

Universidade Federal do Espírito Santo
Centro Tecnológico
Programa de Pós-Graduação em Engenharia Elétrica

Jorge Leonid Aching Samatelo

**Um Sistema de Classificação de Impressões
Digitais**

Vitória
2007

Jorge Leonid Aching Samatelo

Um Sistema de Classificação de Impressões Digitais

Dissertação submetida ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Espírito Santo como requisito parcial para a obtenção do grau de Mestre em Engenharia Elétrica, na área de concentração em Automação.
Orientador: Prof. Dr. Evandro Ottoni Teatini Salles.

Vitória
2007

Dados Internacionais de Catalogação-na-publicação(CIP)
(Biblioteca Central da Universidade Federal do Espírito Santo, ES, Brasil)

Samatelo, Jorge Leonid Aching, 1975-
S187s Um sistema de classificação de impressões digitais /
Jorge Leonid Aching Samatelo. - 2007.
113 f. : il.

Orientador: Evandro Ottoni Teatini Salles.
Dissertação (mestrado): Universidade Federal do
Espírito Santo, Centro Tecnológico.

1. Biometria. 2. Datiloscopia. 3. Classificação. 4.
Morfologia. 5. PCASYS (Sistema classificador de
impressões digitais). I. Evandro Ottoni Teatini Salles.
II. Universidade Federal do Espírito Santo. Centro
Tecnológico. III. Título.

CDU: 621.3

Jorge Leonid Aching Samatelo

**Um Sistema de Classificação de Impressões
Digitais**

Dissertação submetida ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Espírito Santo como requisito parcial para a obtenção do grau de Mestre em Engenharia Elétrica - Automação.

Aprovada em 23/08/2007.

COMISSÃO EXAMINADORA

Prof. Dr. Evandro Ottoni Teatini Salles
Universidade Federal do Espírito Santo
Orientador

Prof. Dr. Thomas Walter Rauber
Universidade Federal do Espírito Santo

Prof. Dr. Jugurta Rosa Montalvão Filho
Universidade Federal de Sergipe

Vitória, ES - BRASIL
Agosto de 2007

“Cada homem é sua própria prisão, mas cada homem pode também conquistar o poder de fugir... Não deixem de lutar”.
Sidarta Gautama-Buda

Dedico esta tesis a mi Abuela Úrsula por estar siempre en mi memoria, a mi Madre por enseñarme a predicar con el ejemplo y no con la palabra, a mi Padre por demostrarme que en esta vida o luchas o luchas y a mi hermano por demostrarme el valor de algo tan sublime como la música.

Agradecimentos

Gostaria de agradecer ao Prof. Evandro por todo apoio dado, e sua judiciousa direção na elaboração deste trabalho. Ao Prof. Sarcinelli por a oportunidade e confiança oferecida. Ao prof Cifuentes por ser um grande amigo. A meus amigos Klaus, Glauber, Janayna, Alexandre, Patrick e Gustavo por ajudar-me na correção da tese. A meus tios Carlos, Zoila e Helga por toda ajuda proporcionada. A meus pais Cesar e Marlene, a meu avô Domingo e a Gabriela porque, não obstante a distância, sempre estiveram presentes.

Resumo

A classificação de impressões digitais é uma tarefa desafiante para o reconhecimento de padrões, despertando o interesse de vários pesquisadores durante os últimos 30 anos. Um número de técnicas e várias estratégias de extração de características foram propostas, desde técnicas baseadas em regras até o uso de sistemas de múltiplos classificadores. Porém, entre as diversas abordagens, a mais promissora é o sistema PCASYS proposta pelos pesquisadores do NIST (*National Institute of Standards and Technology*)

Neste trabalho foi implementado um sistema de classificação de impressões digitais tomando como ponto de partida a arquitetura proposta no PCASYS. O sistema implementado divide-se em duas partes principais: uma dedicada a processar a impressão digital e extrair suas características representativas e outra orientada a realizar a tarefa de classificação. A premissa inicial foi elaborar técnicas robustas na parte de extração de características de maneira que o rendimento do classificador seja maximizado. O banco de imagens utilizado para testar o sistema de classificação implementado foi a banco de imagens de impressões digitais número 4 do NIST, o qual é um padrão na área. O classificador que obteve os melhores resultados foi um classificador Bayessiano Normal Multivariável. Os bons resultados alcançados pela parte de classificação provam que a premissa é válida. Os resultados são comparáveis com outras abordagens na literatura.

Abstract

The fingerprint classification task is a challenging problem to the area of pattern recognition. Many techniques and strategies for feature extraction were proposed, since techniques based on rules to multiple classifiers. The most prominent approach is the PCASYS system, proposed by the NIST researchers (National Institute of Standards and Technology).

In this work, it is implemented a fingerprint classification system based on the PCASYS architecture. The system was divided into two main parts: one of them dedicated to processing and feature extraction of fingerprints and the other one dedicated to the classification task. We suppose that a robust preprocessing and feature extraction stage maximizes the performance of a classifier. The image data bank used was based on the NIST number 4 database, what is benchmark for on the fingerprint classification area.

The best results were obtained by using a Gaussian Multivariate Bayesian classifier and the error rates were compared to other approaches in literature.

Sumário

I	Introdução	1
1	Descrição da Tese	2
1.1	Introdução	2
1.2	Caracterização do Problema	3
1.3	Objetivos	4
1.4	Organização da dissertação	4
II	Fundamentação Teórica	6
2	Definições e Estado do Arte	7
2.1	Introdução	7
2.2	Impressões Digitais	7
2.2.1	Características	7
2.3	Classificação de impressões digitais	10
2.4	Extração de características	16
2.5	Técnicas de classificação	18
2.5.1	Técnicas baseadas em regras	18
2.5.2	Técnicas sintáticas	20
2.5.3	Técnicas Estruturais	21
2.5.4	Técnicas Estatísticas	22
2.5.5	Técnicas baseadas em redes neurais	23
2.5.6	Técnicas Baseadas em vários classificadores	26
2.6	Resumo	27
III	Implementação	30
3	Processamento da impressão digital	31
3.1	Introdução	31
3.2	Cálculo da imagem de regiões	32

