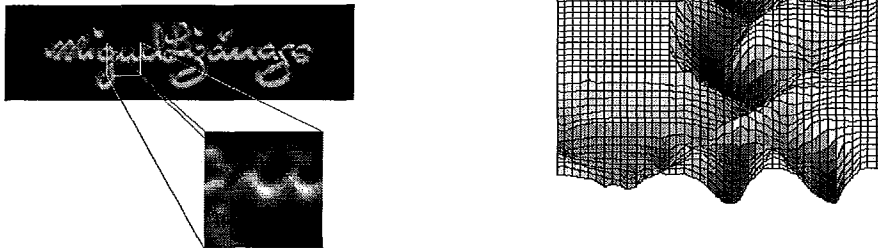


Figura 5.8: Vemos aqui, à esquerda, a imagem resultante da comparação de um jet, com uma assinatura. No centro do quadrado da imagem está o ponto representado pelo Jet, note que as partes mais claras indicam onde a similaridade tem seu valor mais alto. A área do quadrado é extraída, e representada como uma superfície, à direita, por uma vista particular que oferece mais vantagens para se distinguir os vales (eixos trocados). Na figura da superfície, os vales tem uma valor alto de similaridade e o patamar no canto superior esquerdo representa um índice zero de similaridade, pois é uma região onde não se encontra nenhum traço da assinatura, só o fundo. Pode-se perceber com isto porque os Jets se deslocam pela superfície da imagem a fim de buscar o ponto mais similar, que está no fundo do vale.

demorado para métodos simples, como a difusão gráfica de comparação.

Outro ponto positivo para o processamento de assinaturas em particular é a oportunidade de eliminar o fundo durante a comparação, só executando-a entre pontos que pertençam a traços da assinatura testada, tendo uma área menor de busca, e mais eficiência do algoritmo. Ao final do mapeamento, temos uma correspondência entre os pontos e o valor final da função de comparação, o qual pode-se utilizar como o valor de similaridade entre o modelo e assinatura de teste, para, via utilização de um limiar, decidir se esta é verdadeira ou não.

5.3.6 Comparação de Retorno

Uma característica apresentada neste trabalho é da possibilidade de executar uma comparação de retorno entre a assinatura modelo e a de teste.

O modelo armazenado gera para comparação um grafo que representa sua es-

trutura. Este é um grafo particular que incorpora informações desta com o objetivo de representar as características que todas as genuínas possuem. O grafo possui pequenas diferenças em relação as outras assinaturas verdadeiras, embora tenha deformações. Qualquer assinatura genuína deve ser capaz de fornecer estas particularidades, visto que elas são incorporadas a todas, como dito no Capítulo 1.1. Assim, ao se comparar uma verdadeira, esta poderia ser a assinatura que gera o grafo modelo, já que terá de possuir as características do autor.

Se compararmos um grafo obtido da assinatura teste (sendo esta verdadeira) a resposta deve ser tão boa quanto a original. Já ao compararmos com uma falsa, temos duas hipóteses:

- se for uma falsificação habilitada pode-se retirar informação do falsificador (ver Seção 2.1.1). Este pode ter conseguido representar bem as informações gerais da assinatura, mas suas características particulares, se existem, podem diminuir o índice de similaridade por não pertencerem às duas assinaturas.
- se for uma aleatória, ao possuir a maioria das características da verdadeira, ela pode ser aceita. Mas com esta dupla comparação (ver Figura 5.9) teremos a possibilidade maior de capturá-las. O objetivo deste processo é dificultar a aceitação de falsificações quando executa-se duas comparações a possibilidade de uma assinatura randômica ser aceita é atenuada.

Ele permite, como visto na seção 2.1.1 que a assinatura testada gere um novo modelo de grafo para comparação, se ela for uma assinatura verdadeira teremos só uma confirmação, se for falsa, há possibilidades de encontrar características do falsificador que deponham contra a mesma.

Seu ponto positivo está na idéia que o algoritmo não precisa ser modificado para efetuar este novo processamento, só é necessário trocar o rótulo das imagens

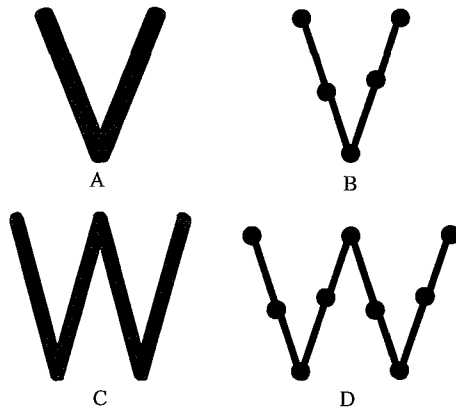


Figura 5.9: Imagine que temos o um grafo como a Figura B obtida do modelo A. Este grafo será testado sobre a imagem C. Como a imagem testada contem um, ou melhor, dois arquétipos do modelo, quando fizermos a busca global e ajustarmos o grafo sobre a imagem encontraremos um alto grau de similaridade, embora esta não seja a imagem procurada. Ao determinar a Figura C como modelo, gera-se o grafo apresentado em D. Ele possui informações que não existem em A, e contribuirão para uma saída baixa.

e continuar o processamento. O ponto negativo é que enquanto sem ele poderíamos armazenar só o grafo com os Jets, que é mais compacto que a imagem; e a duplicação do processo.

Capítulo 6

Resultados e Discussões

“Nós pensamos que o que fazemos é apenas uma gota no oceano.

Mas o oceano, sem essa gota, seria menor.

Madre Teresa.

Realizaram-se dois estudos de verificação. Uma primeira proposta implementada busca uma comparação global, onde pesquisamos o efeito do filtro orientado, introduzido no Capítulo 4 sobre a verificação com duas bases de dados. O segundo estudo foi realizado com as comparações de Gráficos Elásticos, sobre os mesmos grupos.

6.1 Bases de Dados

É imprescindível definir antes de apresentar os resultados obtidos a base de dados, com o objetivo de explicar a proposta de testes efetuados. Os testes de verificação foram efetuados sobre duas bases de dados.

6.1.1 Primeira Base

A primeira base foi adquirida junto ao Prof. Lee Luang Ling, da Unicamp. O conjunto tem resolução de 300dpi e já foi pré-processado no sentido de eliminar ruído. A imagem só teve sua representação modificada de arquivos bitmap para o

formato de imagem utilizado nos algoritmos desta tese. Esta é composta de 740 assinaturas no total, que pode ser dividida em cinco grupos principais:

- Os quatro primeiros grupos são compostos por quatro classes (autores) e contém 100 assinaturas genuínas com 60 falsificações habilitadas.
- O quinto é formado por 100 assinaturas de cerca de 40 pessoas diferentes.

A partir desta distribuição, dividiu-se aleatoriamente o grupo de verdadeiras de cada um dos quatro primeiros grupos, em dois subconjuntos, de 60 e 40 assinaturas. O grupo de 60 genuínas foram processados os testes, e após estes modelos serem processadas pelos algoritmos, ofereciam alguns limiares para os algoritmos de decisão indicarem se uma assinatura é verdadeira ou não. Três grupos foram processados para terem seus valores comparados a estes limiares e gerar as faixas de erros:

- Habilitadas: as 60 assinaturas correspondentes às falsificações especializadas.
- Aleatórias: são 60 assinaturas retiradas do quinto grupo, compondo as assinaturas aleatórias para o teste.
- Teste: 40 assinaturas genuínas restantes.

Os dois primeiros grupos oferecem a possibilidade de testar FAR enquanto o último, a FRR . Esta base é útil para visualizar o comportamento intrapessoal dos algoritmos, na medida que oferece um grande número de exemplos de cada classe.

6.1.2 Segunda Base

Esta foi criada nas dependências da COPPE-UFRJ, a partir de voluntários entre professores e alunos. As assinaturas não foram colhidas em formulários especiais, e sim folhas em branco para não existir qualquer indicação que tolhesse as defomações.

Não houve preocupação com a caneta utilizada, a não ser a indicação que fosse uma esferográfica comum. Foram digitalizadas em um scanner com resolução de 150dpi. A separação foi realizada por meio de um editor de imagens onde não buscou-se centralizar ou alinhar o retângulo com base na assinatura. Não houve pré-processamento prévio a não ser o mesmo realizado com a base 1, descrito nas próximas seções. Com o total de 600 assinaturas, está disposta da seguinte forma: 24 classes (autores), 15 assinaturas verdadeiras para cada uma; e 10 assinaturas falsificadas especializadas.

A proposta, com este grupo, é utilizar oito assinaturas verdadeiras para gerar o limiar e ter três grupos de testes:

- **Habilitadas:** as 10 assinaturas correspondentes às falsificações especializadas determinadas para cada classe.
- **Aleatórias:** são 10 assinaturas retiradas da primeira base de dados, e compõem as assinaturas aleatórias para o teste.
- **Teste:** 7 assinaturas genuínas restantes, de cada classe.