3.2.1	Técnica de segmentação baseada no mínimo valor de intensidade	34
3.3	Melhora da imagem	38
3.3.1	Normalização	39
3.3.2	Melhora no domínio da Freqüência	40
3.4	Cálculo da imagem de orientação	41
3.4.1	Obtenção da imagem de orientação	43
3.5	Registro de impressões digitais	48
3.5.1	Determinação do ponto de referência de uma impressão digital	48
3.6	Resumo	62
4	Classificação	64
4.1	Introdução	64
4.2	Classificação Estatística	65
4.2.1	Geração de Características	65
4.2.2	Redução dimensional	66
4.2.3	Classificação	69
4.3	Resumo	77
IV	Parte Experimental	79
5	Testes e resultados	80
5.1	Introdução	80
5.2	O Ambiente Desenvolvido	80
5.3	Banco de Dados	81
5.4	Rendimento das técnicas de classificação de impressões digitais	84
5.5	Resultados prévios obtidos sobre a DB4	84
5.6	Resultados	86
5.7	Resumo	98
6	Conclusões e Projetos Futuros	100
6.1	Conclusões	100
6.2	Temas a serem pesquisados	102

Lista de Figuras

2.1	Impressão Digital.	8
2.2	Área Padrão e Linhas Tipo.	9
2.3	Pontos singulares (a) configurações do ponto <i>Core</i> ; (b) configurações do ponto <i>Delta</i>	10
2.4	Tipos Minúcias (a) bifurcação; (b) finalização.	10
2.5	Seis classes das impressões digitais (a) <i>Arch</i> ; (b) <i>tented arch</i> ; (c) <i>right loops</i> ; (d) <i>left loops</i> ; (e) <i>plain whorl</i> ; (f) <i>twin loop</i> . Onde o ponto <i>core</i> é indicado pelo círculo e o ponto <i>delta</i> é indicado pelo triângulo.	12
2.6	O esquema de classificação adotado pelo FBI (a) <i>Plain Arch</i> ; (b) <i>Tented Arch</i> ; (c) <i>Radial Loop</i> ; (d) <i>Ulnar Loop</i> ; (e) <i>Plain Whorl</i> ; (f) <i>Double Loop</i> ; (g) <i>Central Pocket</i> ; (h) <i>Accidental Whorl</i>	13
2.7	Três impressões digitais que pertencem a diferentes classes e tem a mesma aparência (pequena variabilidade inter-classes). (a) <i>arch</i> ; (b) <i>Tented Arch</i> ; (c) <i>left loop</i>	14
2.8	Três impressões digitais pertencentes à mesma classe que apresentam características muito diferentes (grande variabilidade intra-classes). (a) <i>twin loop</i> ; (b) <i>Plain Whorl</i> ; (c) <i>Central Pocket</i>	14
2.9	Exemplos de imagens de impressões digitais ruidosas (a) devidas às condições da pele; (b) devidas à técnica de aquisição.	15
2.10	Características usadas com maior freqüência para a classificação de impressões digitais: (a) imagem de orientação; (b) linhas de fluxo de crista; (c) pontos singulares; (d) resposta dos filtros de Gabor.	16
2.11	Técnica de classificação baseada em grafos inexatos [65].	21
2.12	Esquema funcional do PCASYS para a classificação exclusiva.	26
3.1	Representação da imagem como uma superfície tridimensional ou como uma matriz de pontos (píxels) que variam no intervalo de 0 a 255 (escala de cinza).	32
3.2	Regiões nas impressões digitais, (a) imagem original; (b) imagem de regiões (zonas de cor preto são definidas como segundo plano e as brancas como primeiro plano).	32

3.3	Setorização da imagem em blocos retangulares.	35
3.4	(a)Imagem original; (b)imagem equalizada localmente; (c)matriz de valores mínimos antes de filtrar; (d)matriz de valores mínimos depois de filtrar.	37
3.5	Resultados da segmentação. As linhas brancas indicam as fronteiras entre o primeiro e segundo plano definido por Π	38
3.6	Resultados da Normalização: (a) I ; (b) $I_{U_{nf}}$; (c) I_{Nor}	40
3.7	Melhora no domínio da frequência (a) imagem normalizada; (b) imagem melhorada.	42
3.8	Imagem de orientação de uma impressão digital, cada linha representa a orientação média das cristas numa vizinhança de tamanho definido. . .	43
3.9	Imagem de orientação Θ para uma impressão de boa qualidade (a)impressão digital; (b)técnica padrão; (c)técnica proposta em [97]; (d)técnica proposta em [52].	47
3.10	Imagem de orientação Θ para uma impressão de má qualidade (a)impressão digital; (b)técnica padrão; (c)técnica proposta em [97]; (d)técnica proposta em [52].	48
3.11	Ponto de referência numa impressão digital.	49
3.12	Curvas fechadas para a obtenção do índice de Poincaré.	51
3.13	Cálculo do índice de Poincaré para $r = 1$	52
3.14	Detecção de singularidades através do índice de Poincaré (a)imagem de orientação; (b) índice de Poincaré; (c) singularidades válidas; (d); singularidades determinadas (pontos são <i>core</i> e triângulos são <i>deltas</i>).	53
3.15	Detecção do ponto de referência para imagens tipo A (a)imagem de orientação; (b) transformação de covariância da imagem de orientação; (c)componente conectado de maior área da imagem binária da transformação de covariância; (d) ponto de referência da impressão digital tipo A.	54
3.16	Distribuição das cristas sobre uma vizinhança do ponto de referência. . .	56
3.17	(a) I ; (b) Θ quantizado em 32 níveis; (c) Θ_0^B ; (d) Θ_1^B ; (e) Θ_2^B ; (f) Θ_3^B	57
3.18	Imagens Θ_d^B características para as impressões digitais tipo (a)A; (b)TA; (b)RL; (c)LL; (d)W; (e)TL, quando $d = 1$ e $Q = 16$	58
3.19	Imagens Θ_d^C para (a) $d = 0$; (b) $d = 1$; (c) $d = 2$; (d) $d = 3$;	58
3.20	Determinação da matriz binária de pontos médios B_m e as longitudes dos conjuntos de 1 de Θ_d^B ao realizar um percorrido de suas linhas. . . .	60
3.21	(a) Θ quantizado em 32 níveis; (b) Θ_0^B ; (c) Θ_1^B ; (d) Θ_2^B ; (e) Θ_3^B ; (f) B_T ; (g) recálculo do ponto de referência utilizando a matriz de <i>Coh</i> ; (h)ponto de referência final.	62
3.22	Arquitetura da etapa de processamento.	63
4.1	Representação Vetorial de Θ , (a) Θ ; (b) Θ_{reg} ; (c) v ; (d) z	66

4.2	(a)Matriz de observações \mathbf{Z} ; (b)valor médio de \mathbf{Z} , \mathbf{u} ; (c)observações projetadas \mathbf{x} ; (d) matriz de transformação \mathbf{W}	69
4.3	Limites da taxa de erro do k -NN num problema de c classes, dado um número infinito de dados de treino [31].	72
4.4	Topologia de um MLP de três camadas.	74
4.5	Topologia de um PNN.	77
4.6	Arquitetura da etapa de classificação.	78
5.1	Imagens de má qualidade.	82
5.2	Imagens de média qualidade.	83
5.3	Imagens de boa qualidade.	83
5.4	Imagens que apresentam a super segmentação.	87
5.5	(a)impressão digital, (b)matriz de coerência.	89
5.6	A função discriminante do MLP (linha pontuada) passa por todos os pontos de treino (círculos cinza), porém a função discriminante do classificador normal (linha fina) permite uma melhor predição para as novas amostras [31].	96

Lista de Tabelas

2.1	Cronologia das diferentes técnicas de classificação de impressões digitais, cada trabalho é etiquetado de acordo as características usadas (O = imagem de orientação, S = singularidades, R = linhas de fluxo das cristas, G = operadores de Gabor) pelas técnicas de classificação propriamente ditas (Rb = uso de base de regras, Sy = técnicas sintáticas, Str = estruturais, Sta = estatísticas, Nn = redes neurais, Mc = múltiplos classificadores) [67].	29
5.1	Distribuição das impressões digitais na DB4.	82
5.2	Distribuição das impressões digitais da DB4 segundo sua qualidade.	84
5.3	Matriz de confusão das técnicas propostas por : (a) Senior [90] para o problema de classificação em 4 classes e por (b) Capelli, Maio e Maltoni [87] para o problema de classificação em 5 classes, ambas técnicas foram testadas sobre o banco de dados DB4.	85
5.4	Taxas de erro de 14 técnicas testadas com a DB4 para os problemas de 4 e 5 classes, em que o rótulo N significa que os resultados da classificação são ponderados de acordo a distribuição 'natural' das classes e o rótulo B significa que os resultados da classificação não são ponderados, posto que assume-se uma distribuição 'balanceada' de cada classe.	86
5.5	Taxas de erro de precisão quando utiliza-se a técnica de estimação local da orientação padrão, em que o rótulo (MC) faz referência as impressões de má qualidade, o rótulo (MD) faz referência às impressões de média qualidade e o rótulo (BC) faz referência às impressões de boa qualidade.	90
5.6	Taxas de erro de precisão quando utiliza-se a técnica de estimação local da orientação proposta por [97].	90
5.7	Taxas de erro de precisão quando utiliza-se a técnica de estimação local da orientação proposta por [52].	90
5.8	Taxa de erro para os problemas de classificação de 4 e 5 classes, quando é usado um classificador normal, para diferentes valores de n (número de características).	93

5.9	Taxa de erro para os problemas de classificação de 4 e 5 classes, quando é usado um classificador k -NN, para diferentes valores de n	93
5.10	Taxa de erro para os problemas de classificação de 4 e 5 classes, quando é usado um classificador MLP, para diferentes valores de n	93
5.11	Taxa de erro para os problemas de classificação de 4 e 5 classes, quando é usado um classificador PNN, para diferentes valores de n	94
5.12	Matrizes de confusão quando é usado um classificador normal para os problemas de: (a)4 classes com $n = 38$ e registro automático; (b)5 classes com $n = 38$ e registro automático; (c)4 classes com $n = 32$ e registro semi-automático; (d)5 classes com $n = 38$ e registro semi-automático.	97
5.13	Matrizes de confusão quando é usado um classificador normal para os problemas de 5 classes com $n = 38$, em que cada matriz corresponde: (a)as impressões mal segmentadas; (b)as impressões de má qualidade; (c)as impressões de média qualidade; (d)as impressões de boa qualidade; (e)as impressões mal registradas; (d)as impressões mal registradas porém corrigidas manualmente.	99

Parte I

Introdução

Capítulo 1

Descrição da Tese

1.1 Introdução

Com o crescente volume de informações no mundo digital, há uma necessidade de meios cada vez mais seguros para sua proteção. A identificação e autenticação de usuários através de mecanismos de identificação tradicionais, baseados em usuário e senha, já não satisfazem às demandas exigidas, sendo necessária a utilização de meios mais confiáveis para esse fim. Neste sentido, a implementação de técnicas biométricas propicia maior confiabilidade a sistemas que possuem como premissa a identificação pessoal de indivíduos. O termo biometria pode ser definido como "características físicas únicas e mensuráveis de uma pessoa". Todos os indivíduos possuem algumas características que podem ser identificadas de forma única, como por exemplo a impressão digital, a geometria da mão, a formação da face, o DNA e outras.

Existem duas etapas a serem cumpridas para a realização da identificação biométrica. A primeira é o registro do usuário no sistema, onde a característica em questão é capturada, convertida em um modelo matemático, e armazenada em um banco de dados. A segunda etapa é o reconhecimento, onde o sistema compara a característica apresentada pelo usuário com o modelo previamente armazenado.

Os sistemas biométricos baseados em impressões digitais são sistemas que fundamentam suas decisões de reconhecimento tomando como característica pessoal a impressão digital. Em especial, os AFIS (*Automatic Fingerprint Identification System*) são os sistemas de identificação com maior aplicação na atualidade, tendo vários módulos constitutivos, sendo dois dos principais:

Coincidência é a comparação entre duas impressões digitais estabelecendo-se seu grau de similaridade. É uma tarefa extremamente difícil, devido

principalmente à grande variabilidade entre diferentes instâncias da mesma impressão.

Classificação a grande quantidade de impressões digitais capturadas e armazenadas num amplo intervalo de aplicações torna pouco prática a comparação exaustiva de uma impressão desconhecida com todas as registradas no banco de dados. Assim, para reduzir o tempo de busca e a complexidade computacional, é desejável classificar as impressões capturadas de forma consistente e precisa, de tal maneira que a comparação seja feita só com um subconjunto das impressões no banco de dados. A classificação é uma técnica utilizada para catalogar uma impressão digital dentro de uma das classes estabelecidas na literatura.

O desenvolvimento de um sistema biométrico baseado em impressões digitais está intimamente relacionado com o processamento digital de imagens e o reconhecimento de padrões. Por processamento digital de imagens entende-se a manipulação de uma imagem de entrada, de modo que a saída do processo seja uma nova imagem. Analogamente, para o reconhecimento de padrões, apresenta-se à entrada do processo um padrão (a imagem ou uma característica derivada dela) obtendo-se como saída uma categoria ou classe. No caso de impressões digitais o objetivo final é a comparação da impressão digital a ser reconhecida com a armazenada no banco de dados.