Como na base anterior, os dois primeiros grupos informam a taxa de FAR e o terceiro a de FRR. Esta segunda base oferece uma oportunidade de estudar a variabilidade interpessoal, por fornecer um número maior de classes.

6.2 Pré-processamento

6.2.1 Representação Canônica

A fim de executarmos uma comparação global que indique como os filtros orientados podem auxiliar na verificação de assinaturas, é preciso definir um formato padrão. Esta necessidade ocorre por que devemos efetuar medidas de correlação de

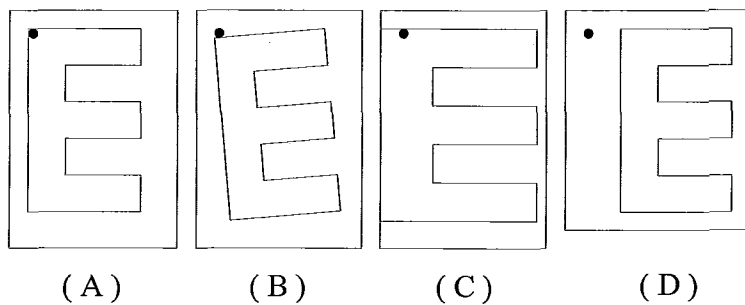


Figura 6.1: O ponto escuro está localizado na mesma região em todas as quatro figuras, ao definir o canto superior esquerdo como referência, e seriam comparados entre si ao utilizar-se características globais, apresentando algumas deformações possíveis: (A) Imagem original, (B) Imagem rotacionada, (C) Deformação de dimensão (D) Translação e tamanho diferente

características globais, pois este tipo de comparação sofre problemas com translação, rotação e escala.

Na Figura 6.1 pode-se visualizar os tipos de deformações. Tome como aspecto de comparação o ponto preto no canto superior esquerdo das figuras A, B, C e D. Comparando A com B já verifica-se uma perda de informação. Percebe-se que na comparação global, as retas de A terão valores baixos de similaridade quando comparadas com as de B, por causa das segundas disporem-se em um ângulo em relação às primeiras. No caso de comparar A com C acontece um problema semelhante, pois a translação existe devido ao fato do ponto estar em um “canto” do objeto que é a referência e no testado, estar no meio em uma comparação global, pois as bordas das imagens estarem completamente fora de padrão entre si. Com o processo empregado, elimina-se o problema, ao centralizar a assinatura. Além disto, há uma diminuição do efeito de rotação (caso exista), com o decréscimo do tamanho. Um formato canônico elimina os efeitos de translação, além de minimizar os efeitos de escala e rotação.

Assim, um pré-processamento é executado sobre o grupo a ser estudado levando-o a um tamanho padrão. O primeiro fator é determinar, a partir de um grupo de modelos, um tamanho mínimo padrão. Foi definido um mínimo e não um máximo,

porque assim o custo de comparação, que é associado ao número de pontos, é menor.

O algoritmo global não precisou utilizar um alinhamento entre os pontos ou regiões. A segmentação da assinatura eliminou o problema de translação ao retirar o fundo que poderia atuar deslocando-a em relação ao referencial, como apresentado acima. Outra vantagem desta compressão é a padronização das assinaturas verdadeiras, que, devido a variação de escala, possuíam uma deformação. Com esta compressão a deformação total é minimizada, embora a assinatura continue com as deformações locais, que não são eliminadas pela contração total. A rotação pode ser atenuada com a representação no tamanho menor, pois quando diminuimos uma imagem, ela terá um deslocamento menor em relação ao eixo vertical.

Nesta representação foram eliminados os espaços em branco entre nome e sobrenome, na medida que, estes espaços podem ser confundidos com o espaçamento dentro da assinatura (no caso do autor da assinatura escrever sem ligar as letras). Isto também testa o poder de discriminação do algoritmo permitindo-o ser comparado com um número maior de assinaturas de teste, que seriam eliminadas se fosse tomado como característica o número de palavras que o nome contém.

O algoritmo de compressão segue os seguintes passos:

- Eliminação das partes “brancas” (margens inúteis) do modelo tratado e determinação do tamanho total da informação armazenada do mesmo.
- Verifica-se o tamanho de um grupo¹ de assinaturas modelo e obtêm-se a informação sobre o tamanho da assinatura, sendo só a parte da imagem que apresenta traços de escrita avaliada para determinar este valor. Seleciona-se o de menor tamanho e determina-se que este é o tamanho padrão.

¹ Ele pode ser o total, ou um grupo menor, e determinar que 80% desta é o tamanho ideal, por exemplo

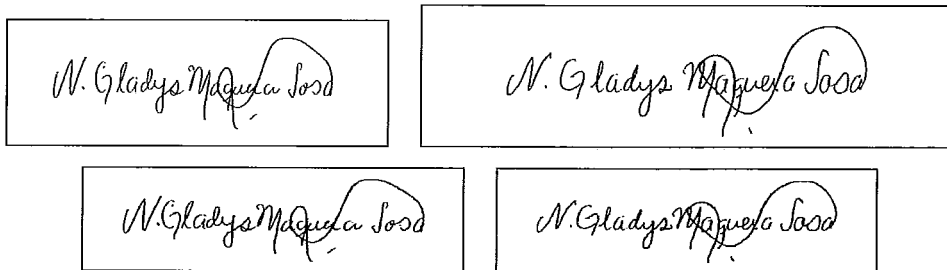


Figura 6.2: As assinaturas superiores são as originais, onde pode-se perceber a diferença de dimensões. As assinaturas abaixo destas são as mesmas pré-processadas, possuindo mesma dimensão e alinhamento.

- Com este tamanho padrão comprime-se as outras assinaturas.

As assinaturas da Figura 6.2 são apresentadas primeiro em sua forma natural e depois com o pré-processamento. Se houver no grupo assinaturas que forem menores que este formato serão acrescentadas bordas em branco ao invés da compressão, até que atinjam tamanho padrão desejado.

6.3 Fórmulas e Limiar de Decisão:

Um programa de filtros orientados foi aplicado na assinatura, produzindo como saída a soma dos valores da resposta ao filtro em oito orientações formando uma imagem, veja alguns exemplo de imagens geradas são as Figuras 6.3, 6.4, 6.5. Assim, armazenou-se um conjunto de dados com o intuito de proporcionar uma visão das orientações e espessuras dos traços que compõem a assinatura. Cada uma das vistas particulares dos filtros, que formam a *Resposta Composta*², enfatiza uma orientação particular de escrita (uma orientação do traçado do escritor) em detrimento aos outros, devido à orientação do filtro. O tamanho e orientação do filtro aplicado definem a forma de seleção destas características, conforme explicou-se no Capítulo 4. Iniciou-se o processo com um filtro de tamanho pequeno, que seleciona a área

²A partir deste ponto, utilizaremos a denominação Resposta Composta para indicar a imagem formada pela soma das 8 saídas dos filtros orientados.

próxima do ponto, que depois cresceu, para objetivar a importância do tamanho do filtro no processo.

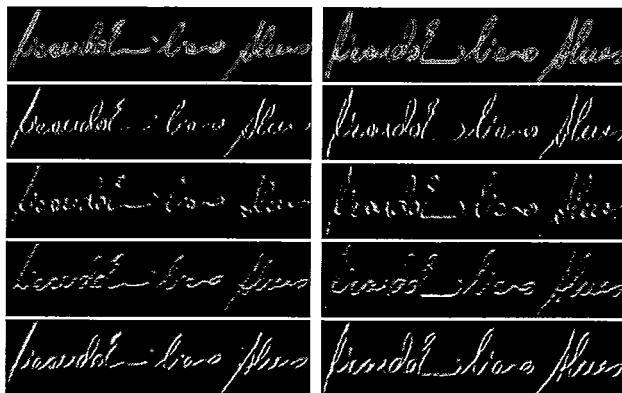


Figura 6.3: Imagens das Orientações e Resposta Composta dos Modelos 1 e 2, saída do Filtro Pequeno - 0.5/1.5

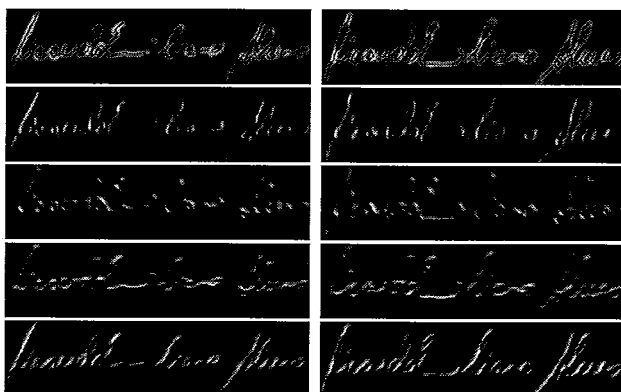


Figura 6.4: Imagens das Orientações e Resposta Composta dos Modelos 1 e 2, saída do Filtro Médio - 1.5/3.0