1.2 Caracterização do Problema

A problemática que afronta a classificação de impressões digitais é: como realizar a tarefa de coincidência num tempo admissível para um banco de dados de milhões de impressões? por exemplo, supondo um banco de dados de 1 milhão de impressões e considerando que o tempo de comparação seja 0.1s, realizar a busca em todo o banco de dados levará 1 dia, se o banco de dados é particionado em 5 classes, tal tarefa levará 5 horas, sendo a redução de tempo drástica. Neste ponto, surgem outros problemas como:

- que critérios tem-se para realizar o particionamento do banco de dados,
- tais critérios permitem estabelecer processos automatizados que tenha um máximo de rendimento ao momento de efetuar a classificação de uma impressão,
- as características a serem utilizadas para a classificação são robustas à qualidade da impressão, erros de aquisição?

Por estes motivos e por ser um problema com uma aplicação real no mundo atual, este problema despertou o interesse de vários pesquisadores durante os últimos 30 anos.

1.3 Objetivos

O objetivo geral deste trabalho é estudar e implementar um sistema de classificação de impressões digitais tomando como ponto de partida a arquitetura proposta no PCASYS. Em tal sentido, o sistema implementado divide-se em duas partes principais, uma dedicada a processar a impressão digital e extrair suas características representativas e outra orientada a realizar a tarefa de classificação.

O objetivo específico é elaborar técnicas robustas na etapa de processamento da impressão digital de maneira que o rendimento do classificador seja maximizado. Para realizar tal tarefa são propostas técnicas originais para a maioria dos componentes constitutivos da etapa de processamento de maneira que o efeito das diferentes fontes que geram uma má classificação sejam atenuados.

1.4 Organização da dissertação

Este trabalho está organizado da seguinte maneira:

- No capítulo II estabelece os conceitos básicos e define um marco referencial da problemática da classificação de impressões digitais, detalham-se as diferentes soluções propostas pela comunidade científica para a solução deste problema, começando por uma resenha histórica e logo uma exposição do estado do arte das técnicas de classificação propostas até a atualidade.
- No capítulo III, detalham-se os passos prévios para etapa de classificação de impressões digitais, definida como a etapa de processamento da impressão digital. Aqui é explicado o cálculo da imagem de regiões, como é melhorada a imagem de entrada, como é realizada a determinação da imagem de orientação e, por último, detalham-se as técnicas que permitem registrar a imagem de orientação para eliminar as variações devidas a uma possível translação da imagem de entrada.

- No capítulo IV, são apresentados os componentes do classificador implementado, explicando-se cada uma de suas partes constitutivas: Geração de características, redução dimensional e os classificadores usados.
- No capítulo V, são apresentados os resultados experimentais feitos para a avaliação do sistema de classificação implementado, usando um banco de dados especializado para testar sistemas de classificação de impressões digitais, de modo que os resultados obtidos sejam comparáveis àqueles publicados por outros pesquisadores.
- No capítulo VI, se apresentam as conclusões deste trabalho junto com as propostas de trabalhos futuros.

Parte II

Fundamentação Teórica

Capítulo 2

Definições e Estado do Arte

2.1 Introdução

Este capítulo tem por finalidade estabelecer os conceitos básicos e definir um marco referencial da problemática da classificação de impressões digitais. Portanto, o capítulo inicia-se com a definição de uma impressão digital e como ela é caracterizada, em continuação detalham-se as diferentes soluções propostas pela comunidade científica para a solução deste problema, começando por uma resenha histórica e logo uma exposição do estado da arte das técnicas de classificação propostas até a atualidade.

2.2 Impressões Digitais

Uma impressão digital é a reprodução sobre uma superfície plana das papilas (elevações da pele) presentes nas polpas dos dedos das mãos, armazenada em formato analógico (papel) ou digital (arquivo). Nestas, as cristas papilares apresentam-se como as linhas mais escuras e os vales interpapilares como as linhas mais claras (ver Figura 2.1).

2.2.1 Características

Os especialistas definem dois tipos de características presentes numa impressão digital, comumente denominadas globais e locais, elas são descritas a seguir.

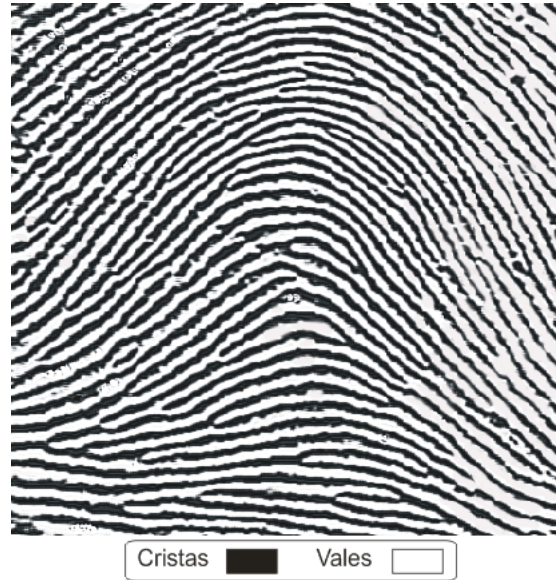


Figura 2.1: Impressão Digital.

Características globais

São os tipos de padrões geométricos das cristas que são reconhecíveis por simples observação, baseados nestes padrões as impressões digitais dividem-se em três grupos principais: *Loop*, *Arch* e *Whorl*¹. Estes grupos também se subdividem em grupos menores, como é o caso do *Loop* que se subdivide em *Left loop* e *Right loop*². Geralmente a determinação do padrão a que pertence a impressão digital obtêm-se mediante o conhecimento de seus pontos singulares. É necessário, portanto, explicar os conceitos relacionados.

Área Padrão É a parte principal da impressão digital e está constituída pelas cristas e todas as suas características.

Linhas Tipo São definidas como duas cristas que iniciam-se paralelamente e divergem sobre o área padrão. Estas cristas podem ser contínuas ou não, em caso de que ocorra alguma ruptura (ver Figura 2.2).

¹A tradução destes termos ao português são presilha, arco e verticilo respectivamente. Posto que na literatura especializada usam-se os términos em inglês estes serão mantidos ao longo do texto

²presilha esquerda ou interna e presilha direita ou externa, respectivamente

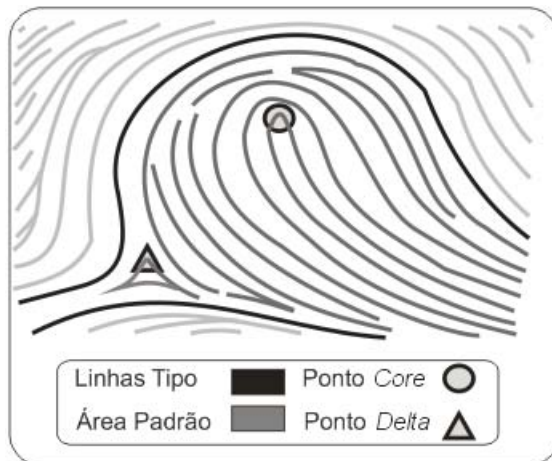


Figura 2.2: Área Padrão e Linhas Tipo.

Ponto Core Está localizado dentro da área padrão em que as cristas apresentam uma maior curvatura. Ele é definida como o ponto superior da linha da crista mais interna. Devido à grande variação nas configurações das cristas (ver Figura 2.3.a), as técnicas para sua determinação automática são complexas.

Ponto Delta É o ponto de divergência das linhas tipo mais internas que tendem a envolver a área padrão. Um Delta é um triângulo constituído pelas cristas papilares que podem formar-se de duas maneiras: pela bifurcação de uma linha simples ou pela brusca divergência de duas linhas paralelas (ver Figura 2.3.b).

Características locais

As cristas em uma impressão digital não são contínuas nem retas, elas trocam de direção, cortando-se e bifurcando-se. Os pontos em que estas trocas ocorrem são denominados minúcias, que provêem a informação suficiente para determinar a individualidade de uma impressão digital. Estes pontos permitem a um perito (papiloscopista) identificar uma pessoa de forma bastante confiável. As minúcias de finalização e bifurcação são consideradas como os tipos mais importantes (ver Figura 2.4), posto que elas geram todos os outros tipos de minúcias, além de serem as mais adequadas para um processo de detecção automática. Os principais atributos de uma minúcia são: tipo, posição e orientação, sendo estes os dados armazenados para a tarefa de reconhecimento.

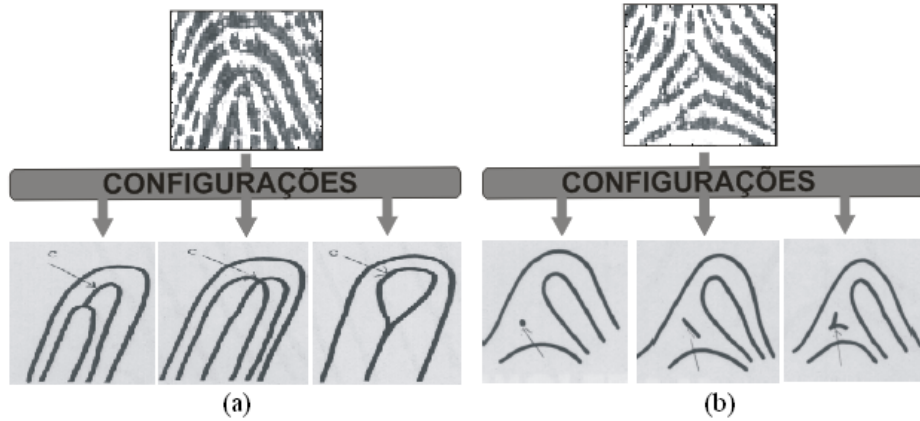


Figura 2.3: Pontos singulares (a) configurações do ponto *Core*; (b) configurações do ponto *Delta*.

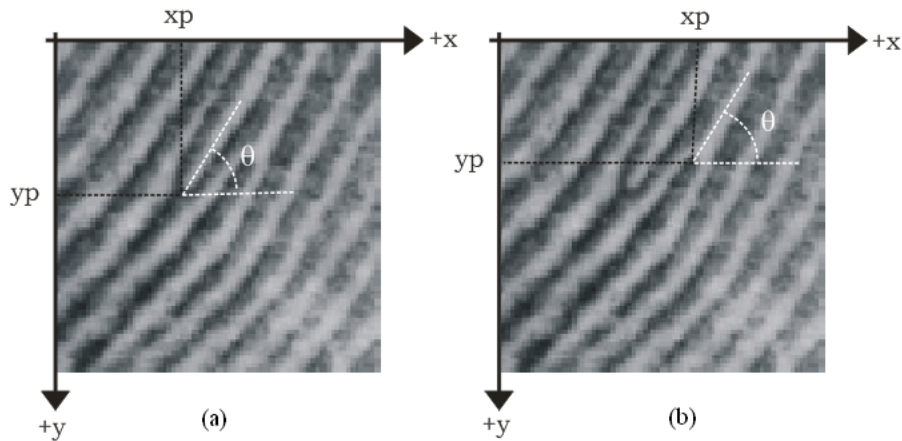


Figura 2.4: Tipos Minucias (a) bifurcação; (b) finalização.

2.3 Classificação de impressões digitais

A identificação de uma pessoa requer uma comparação de sua impressão digital versus todas as impressões do banco de dados. Em aplicações forenses e cíveis, este banco de dados pode ser de milhões de indivíduos (por exemplo o banco de dados do sistema de reconhecimento do FBI (*Federal Bureau of Investigation*) consta de 46 milhões de impressões digitais [59]). Em tal situação, se é comparada uma impressão contra todo o banco de dados, o tempo de resposta é inaceitável. Por outro lado, o processo de identificação pode ser acelerado pela redução do número de comparações. Algumas vezes, a informação do sexo, raça, idade e outros dados relacionados ao indivíduo

estão disponíveis e a porção do banco de dados onde a procura será efetuada pode ser reduzida; porém, esta informação não é sempre acessível (por exemplo, a identificação criminal baseada em impressões digitais latentes). Uma estratégia comum para ter um tempo de comparação aceitável é dividir o banco de dados em diferentes classes, assim, uma impressão digital que requer identificação será comparada unicamente com as impressões digitais pertencentes àquela classe em questão.