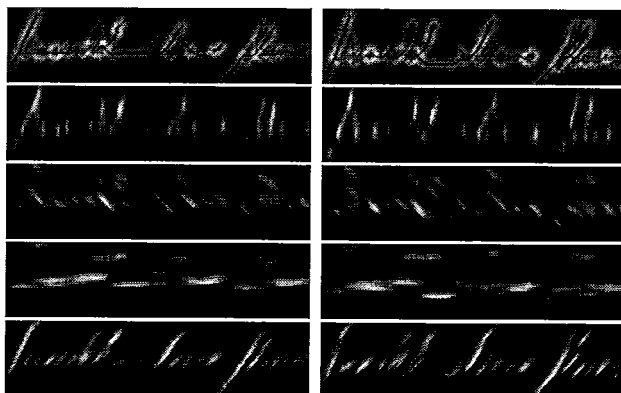


Figura 6.5: Imagens das Orientações e Resposta Composta dos Modelos 1 e 2, saída do Filtro Maior - 3.0/9.0

Uma vez de posse das imagens pré-processadas, faz-se necessário compará-las entre si. Para proceder a comparação foi estabelecida a Fórmula 6.1 para medir a diferenças entre imagens:

$$D(A, B) = \sum_{k=1}^n \frac{|A_k - B_k|}{|A_k| + |B_k|} \quad (6.1)$$

Onde, n é o número de pontos da imagem, e A_k e B_k são cada um destes pontos que compõem a mesma. Esta fórmula está relacionada com a comparação dos valores absolutos da imagem obtida pela filtragem da imagem da assinatura. Note que este é um valor de diferença, então para se obter a similaridade é necessário efetuar a Fórmula 6.2, obtendo-se $S(A, B)$.

$$S(A, B) = 1 - D(A, B) \quad (6.2)$$

Há um problema com a implementação desta fórmula. No caso, ela estabelece uma comparação ponto a ponto que irá atuar também no fundo da imagem (pontos de valor igual a zero). Como este é igual para todas, estabelece-se uma comparação errônea, resultante do fato que, tendo mesmas características, os pontos do fundo formariam um falso acréscimo ao valor. Para resolver este problema e também com a idéia de encontrar uma forma de realçar os pontos mais influentes da assinatura foram realizados testes com uma seleção baseada em limiares definidos da seguinte forma:

Considere que $P(x, y)$ é um ponto, $I_m(x, y)$ é o valor deste ponto na imagem modelo, $I_t(x, y)$ é o valor na imagem testada, C_p é o conjunto de pontos selecionados como pertencentes a assinatura, a fim de efetuar a comparação, e th é um limiar que indica a segmentação entre imagem e fundo, teremos as comparações descritas da seguinte forma:

- AND: Compara pontos que possuam um valor superior ao limiar nas duas imagens filtradas. Ou seja,

$$\begin{aligned} \text{Se } I_m(x, y) > th \text{ e} \\ I_i(x, y) > th, \\ P(x, y) \in C_p \end{aligned}$$

Veja Figura 6.6.

- OU: Compara todos os pontos que possuam um valor não nulo em qualquer uma das 2 imagens:

$$\begin{aligned} \text{Se } I_m(x, y) > th \text{ ou} \\ I_i(x, y) > th, \\ P(x, y) \in C_p \end{aligned}$$

Veja Figura 6.7.

- TEMP: Compara todos os pontos de valor superior a zero na imagem de referência (como uma máscara):

$$\begin{aligned} \text{Se } I_m(x, y) > th, \\ P(x, y) \in C_p \end{aligned}$$

Veja Figura 6.8.

Alguns valores foram testados para este limiar th , o que obteve melhor resultados foi utilizar um valor bem próximo de zero.

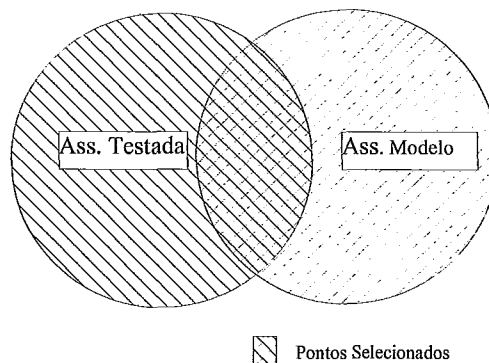
A fórmula de correlação vetorial 6.3 pode ser aplicada também. Ela produz uma comparação que é capaz de eliminar o problema do fundo, pois não inclui para medida de similaridade as contribuições de valor zero que não contribui com

informação para o julgamento da similaridade, implementando automaticamente uma estratégia similar a AND.

$$S(A, B) = \sum_{k=1}^n \frac{A_k B_k}{\sqrt{A_k^2 B_k^2}} \quad (6.3)$$

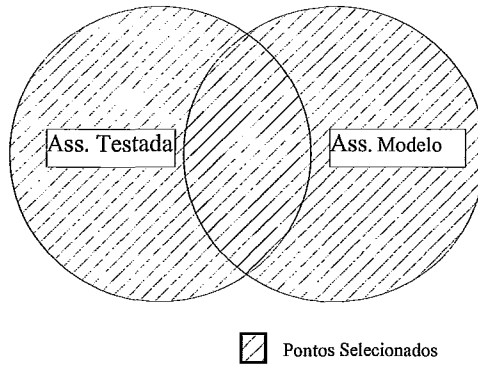
Os valores dos limiares de decisão indicam o índice de similaridade mínimo exigido para se considerar a assinatura verdadeira. Eles surgiram de duas formas, a partir do grupo de 60 assinaturas:

- pelos 2 valores mínimos resultantes das comparação com o grupo de 60 modelos. A ideia de buscar estes valores advêm do fato do primeiro mínimo indicar que todo o grupo definido passaria nos testes, e o segundo é utilizado no caso do mínimo encontrado ser resultado de uma assinatura que por algum motivo, seja muito distorçida, gerando um valor muito baixo. Veja a Figura 6.9
- o limiar relacionado com o desvio padrão é extraído a partir do desvio dos valores de respostas. A proposta foi utilizar 3 valores: $\mu - \sigma$, $\mu - 2\sigma$ e $\mu - 3\sigma$ (sendo μ a média e σ o desvio padrão do grupo de 60 assinaturas tomado para se extrair o limiar de decisão). O motivo para usar estes valores é que se a distribuição dos valores de similaridade de assinaturas segue uma distribuição normal, isto determinaria porcentagens superiores a 68%, 95% e 99%



Conjunto AND

Figura 6.6: Visualização do Processo AND, indicando quais pontos são selecionados.



Conjunto OU

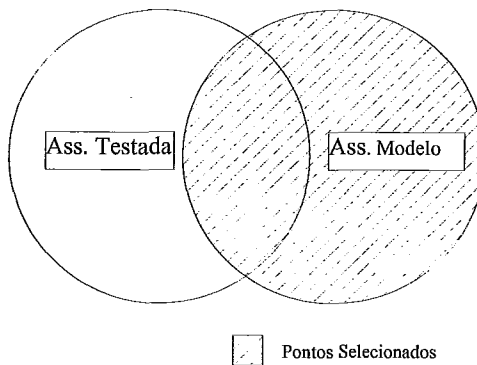
Figura 6.7: Visualização do Processo OU, indicando quais pontos são selecionados.

de assinaturas reconhecidas pelo algoritmo. Veja a Figura 6.10.

6.4 Comparações Globais

Passaremos a apresentar os resultados das comparações globais, tendo de informar alguns parâmetros.

O valor escolhido para se apresentar as tabelas dos resultados foi o de $\mu - 2\sigma$, porque o índice de $\mu - 3\sigma$ revelou um aumento de FAR superior a faixa de decréscimo de FRR conseguida, o que não é desejável, visto que a faixa de FRR para o valor já é baixa. Valores para $\mu - \sigma$ apresentaram taxas altas de FRR, por isso os descartamos. Os valores de mínimos foram também excluídos por não apresentarem melhores



Conjunto TEMP

Figura 6.8: Visualização do Processo TEMP, indicando quais pontos são selecionados.