As regras para a classificação de impressões digitais foram propostas por Purkinje em 1823, que classificava as impressões digitais em nove categorias (*transverse curve*, *central longitudinal stria*, *oblique stripe*, *oblique loop*, *almond whorl*, *spiral whorl*, *ellipse*, *circle* e *double whorl*) de acordo com as configurações globais das cristas. O primeiro estudo científico em profundidade sobre classificação de impressões digitais foi realizado por Francis Galton, que dividiu as impressões digitais dentro de três classes (*arch*, *loop* e *whorl*) e dividiu cada classe em sub-classes, paralelamente, Juan Vucetich, da polícia federal da Argentina, desenvolveu um sistema diferente de classificação; o sistema de classificação de Vucetich. Vucetich também foi o primeiro a fazer uma identificação de impressões digitais de um suspeito, em 1892. Dez anos depois, Edward Henry redefiniu a classificação de Galton ao incrementar o número de classes. Todos os esquemas atualmente usados pelas agências policiais são variações do chamado esquema de classificação de Galton-Henry. O esquema de classificação de Galton-Henry classifica as impressões digitais em cinco classes:

- Uma impressão digital *Arch* possui cristas que entram por um lado, sobem uma pequena protuberância e saem pelo lado oposto. Estas impressões não tem pontos singulares notórios *core* nem *delta* (ver Figura 2.5.a).
- Uma impressão digital *tented arch*³ é similar a uma *arch*, salvo que nesta as cristas mais profundas exibem uma alta curvatura e apresentam um ponto singular *core* e um ponto singular *delta* (ver Figura 2.5.b).
- Uma impressão digital tipo *loop* tem uma ou mais cristas que entram de um lado, curvam-se e retrocedem, para logo sair pelo mesmo lugar por onde entraram, apresentam um ponto singular *core* e um ponto singular *delta*, assume-se que o *delta* está ao sul do *core*. O tipo *Loop* pode ser subdividido em *Loops* que têm cristas que entram e saem do lado esquerdo, denominadas *left loops* (ver Figura 2.5.d), e *loops* que

³Sua respectiva tradução é arco angular.

têm cristas que entram e saem do lado direito, denominadas *right loops* (ver Figura 2.5.c).

- Em uma impressão digital tipo *whorl*, as cristas mais profundas realizam um giro de 360° ao redor do centro da impressão digital. Dois pontos singulares *core* e dois pontos singulares *delta* podem encontrar-se neste tipo de impressões. A classe *whorl* é bastante complexa e alguns esquemas de classificação realizam uma divisão em duas sub-classes: *plain whorl* e *twin loop* (ver figuras 2.5.e e 2.5.f).

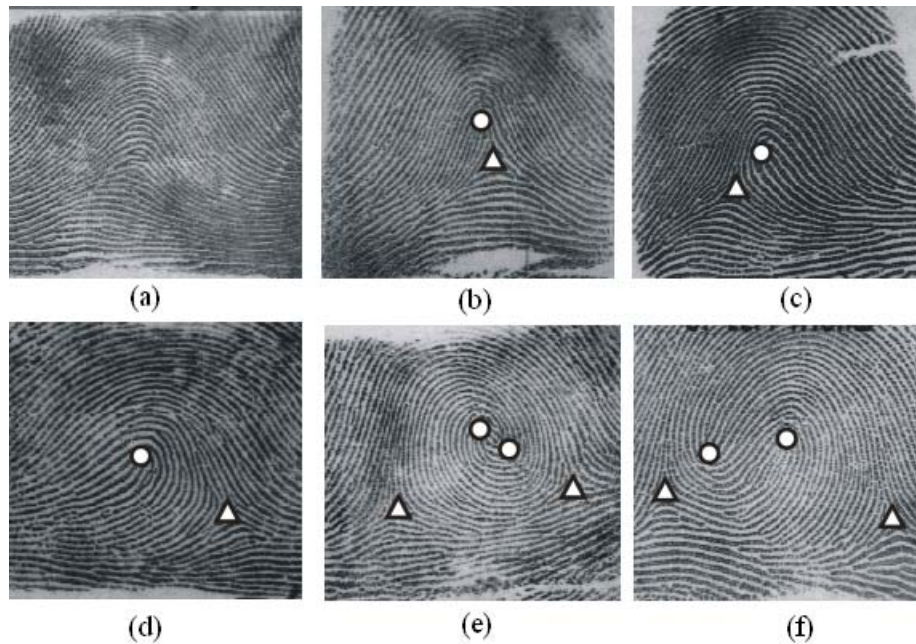


Figura 2.5: Seis classes das impressões digitais (a) *Arch*; (b) *tented arch*; (c) *right loops*; (d) *left loops*; (e) *plain whorl*; (f) *twin loop*. Onde o ponto *core* é indicado pelo círculo e o ponto *delta* é indicado pelo triângulo.

O esquema de classificação adotado pelo FBI define três classes, cada uma das quais pode ser dividida dentro de duas ou mais sub-classes, fazendo um número total de oito: *Plain Arch*, *Tented Arch*, *Radial Loop*, *Ulnar Loop*, *Plain Whorl*, *Double Loop*, *Central Pocket* e *Accidental Whorl*⁴ (ver Figura 2.6).

A classificação de impressões digitais é um problema difícil de reconhecimento devido a:

⁴as respectivas traduções são: arco plano, arco angular, presilha radial, presilha cubital, verticilo plano, presilha dupla, bolso central, verticilo accidental.