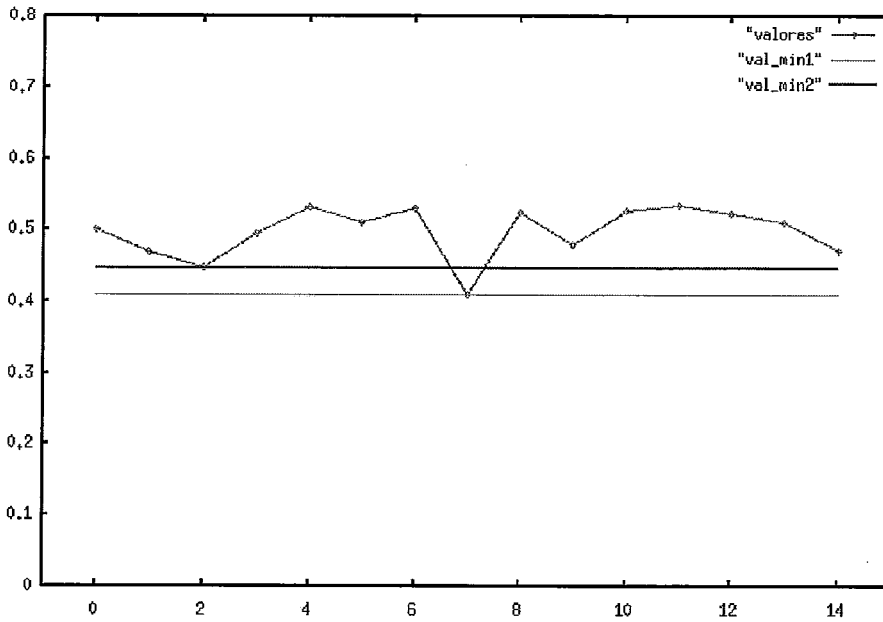


Figura 6.9: Para visualização é apresentada uma curva de 15 pontos que são os resultados de comparações com assinaturas genuínas. Os dois mínimos da curva são tomados por limiares e cortados por uma reta, que é a reta de decisão. A mais baixa para o limiar definido pelo valor mínimo da curva (val_min1) e a segunda (val_min2) é determinado pelo segundo valor mínimo da curva.

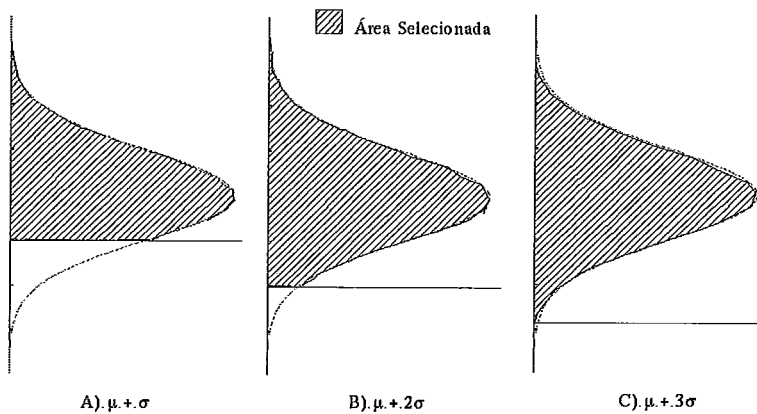
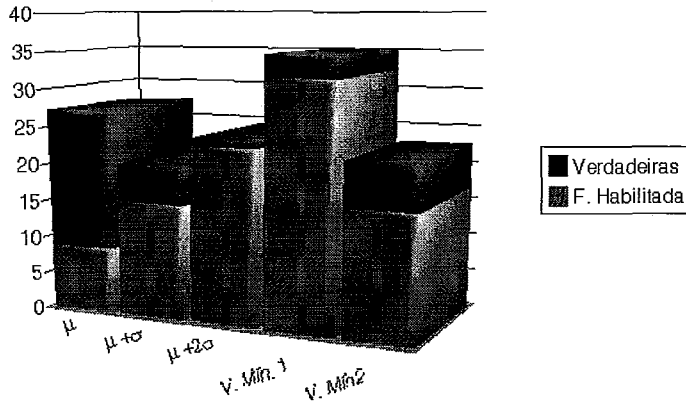


Figura 6.10: Os limiares são relativos ao desvio padrão das respostas das assinaturas verdadeiras, em (A) temos a região selecionada com o primeiro limiar de $\mu - \sigma$, em (B) $\mu - 2\sigma$ e (C) $\mu - 3\sigma$. Veja que a área selecionada, que determina o número de assinaturas verdadeiras cresce, pois o limiar é cada vez menos restritivo.

valores que o adotado. Veja a Figura 6.11.

Por simplicidade, apresentaremos os resultados em função do índice de FTR , sendo este a média simples dos dois primeiros, $FTR = \frac{FAR+FRR}{2}$. Pois assim não

Faixas de Erros



Faixas de Erros

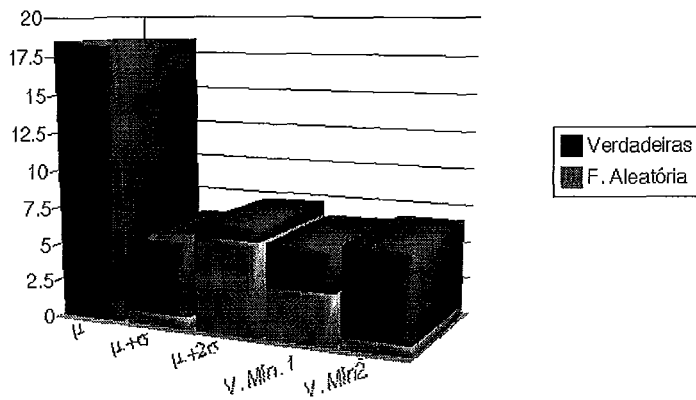


Figura 6.11: A imagem apresenta dois gráficos com os valores de erros. O superior demonstra as típicas faixas de erros das falsificações habilitadas. Sob um tom mais escuro temos a FTR e sob o mais claro a FAR. Como as taxas dos limiares μ e *V. Mín. 1* são altas, não são utilizadas, por isso. O limiar de $\mu - 2\sigma$ apresenta uma taxa alta de falsificações aceitas. Em geral, este acréscimo é superior ao decréscimo obtido na taxa de rejeição de verdadeiras, definindo um desempenho pior. Por fim, o limiar de *V. Mín. 2* foi recusado por apresentar, em média resultados piores que $\mu - \sigma$. As mesmas conclusões podem ser obtidas no segundo gráfico que representa os resultados das Aleatórias.

precisaremos nos preocupar com a comparação dos dois termos e sim com o valor total.

Como já foi dito, extraímos nove imagens pela resposta da passagem de 8 filtros sobre a assinatura (orientações simples) e uma nona relativa a combinação destas. Assim, pode-se executar a comparação sobre qualquer uma destas imagens. Porém

as realizadas com as imagens formadas por orientações simples não obtiveram, na quase totalidade dos casos, valores superiores a da composta. Isto se dá porque elimina-se muitos pontos com a segmentação em particular, pois só resta, naquela imagem, os traços que seguem um sentido. Se a assinatura tiver uma inclinação no ângulo de escrita, ela possuirá uma resposta melhor para esta inclinação, e valores baixos para as outras. Por isso, os valores apresentados pelas comparações globais são resultados de imagens de Respostas Compostas.

Para cada comparação executou-se duas variações, a primeira utilizou um conjunto de modelos unitário, que é apresentado na tabela sob o título “Um Modelo”. Outra experiência foi conduzida com a resposta obtida pela comparação com 3 assinaturas verdadeiras e o resultado final será a média das comparações em separado com cada modelo. Isto é mais custoso, é claro, embora reflita melhor a comparação humana. A única diferença entre os algoritmos é relativa a esta média sendo no resto similares. O objetivo desta segunda é a possibilidade de amenizar a execução da comparação com modelos potencialmente ruins.

A segunda base estudada, por não ter um pré-processamento para eliminar ruído, é mais sujeita a erros, por isso o crescimento dos índices. Em segundo lugar, era provável encontrar valores mais altos nesta faixa devido ao maior número de classes estudadas, como foi comentado na Seção 2.4.5. Outro fator para um resultado pior é o menor número de elementos utilizados na obtenção do limiar (número menor de testes).

6.4.1 Comparação AND

A Tabela 6.1 demonstra os resultados do processamento global sobre os grupos de assinaturas. São indicados os valores dos filtros utilizados.

Se duas assinaturas fossem completamente iguais, esta comparação forneceria

Tamanho de Filtro/FT	0.5/1.5	1.5/3.0	3.0/9.0
	Primeira Base de Dados		
	Um Modelo		
HABILITADA	41.40	27.54	17.25
ALEATÓRIA	35.85	17.29	6.12
	Três Modelos		
HABILITADA	27.96	13.79	10.60
ALEATÓRIA	22.46	7.58	3.81
	Segunda Base de Dados		
	Um Modelo		
HABILITADA	31.68	30.04	22.95
ALEATÓRIA	15.29	15.14	9.43
	Três Modelos		
HABILITADA	48.46	39.45	22.04
ALEATÓRIA	34.92	19.33	5.71

Tabela 6.1: Média dos Resultados FTR para Comparação AND

um valor alto de saída, porém devido às distorções locais, temos saídas menores. No caso, o processamento AND, aplicou-se uma comparação com objetivo de só verificar os pontos que sejam representados no modelo e imagem e eliminar as regiões onde não houve um “ajuste” bom.

Estudando-se os valores da tabela, vê-se que os valores de FTR para as habilitadas são maiores que em relação as aleatórias e se distanciam mais com o aumento da dimensão do filtro. Isto demonstra como a dimensão do mesmo pode auxiliar na verificação, visto que a diferença entre habilitadas e aleatórias cresceu.

Percebe-se, nas tabelas apresentadas, que as respostas à aplicação dos filtros de tamanho pequeno tem valores de erros altos (no caso pior, podendo errar em quase 50%) em relação aos posteriores, que são as respostas a filtros de dimensões crescentes. O filtro age sobre uma assinatura formando uma imagem onde cada ponto contém informação não só sobre si mesmo, mas bem como sobre a área vizinha. Isto aumenta a informação local, permitindo a absorção de pequenas deformações.

Isto decorre porque a proporção que, para cada ponto, o filtro seleciona uma área próxima deste, pontos vizinhos selecionarão áreas aproximadamente similares, pois, suas respostas são influenciadas pela região a qual pertencem, representando áreas que diferem entre si por alguns pixels. Veja a Figura 6.12.

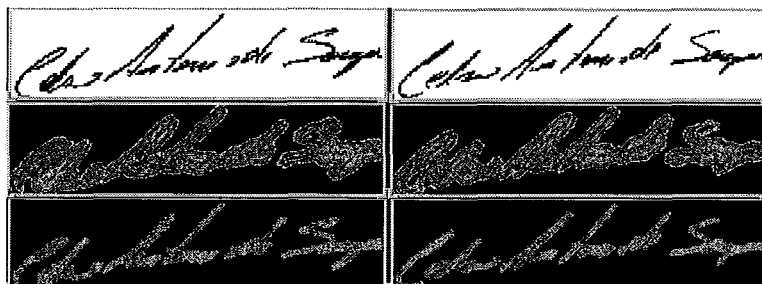


Figura 6.12: Apresenta-se duas assinaturas genuínas, com duas imagens resultantes da filtragem das mesmas, a superior envolve um filtro maior e, a inferior, um menor. Note que para o menor filtro, a área de informação (os pontos cinzas) é menor que a gerada pelo filtro mais alto. Assim, pequenos deslocamentos podem continuar para o primeiro filtro, mas para o segundo, que possui áreas maiores de correspondência, podem ser atenuados.

Quanto maior for o tamanho do filtro, maior será a região que ele irá abranjer, possibilitando que pontos próximos tenham valores mais similares, na medida que a diferença de vizinhança será menor em relação a área que o filtro atua. Com isto, uma comparação global, será auxiliada, devido ao fato que utiliza-se para comparação valores que representam características de regiões provavelmente similares, por diferirem entre si em poucos pixels.

Outro fator importante da utilização de filtros sobre uma imagem é permitir que pontos onde não existia informação, ou seja que não eram pontos formados pelo traçado da caneta na assinatura, apresentassem dados também (seria como se borrasse a imagem da assinatura) formando um “desenho” mais informativo. Isto ocorre porque nas regiões próximas ao traço da assinatura irão apresentar valores de acordo com o tamanho de filtro utilizado.

Por fim, vale a pena salientar que este primeiro algoritmo irá selecionar um con-

junto pequeno de pontos, visto que só os que forem não nulos nas duas imagens serão comparados. Com isto, pode ocorrer de uma assinatura completamente diferente da assinatura original ser aceita, por aqueles poucos pontos apresentarem valores similares aos do modelo. A partir do maior número de pontos que é produzido pelos filtros de maiores dimensões, os valores destes tendem a ser mais realistas. Só a título de informação pode-se apresentar uma média de porcentagem 15% para o filtro menor, cerca de 30% para o médio e 60% para o maior. O problema dos filtros maiores é aumentar os pontos da comparação elevando o custo, além de ter maior tempo de processamento para gerar a imagem.

Para a comparação que verifica a similaridade da assinatura testada com três assinaturas no conjunto de modelos percebe-se bem como um modelo de assinatura “ruim” pode prejudicar o processo. Isto ocorre porque mesmo as assinaturas verdadeiras conterão um conjunto de divergências entre si, e este conjunto pode ser grande ou pequeno, dependendo das várias condições já apresentadas na Seção 2.1.1. Com a utilização de três protótipos teremos é claro os conjuntos de dissimilaridades entre cada uma em particular, mas este conjunto será representado por uma média, que deve ser mais informativa que um valor. E mesmo que entre a assinatura teste e um dos modelos o conjunto de divergências seja grande, provavelmente não será entre todas. Isto pode ser visto nos valores da base 1. No caso da base 2, como possui uma diversidade grande de formatos, onde pode-se encontrar assinaturas pequenas que não fornecem muitos pontos para comparação, além do fato da forma de criação (veja a Seção 6.1.2) permitir altas deformações, o que, possivelmente, produz os resultados piores para o conjunto de 3 modelos em relação aos filtros mais baixos. Com base nesta idéia, o crescimento da dimensão dos filtros aprimora os modelos, determinando uma performance superior a encontrada com um só modelo.

6.4.2 Comparação OU

A Tabela 6.2 apresenta os valores para o algoritmo OU, que seleciona todos os pontos que coexistam nas assinaturas modelo e teste. As observações sobre a atuação do tamanho do filtro da seção anterior continuam, tendo o decréscimo do erro para filtros maiores, isto também será verificado para os algoritmos a frente. O distanciamento entre os valores dos erros com o acréscimo da dimensão dos filtros não ocorre de forma tão acentuada pois os índices dos primeiros filtros já são mais razoáveis. O fato do resultado para o filtro menor já ser bem mais baixo ocorre porque as regiões que não “coincidem” na imagem, e que eram eliminadas na comparação AND, agora são incorporadas fornecendo uma maior informação. Assim, mesmo para filtros pequenos mais pontos são comparados (cerca de 40% já no primeiro filtro e crescendo com o valor do filtro).

Tamanho de Filtro/FT	0.5/1.5	1.5/3.0	3.0/9.0
	Primeira Base de Dados		
	Um Modelo		
HABILITADA	23.85	19.08	15.94
ALEATÓRIA	11.06	7.21	4.23
	Três Modelos		
HABILITADA	12.94	10.81	9.98
ALEATÓRIA	5.27	4.02	3.19
	Segunda Base de Dados		
	Um Modelo		
HABILITADA	26.22	26.37	20.17
ALEATÓRIA	11.31	10.91	7.25
	Três Modelos		
HABILITADA	35.60	26.70	19.15
ALEATÓRIA	12.98	7.58	4.94

Tabela 6.2: Média dos Resultados FTR para Comparação OU

Pode-se concluir que penaliza-se as assinaturas com formato diferente, mas todas discrepâncias entre verdadeiras (leia-se variações livres que existem entre mode-

los, descritos anteriormente) são também somadas negativamente, o que indica seu problema: prejudicar o reconhecimento para assinaturas verdadeiras. Em outras palavras, as deformações impedem que uma característica seja encontrada na assinatura testada. Esta característica distorcida, quando buscada no modelo, também determinará um decréscimo na similaridade.

6.4.3 Comparação TEMP

Os valores desta comparação estão apresentados na Tabela 6.3. A comparação com o TEMP tem a vantagem de pré-determinar o número de pontos que será comparado. Este é igual para o teste em qualquer assinatura, devido ao fato de ser determinado só pelo modelo. Isto é um fator positivo, sobre um aspecto de vista; ele definirá uma esperança de valores melhores, por comparar imagens com um padrão fixo que demonstra a perspectiva de onde provavelmente estarão os pontos, o que favoreceria a eliminação de algumas falsificações habilitadas que não “seguem” corretamente a forma da assinatura.

Para as falsificações aleatórias que não possuam um grande número de características fornece uma baixa taxa de FTR, porém, podem existir aleatórias que possuam um conjunto destas que sejam aceitas, por isso o valor apresentado (além da FAR fornecer também parte do índice).

Continua-se com o problema das assinaturas geradas por um mesmo indivíduo apresentarem grande variações entre si em vários pontos e, neste processo, este fator é enfatizado, sendo a qualidade dos resultados definido somente pela própria qualidade do modelo, como ocorre em qualquer verificação. O problema deste algoritmo é que se a assinatura modelo encontrar um bom conjunto de correspondências com a teste, esta será aceita, ainda que a mesma seja falsa, e possua características adicionais diferentes, pois estas não serão tomadas em consideração para a resposta

Tamanho de Filtro/FT	0.5/1.5	1.5/3.0	3.0/9.0
	Primeira Base de Dados		
	Um Modelo		
HABILITADA	29.19	23.67	17.52
ALEATÓRIA	14.23	8.58	4.73
	Três Modelos		
HABILITADA	17.50	12.73	10.48
ALEATÓRIA	7.38	4.99	3.27
	Segunda Base de Dados		
	Um Modelo		
HABILITADA	26.09	25.37	21.14
ALEATÓRIA	12.61	8.96	7.14
	Três Modelos		
HABILITADA	35.99	29.28	20.61
ALEATÓRIA	8.7	6.57	4.74

Tabela 6.3: Média dos Resultados FTR para Comparação TEMP

do algoritmo.