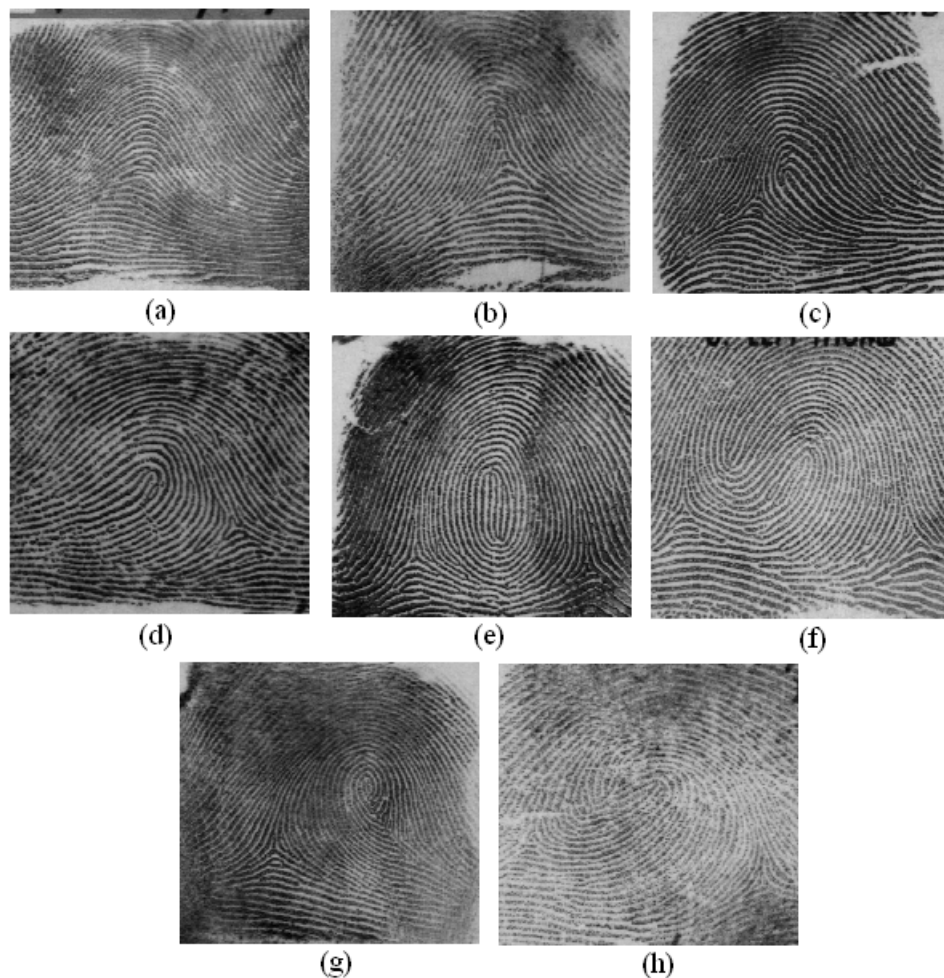


Figura 2.6: O esquema de classificação adotado pelo FBI (a) *Plain Arch*; (b) *Tented Arch*; (c) *Radial Loop*; (d) *Ulnar Loop*; (e) *Plain Whorl*; (f) *Double Loop*; (g) *Central Pocket*; (h) *Accidental Whorl*.

- pequena variabilidade⁵ entre impressões de diferentes classes, conhecido como variação inter-classes (ver Figura 2.7);
- grande variabilidade entre impressões da mesma classe, conhecido como variação intra-classes (ver Figura 2.8);

⁵A variabilidade é considerada respeito ao padrão de linhas de cristas presente na impressão digital, ou seja, as impressões que pertencem a uma mesma classe apresentaram um padrão de linhas similar (pequena variabilidade entre impressões da mesma classe), as impressões que pertencem a diferentes classes apresentaram diferentes padrões de linhas (grande variabilidade entre impressões de diferentes classes).

- ruído presente nas impressões digitais⁶, o qual faz com que a tarefa de classificação seja mais complexa (ver Figura 2.9).

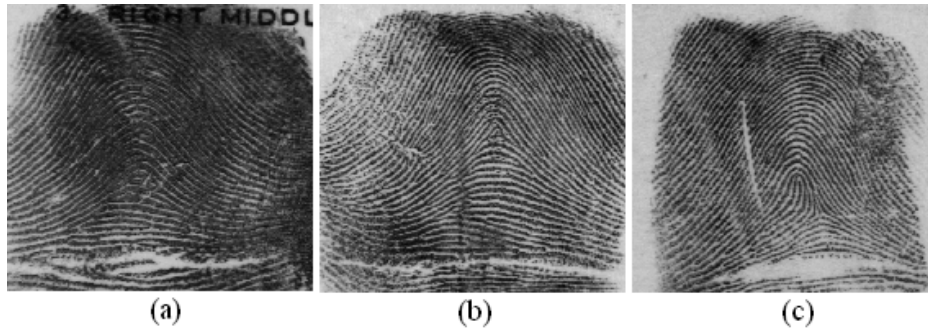


Figura 2.7: Três impressões digitais que pertencem a diferentes classes e tem a mesma aparência (pequena variabilidade inter-classes). (a) *arch*; (b) *Tented Arch*; (c) *left loop*.

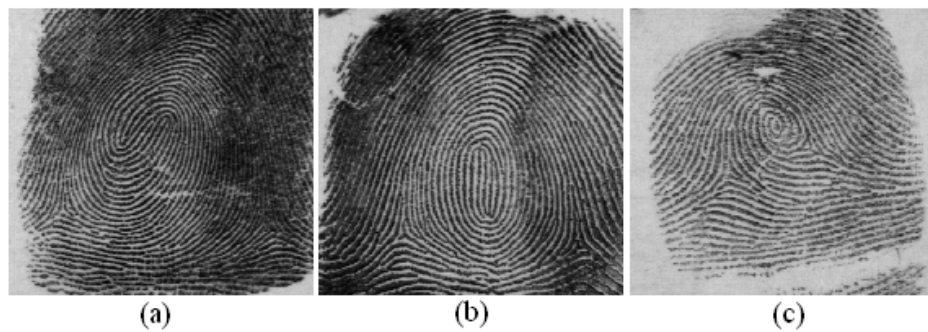


Figura 2.8: Três impressões digitais pertencentes à mesma classe que apresentam características muito diferentes (grande variabilidade intra-classes). (a) *twin loop* ;(b) *Plain Whorl*; (c) *Central Pocket*.

A maioria de AFIS usam uma classificação exclusiva. Ou seja, as impressões digitais são particionadas dentro de algumas classes predefinidas em relação a suas macro-estruturas (número de pontos singulares, formas das linhas tipo). Estas técnicas dependem do número de classes e a distribuição natural das impressões digitais nestas classes. Infelizmente, o número de classes usado é comumente pequeno e as impressões digitais não são distribuídas uniformemente nas classes. Por exemplo, a maioria dos sistemas automáticos consideram cinco classes (ou seja, *arch*, *tented arch*, *left loop*, *right loop* e *whorl*) e a proporção natural das impressões digitais nestas classes

⁶Este pode dever-se à técnica de aquisição usada, características do dispositivo de captura, defeitos na imagem e condições da pele.