6.4.4 Comparação

Com a utilização da fórmula de correlação da equação 6.3 ocorre uma comparação que atua em todos os pontos da imagem. Isto irá determinar uma menor dependência da qualidade do modelo, amenizando o problema das variações no conjunto de assinaturas verdadeiras para pequenas mudanças, o grande número destas é que não poderá ser absorvido. O aumento do filtro também não atua de forma tão crítica no reconhecimento, embora ainda auxilie este. Como sua fórmula de comparação (6.3) não compara regiões que possuam zero em um valor e um valor qualquer em outra, ela despreza esta diferença no momento de apresentar a correspondência total. Finalmente, o método tem maior custo que os três anteriores, por comparar todos os pontos das imagens.

Tamanho de Filtro/FT	0.5/1.5	1.5/3.0	3.0/9.0
	Primeira Base de Dados		
	Um Modelo		
HABILITADA	25.58	22.99	17.12
ALEATÓRIA	13.50	9.15	5.38
	Três Modelos		
HABILITADA	18.92	14.33	11.62
ALEATÓRIA	7.38	4.99	3.27
	Segunda Base de Dados		
	Um Modelo		
HABILITADA	27.81	27.14	24.39
ALEATÓRIA	15.59	16.84	11.86
	Três Modelos		
HABILITADA	34.10	27.89	20.33
ALEATÓRIA	8.7	6.57	4.74

Tabela 6.4: Média dos Resultados FTR para Comparação

6.5 Comparação com EGM

Pode-se dizer que a comparação padrão de EGM leva uma certa similaridade com a TEMP acima descrita, ao selecionar alguns pontos para serem comparados, e tem como característica superior uma adaptação local ao formato global. Como esta comparação não obteve bons resultados foi implementada a idéia da *Comparação com Retorno* (Seção 5.3.6) modificando a similaridade para uma comparação OU, que como vimos, tem uma melhor resposta. Embora com valores próximos para algumas comparações globais anteriormente apresentadas, devemos mencionar que os mesmos tiveram um comportamento particular melhor. O que deseja-se dizer com isto é que enquanto comparações globais possuem resultados muito bons para as assinaturas dos grupos que não possuíam uma distorção local grande, para aquelas que o possuíam seus resultados eram bem piores. No caso do EGM, isto não ocorreu, sendo que ele teve um desempenho mais constante por sua própria natureza. Os valores são apresentados na Tabela 6.5.

As comparações locais (veja Seção 5.2.1) foram executadas com a Fórmula 6.2 que apresentou melhores resultado sobre a 6.3 pois como não tínhamos que nos preocupar em comparar o fundo nas diferenças que existem entre o Jet do modelo e o da imagem nos quais um dos termos é zero é computado.

Assinaturas pequenas, ou seja, as que se assemelhavam a rúbricas, tem poucos pontos a serem extraídos, e por isso foram ampliadas para obter-se um número razoável de pontos a fim de executar as comparações. O número de nós por assinatura foi de aproximadamente 60, um número baixo de pontos, como 10 resultou em um índice alto de erro.

O algoritmo só trabalhou sobre a área onde existissem traços da assinatura pois tinha de buscar correlações entre o Jet e regiões escritas, o que define uma estratégia mais rápida de reconhecimento em relação a uma mesma executada com uma busca aberta a todos os pontos da imagem.

Tamanho de Filtro/FT	Um Modelo	Três Modelos	Peso
Primeira Base de Dados			
HABILITADA	12.75	9.67	12.67
ALEATÓRIA	3.00	2.98	2.67
Segunda Base de Dados			
HABILITADA	20.08	15.65	19.19
ALEATÓRIA	8.12	5.07	7.07

Tabela 6.5: Média dos Resultados FTR para Comparação EGM

O objetivo do EGM é abranger todos os possíveis estilos de escrita, mas como não se pode garantir esta cobertura, deve-se procurar métodos que aumentem a chance de cobrir os possíveis estilos devem ser tentados. Com base nesta idéia, utilizou-se mais de um modelo como teste, incrementando a oportunidade de ter modelos mais representativos (três assinaturas devem possuir mais informação que só uma). Entretanto, isto gera um segundo problema. O tamanho da base de modelos é restrito,

por razões de hardware, impedindo que utilize-se dezenas de modelos. Adicionar modelos, também implica em ter-se mais modelos para comparar, aumentando a complexidade do algoritmo e o tempo de execução.

6.5.1 EGM com Peso

Uma característica do EM é que cada ponto do modelo é considerado de igual importância na fórmula original. A diferença total é a soma das diferenças para cada ponto que contribui com uma parte equivalente. É frequentemente desejável que algumas partes sejam mais importantes (ou cruciais) que outras, e permitir que concentrar nosso esforço nestas resulte num melhor desempenho do algoritmo [25]. Com isto, o EGM com Peso, teoricamente, possui mais informação que um só modelo e é mais rápido que três modelos.

Esta noção de realçar a importância de alguns pontos pode ser adicionada em EM pela implementação de pesos [40]. A introdução destes muda o padrão de comparação, ao modificar a contribuição de cada ponto (ou região) para a solução. Agora, os pontos mais importantes contribuirão com um peso maior para a resposta total. Isto quer dizer que penaliza-se mais alguns pontos que outros, na medida que para alguns não são permitidas grandes deformações.

Comparações com peso foram implementadas com uma mudança local no momento de julgar similaridades entre os Jets. O processo segue a seguinte forma:

- O grafo, antes de ser aplicado nas comparações foi processado com um grupo de três modelos. O modelo que obteve uma melhor resposta em relação aos demais foi escolhido para formar o grafo.
- Os Jets do grafo receberam, cada qual, um peso determinado pela similaridade alcançada na comparação com os outros dois modelos. Isto é, executou-se uma

comparação EGM com este grafo sobre os mesmos. No final, com a máscara já com sua maior disposição final de maior similaridade, retira-se o valor da similaridade de Jets e utilizamos estes para compor os índices de pesos.

- Nas comparações posteriores para verificação das faixas de erros, cada Jet tem este peso incorporado na Fórmula 5.3 e multiplicado por λ aumentando o peso da deformação para os Jets que tivessem sido melhor reconhecidos na comparação para obtenção dos pesos, considerando-os mais importantes.

Os resultados preliminares demonstram uma melhora em relação ao valor das comparações simples, mas piores que com três modelos, indicando que deve-se procurar uma melhor forma de representação X codificação das assinatura para conseguir instanciar as informações destas em um modelo.

6.5.2 Global X EGM

A comparação OU teve seus valores, na maioria dos casos, melhores que todas as outras por utilizar a informação topológica não só do modelo, mas também da assinatura testada. Por isso, seus resultados são demonstrados na Tabela 6.6 para a comparação com o EGM. Pode-se ver que os valores do EGM para a primeira base são melhores que a segunda, devido aos fatores descritos na Seção 6.1.2.

Uma pequena diferença em favor do EGM aparece, principalmente na primeira base. O algoritmo de EGM além de possuir um custo computacional menor, permite que execute a Comparação de EGM com peso, que assimila mais informação das assinaturas.

Primeira Base de Dados					
Número de Modelos	Um Modelo		Três Modelos		Peso
Comparação	Global	EGM	Global	EGM	Peso
HABILITADA	15.94	12.75	9.98	9.67	12.67
ALEATÓRIA	4.23	3.00	3.19	2.98	2.67
Segunda Base de Dados					
Número de Modelos	Um Modelo		Três Modelos		Peso
Comparação	Global	EGM	Global	EGM	Peso
HABILITADA	20.17	20.08	19.15	15.65	19.19
ALEATÓRIA	7.25	8.12	4.94	5.07	7.07

Tabela 6.6: Comparação dos Resultados FTR para as Comparações Globais (representada pelo algoritmo OU) e EGM

Capítulo 7

Conclusões e Trabalhos Futuros

“Primeiro apure os fatos.

Depois, pode distorcê-los à vontade.”

Mark Twain

O problema de verificação de assinaturas pode ser definido da seguinte forma: apresentada uma assinatura e supondo conhecer a identidade do autor desta, via número da conta, nome, ou etc., verificar se a assinatura é realmente da pessoa ou uma falsificação.

A tarefa de reduzir a medida de discrepância entre assinaturas verdadeiras e produzir uma forte separação entre as genuínas e falsificações para cada pessoa não é fácil de se realizar, pois é difícil determinar quais as características que discriminam cada uma destas classes, mesmo porque a classe de falsificações que podem ocorrer não é totalmente conhecida.

A descrição da imagem de uma assinatura pode envolver propriedades da mesma ou suas partes e relações entre as mesmas. Na proposta apresentada nesta tese, características locais são processadas com filtros orientados que são inspirados em aspectos da visão, e fornecem indicações da orientação dos traços em cada ponto. Para ser consistente com a fisiologia da visão humana, várias resoluções da repre-

sentação de forma são utilizados para fornecer uma representação retinotópica das imagens para vários graus de escalas.

Duas formas de estudo foram implementadas. A primeira estudou, com características globais, como a aplicação e o tamanho dos filtros interfere na representação das imagens, constatando que filtros maiores possuíam a capacidade de minimizar os efeitos de deformações locais existentes entre as assinaturas genuínas. Assim, quanto maior o filtro melhor a saída. Esta afirmação envolve problemas, como o número de pontos comparados, que cresce com o tamanho do filtro, além do processamento necessário para gerar a imagem filtrada que é proporcional ao tamanho do filtro.

Em um segundo método, os testes com comparações elásticas são também importantes no sentido de proporem uma metodologia compacta da descrição derivada do conteúdo pictorial, com um número de pontos menor a ser buscado, e uma procura mais “inteligente” dos pontos representados na imagem testada. Esta procura consegue absorver pequenas deformações locais das imagens.

Como as características extraídas da assinatura são particulares ao formato buscado utilizou-se uma inovação do método original de EGM que é a comparação de retorno. Esta utiliza características da imagem testada para serem buscadas no modelo, oferecendo a possibilidade de verificar se o objeto testado não possui características que possam não ser encontradas no modelo, o que configuraria uma falsificação que não conseguiu “esconder” suas características, veja a Seção 2.1.1.

Um ponto que se deve indicar é que a técnica de EGM, por gerar um “alinhamento” entre os pontos do modelo e imagem testada, pode ser utilizada como um indicador de quais regiões se relacionam em um algoritmo dinâmico, além de poder ser processado em paralelo com o mesmo, contribuindo para uma saída final mais complexa. Vejamos melhor isto: quando aplicamos o algoritmo de EGM, busca-se

encontrar a melhor correspondência de similaridade entre os pontos do modelo e imagem testada. Assim, ao término do processo existe uma relação de quais pontos são correlacionados entre as duas assinaturas. Com base neste dado, os algoritmos dinâmicos podem alinhar os segmentos destas.

O algoritmo de Comparação Global pode ter um desempenho quase similar ao do EGM para os melhores resultados, por se adaptar melhor a formas de modelo. Porém o EGM permite que informações de mais de uma assinatura sejam incorporadas ao modelo para se obter uma melhor performance nos resultados. O EGM é um algoritmo que necessita de um grafo modelo que represente bem a assinatura a fim de poder realizar a verificação, porém a forma de criar um bom grafo não é muito simples para assinaturas, visto que podem existir assinaturas que sejam similares a rubricas que forneçam poucos pontos para comparação.

A memória necessária para armazenar os modelos também deve ser salientada, no caso da Comparação Global, toda a assinatura deve ser armazenada, mas com o EGM, os Jets, com suas “ligações” relativas já tem a informação necessária, acarretando em um decréscimo de espaço de memória.

Os custos computacionais entre as Comparações globais e EGM podem ser vistos no número de pontos que cada um trabalha, enquanto o EGM atua somente na área da assinatura, e por isso, utiliza menos pontos, a Comparação Global tem de utilizar vários pontos que não pertencem a assinatura, mas que são informativos pela passagem dos filtros maiores. Também deve ser notado a forma de pré-processamento pela filtragem, enquanto a Comparação Global envolve a passagem de filtros por toda imagem, o EGM permite que só os pontos onde há tracos sejam processados, embora tenha de passar pela mesma um número maior de filtros.

Antes de comparar os métodos utilizados nesta tese com alguns trabalhos, vale

a pena realçar as condições que aparecem na Seção 2.4.5. Mesmo com esta informação, pode-se perceber que os principais resultados obtidos nesta tese estão apresentados na Tabela 6.6 e se apresentam próximos dos valores da literatura, indicados na Tabela 3.1. Isto demonstra que novos estudos podem ser efeturados para aprimorar os algoritmos, principalmente o de EGM, a fim de se conseguir um melhor desempenho.

7.1 Trabalhos Futuros

Devido aos testes empregados, várias propostas podem ser implementadas para buscar melhores resultados. A primeira pode ser simplesmente aplicar o pré-processamento baseado em Filtros Orientados nas entradas de outros algoritmos, como os descritos no Capítulo 3. Em trabalhos que utilizem os mesmos algoritmos aqui apresentados, podem ser executados alguns itens, como:

- Novas dimensões de filtros podem ser implementadas.
- Outros filtros podem ser utilizados para representar as características dos traços, como exemplo, temos os “Banana-wavelets” que podem indicar as curvas das assinaturas [55].
- Um estudo sobre assinaturas que não fossem pré-processadas no sentido de transformá-las para uma escala similar pode ser realizado com as comparações globais, desde que aplique-se um método de relacionar regiões; assim como no EGM, onde só precisaríamos implementá-lo com um movimento global que adotasse a busca com deformações rígidas.
- Uma estrutura que busque características globais em regiões processadas com filtros pode ser implementada.

Devido a modularidade do algoritmo de EGM, há grandes possibilidades particulares para o mesmo, além das acima citadas, de implementar modificações e descobrir novas respostas sem modificar a idéia principal.

- Novas formas de busca podem ser implementadas no EGM. Por exemplo, utilizando-se uma busca partindo de um subconjunto inicial bom, descrita na seção 5.2.1.
- Outras formas de busca para encontrar o mínimo global com o gradiente descendente, ou “simulated annealing” onde o tamanho da área diminuiria durante o processo, podem ser testadas (veja a Seção 5.2.1).
- As comparações entre elásticos podem ser relativas. Ao ocorrer uma rotação linear em relação a extensão da assinatura, as posições dos nós e dos elásticos sofrerá um deslocamento linear também. Assim, ao se comparar as deformações dos elásticos de uma forma menos rígida, estudando uma compensação, melhores soluções podem ocorrer.
- A formação de um modelo de pesos mais robusto pode ser elaborada. O conjunto de assinaturas pode prover um modelo que não seja uma representação da topologia de uma assinatura em particular, mas sim de um conjunto destas. Por exemplo, pode-se tentar iniciar com um grafo extraído de uma assinatura e, aplicando-o sobre outros modelos, incorporar novos nós ou deletar alguns, assim como modificar a disposição dos elásticos.

Referências Bibliográficas

- [1] GOMES, H. M., FILHO, E. C. B. C. “Reconhecimento Off-line de Assinaturas Utilizando MLP-Backpropagation e Momentum: um Estudo Comparativo.” Proceedings on I Congresso Brasileiro de Redes Neurais, Escola Federal de Engenharia de Itajubá, Itajubá, pp 133-138, outubro, 1994.
- [2] MCCABE, A., *A Review of Dynamic Handwritten Signature Verification*. Honours Project, Department of Computer Science, James Cook University, Townsville, Austrália, Set. 1997.
- [3] SENIOR, A. W., *Off-line Handwriting Recognition: A Review and Experiments*, Technical Report CUED/F-INFENG/TR 105, Dec 1992.
- [4] PLAMONDON, R., LORETTE, G., “Automatic Signature Verification and Writer Identification - The State of the Art”, *Pattern Recognition*, v. 22, pp. 107-131, 1989.
- [5] MILLER, B., “Vital Signs of Identity”, *IEEE Spectrum*, pp. 22-30, V. 31 N. 2, Feb. 1994.
- [6] WISKOTT, L., FELLOUS, J.-M., KRÜGER, *et al.*, “Face Recognition by Elastic Bunch Graph Matching.” *IEEE Transactions on Pattern Analysis and Machine Intelligence* v. 19, n. 7, pp. 775-779, Jul. 1997.

- [7] EXEL, S., PESSOA, L., “Space-variant representation for active object recognition.” Proceedings of the International Symposium on Computer Graphics, Image Processing and Vision, SIBGRAP’98, Rio de Janeiro, Brasil, 1998.
- [8] PESSOA, L., EXEL, S., ROQUE, A., LEITÃO, A., “Atentive Vision Recognition for Scene Exploration.” Proceedings of the International Conference on Neural Information Processing, ICONIP’98, Japão, 1998.
- [9] HERBEST, N. M., LIU, C. N., “Automatic Signature Verification Based on Accelerometry” *IBM Journal of Research Development*, V. 21, N. 3, pp 245-253, May 1977.
- [10] AMMAR, M. , YOSHIDA, Y., FUKUMURA, T., “Structural Description and Classification of Signature Images” *Pattern Recognition*, v.23, n. 7, pp. 697-710, 1990.
- [11] HILTON, O., “Signatures - Review and a New View” *Journal of Forensic Sciences*, v.37, n. 1, pp. 125-129, Jan 1992.
- [12] SCHALKOFF, R., *Structural and Neural Approaches.*, John Wiley & Sons, Inc, USA, 1992.
- [13] OJA, E., LAMPINEN, J. “Unsupervised Learning for Feature Extraction.”, *In: Computational Inteligence Imitating Life*, 1 ed., IEEE Press, pp. 13-21, 1994.
- [14] PESSOA, L. A. F. C., *Aprendizado Não-Supervisionado em Redes Neurais*. Tese de M. Sc., COPPE-UFRJ, Rio de Janeiro, RJ, Brasil, Jul. 1990.
- [15] QI, Y., HUNT, B. R., “Signature Verification Using Global and Grid Features.” *Pattern Recognition*, v. 27, n. 12, pp 1621-1629, 1994.
- [16] LING, L. L., “Sistemas Automáticos de Verificação de Assinaturas Dinâmicas em Tempo Real” *SBA Controle & Automação* v. 8, n. 3, pp 95-104, Dez. 1997.

- [17] DROUHARD, J. P., SABOURIN, R., GODBOUT M., "A Neural Network Approach to Off-line Signature Verification Using Directional PDF" *Pattern Recognition*, v.29, n. 3, pp. 415-424, 1996.
- [18] LECLERC, F., PLAMONDON, R., "Automatic Signature Verification the State of The Art -1989-1993" *International Journal of Pattern Recognition and Artificial Intelligence* v. 8, n. 3, pp. 643-660, 1994.
- [19] YANG, L., WIDJAJA, B. K., PRASAD, R., "Application of Hidden Markov Models for Signature Verification." *Pattern Recognition*, v. 28, n.2, pp 161-170, 1995.
- [20] NAGEL, R. N., ROSENFELD, A., "Computer Detection of Freehand Forgeries" *IEEE Transactions on Computers* v. C-26, n. 9, pp 895-905, Set. 1977.
- [21] LEE, S., PAN J. C., "Offline Tracing and Representation of Signatures" *IEEE Transactions On systems, Man, and Cybernetics* v. 22, n. 4, pp 755-771, Jul-Ag. 1992.
- [22] SABOURIN, R., GINETTE, G, PRÊTEUX, F. J., "Off-line Signature Verification by Local Granulometric Size Distributions" *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 19, n. 9, pp. 976-987, Set. 1997.
- [23] NEMCEK, W. F., LIN, W. C., "Experimental Investigation of Automatic Signature Verification" *IEEE Transactions On systems, Man, and Cybernetics*, pp. 121-126, Jan. 1974.
- [24] RIGOLL, G., KOSMALA, A., "A Systematic Comparasion Between On-line and Off-line Methods for Signature Verification with Hidden Markov Models" *Proceedings of 14th International Conference on Pattern Recognition*, pp. 1755 a 1757, Brisbane, Austrália, 1988.

- [25] BRAULT, J., PLAMONDON, R., "Segmenting Handwritten Signatures at Their Perceptually Important Points", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v.15, n. 9, pp 953-957, Set. 1993.
- [26] RIBEIRO, J. N. G., VASCONCELOS, G. C., "Verificação Off-line de Assinaturas Baseada em uma Arquitetura Cascade-Correlation Autoassociativa." *In: Anais do V Simpósio Brasileiro de Redes Neurais*, pp. 305-310, Belo Horizonte, M.G., Brasil, Dez. 1998.
- [27] LEE, L. L., BERGER, T., AVICZER, E., "Reliable On-line Human Signature Verification Systems" *IEEE Transactions on Pattern Analysis and Machine Intelligence* v. 18, n. 6, pp 643-647, Jun 1996.
- [28] PARIZEAU, M., PLAMONDON, R. "A Comparative Analysis of Regional Correlation Dynamic Time Warping, and Skeletal Tree Matching for Signature Verification." *IEEE Transactions on Pattern Analysis and Machine Intelligence* v. 12 n. 7, pp. 710-717, Jul 1990.
- [29] LIU, C. N., HERBST, N. M., "Automatic Signature Verification: System Description and Field Test Results", *IEEE Transactions on Systems, Man, and Cybernetics*, v. 9, n. 1, pp. 35-38, Jan. 1979.
- [30] PLAMONDON, R., MAARSE, F. J., "An Evaluation of Motor models of Handwriting. *IEEE Transactions on Systems, Man, and Cybernetics*", v. 19, n. 5, pp. 1060-1072, Set./Out. 1989.
- [31] CRAIZER, M., SILVA, E. A. B. , "Processamento Digital de Imagens" *XXI Congresso Nacional de Matemática Aplicada e Computacional*, SBMAC, Caxambu, MG, 1998.

- [32] MASCARENHAS, N. D. A., VELASCO, F. R. D., *Processamento Digital de Imagens*, IV Escola Brasileiro-Argentina de Informática (IEBAI), Universidade Católica de Santiago del Estero, Kapelus S. A., 2 ed., Buenos Aires, Argentina, Jan 1989.
- [33] PAZ, E. P., CUNHA, T. N., *Iniciação ao Processamento Digital de Imagens*. VII JAI-Jornada de Atualização em Informática, VIII Congresso da Sociedade Brasileira de Computação, Rio de Janeiro, jul. 1988.
- [34] NIBLACK, W., *An Introduction to Digital Image Processing*. Prentice Hall International, New Jersey, USA, 1986.
- [35] NALWA, V. S., *A Guided Tour Of Computer Vision* Addison Wesley, 1993.
- [36] PESSOA, L. A. F. C. . “Introduction to the Boundary Contour System/Feature Contour System Theory of Vision”, Artigo não publicado, 1997.
- [37] KEFALEA, E., REHSE, O., MALSBERG, C. v.d., “Object Classification Based on Contours with Elastic Graph Matching. In: *Proceedings of the 3rd International Workshop on Visual Form-IWVF3 1997*, World Scientific, Singapore, 1997.
- [38] BASRI, R., COSTA L., GEIGER D., JACOBS D., “Determining the similarity of deformable shapes”, *Vision Research*, v.38, pp. 2365-2385, 1998.
- [39] WÜRTZ, R., “Object Recognition Robust Under Translations, Deformations, and Changes in Background.” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 19, n. 7, pp 769-775, Jul. 1997.
- [40] SCATTOLIN, P., *Recognition of Handwritten Numerals Using Elastic Matching*. M. Sc. dissertation, Concordia, Quebec, Canada, Nov. 1995.

- [41] NIEMANN, H., *Patter Analysis and Understaning*. Springer Series in Information Science., V. 4, 2 ed., Berlin, Germany,1990.
- [42] TAPPERT, C. C., "Cursive Script Recognition by Elastic Matching" *IBM J. Res. Develop.* v. 26, n. 6, pp 765-770, Nov. 1982.
- [43] LAM, L., SUEN, C. Y., "A dynamic shape preserving thinning algorithm" *Signal Processing*, v. 22, pp. 199-208, 1991.
- [44] BAJA, G. S. d., THIEL, E., "Skeletonization algorithm running on path-based distance maps" *Imagem and Vision Computing*,v. 14, pp. 47-57, 1996.
- [45] LAM, L., LEE, S.-W., "Thining Methodologies- A comprehensive Survey", *IEEE Transactions on Pattern Analysis and Machine Intelligence* v. 14, n. 9, pp. 869-885, 1992.
- [46] WISKOTT, L., "Phantom Faces for Face Analysis" In: *Proceedings of 7th. International Conference on Computer Analysis of Images and Patterns*. Kiel, Germany, Set. 1997.
- [47] WISKOTT, L., "Phantom Faces for Face Analysis" *Pattern Recognition.* v. 30, n. 6, pp. 837-846, 1997.
- [48] PÖTZSCH, M, KRÜGER, N., MALSBURG, C. v. d., "Improving Object Recognition by Transforming Gabor Filter Responses", *Network: Computation in Neural Systems*, v. 7, n. 2, pp. 341-347, 1996.
- [49] WISKOTT, L., FELLOUS, J.-M., KRÜGER, N., *et al.*, "Face Recognition and Gender Determination", In:*Proceedings of International Workshop on Automatic Face- and Gesture-Recognition, IWAFGR'95*, pp. 92-97, ed. Martin Bichsel, publ. MultiMedia Laboratory, University of Zurich, Zurich, Switzerland, Jun. 1995.

- [50] FISCHLER, M. A., ELSCHLAGER, R. A., “The Representation and Matching of Pictorial Structures” *Computer Methods in Image Analysis.*, pp. 31-57 , IEEE Press, N.Y., USA, 1977.
- [51] ROSENFELD, A., KAT, A. C. *Digital Picture Processing.*, v. 1, *Computer Science and Applied Mathematics*, 2 ed., Academic Press Inc., San Diego, USA, 1982.
- [52] DOURSAT, R., KONEN, W., LADES, M. , et. al., *Neural Mechanisms of Elastic Pattern Matching* Internal Report IR-INI 93-01, Institut für Neuroinformatik, Ruhr-Universität Bochum, D-44780 Bochum, Germany, Jan. 1993.
- [53] RUSSELL, S., NORVIG, P. *Artificial Intelligence: A Modern Approach.* Prentice Hall Series in Artificial Intelligence, New Jersey, 1995.
- [54] LADES, M., VORBRÜGGEN, J. C., BUHMA, J., et al., “Distortion Invariant Object Recognition in the Dynamic Link Architecture” *IEEE Transactions On Computers*, v. 42, n. 3, pp. 300-311, Mar. 1993.
- [55] PETERS, G., KRÜGER, N., “Learning Object Representations by Clustering Banana Wavelet Responses”, In. *Proceedings of the 1st International Workshop on Statistical Techniques In Pattern Recognition (STIPR97)*, P. Pudil, J. Novovicova, and J. Grim, editors, pp. 113-118, Prague, June 1997.