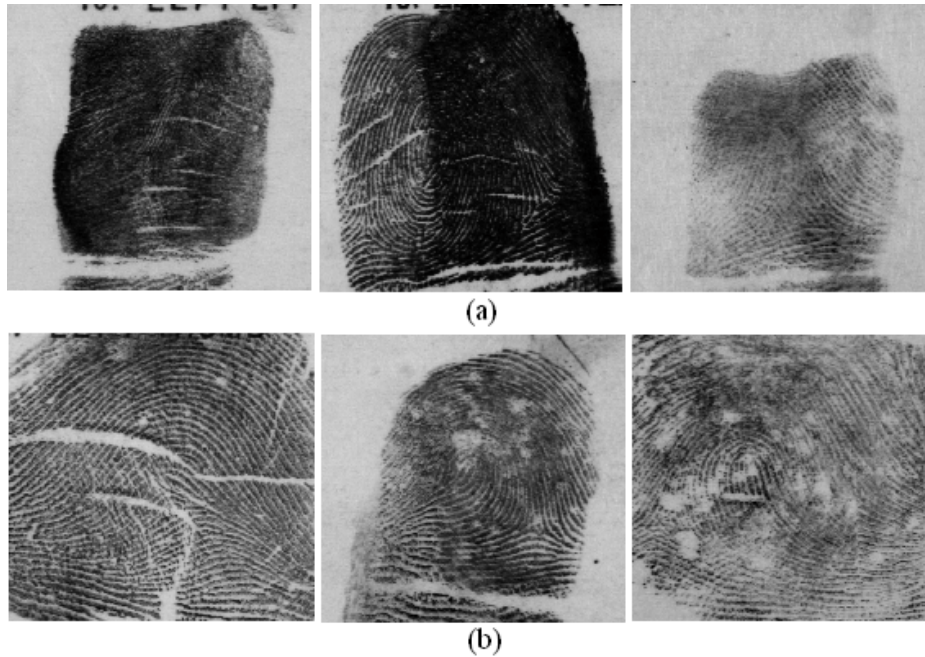


Figura 2.9: Exemplos de imagens de impressões digitais ruidosas (a) devidas às condições da pele; (b) devidas à técnica de aquisição.

são 3,7%, 2,9%, 33,8%, 31,7% e 29,9%, respectivamente [103]. Além disto, têm-se muitas impressões digitais ambíguas, cuja pertinência a uma única classe não pode ser confiavelmente estabelecida por um especialista forense. Não obstante, a classificação exclusiva permite melhorar a eficiência na identificação sobre dez impressões, posto que o conhecimento das classes das dez impressões digitais pode ser usado como um código para a redução do número de comparações no momento de realizar o processo de coincidência de minúcias. Por outro lado, as técnicas de classificação de impressões digitais não oferecem suficiente seletividade na procura de impressões digitais latentes. Para aplicações em que não é necessário complicar-se com o esquema de classificação exclusiva, alguns autores propõem técnicas baseadas na classificação contínua. Uma técnica de classificação contínua caracteriza cada impressão digital como um vetor numérico, cujos componentes denotam o grau de similaridade com respeito a um conjunto pre definido de classes protótipo. A vantagem da classificação contínua versus a classificação exclusiva é discutida em [63].

2.4 Extração de características

O problema da classificação de impressões digitais tem recebido a atenção da comunidade científica devido à sua importância e à sua dificuldade intrínseca. Um grande número de trabalhos foram publicados sobre este tema durante os últimos 30 anos. Apesar de que uma grande variedade de técnicas terem sido propostas para este problema, relativamente um pequeno número de características extraídas a partir das impressões digitais são usadas pela maioria de autores. Em particular quase todas as técnicas são baseadas numa das seguintes características: imagem de orientação, linhas de fluxo de crista, pontos singulares e análise dos filtros de Gabor, como é observado na Figura 2.10.

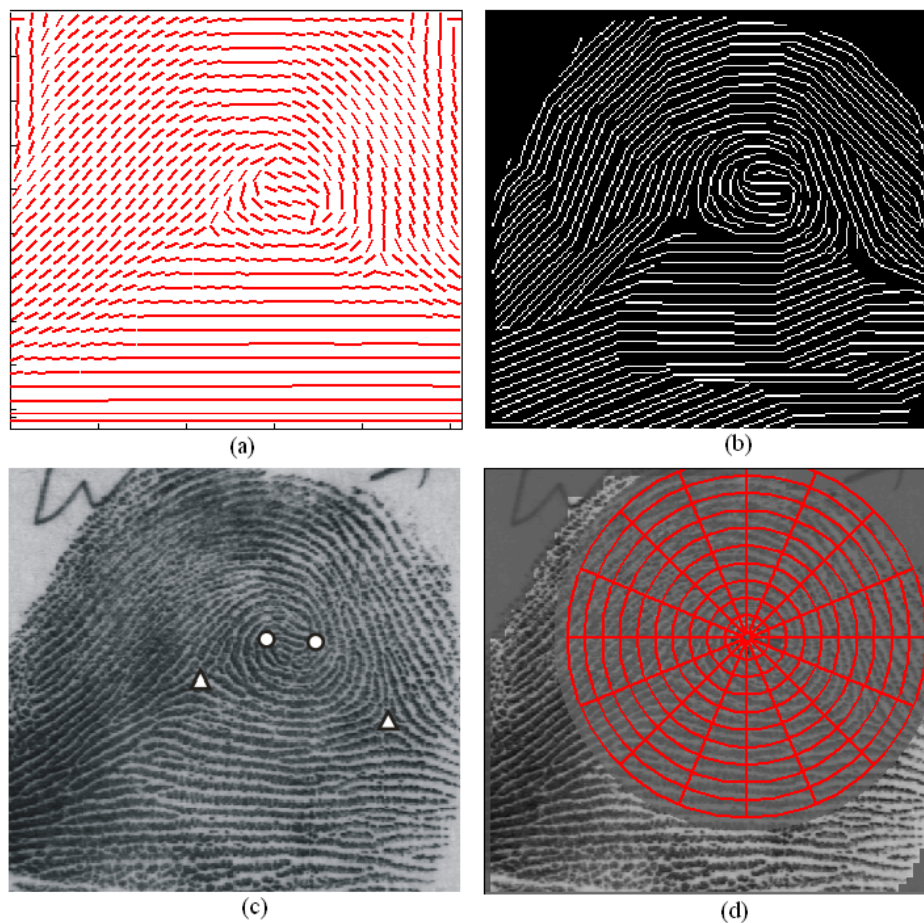


Figura 2.10: Características usadas com maior frequência para a classificação de impressões digitais: (a) imagem de orientação; (b) linhas de fluxo de crista; (c) pontos singulares; (d) resposta dos filtros de Gabor.

imagem de orientação É uma matriz discreta cujos elementos representam a orientação média local das cristas da impressão digital (ver Figura 2.10.a). A imagem de orientação resume com eficácia a informação contida no padrão da impressão digital e pode ser calculada com facilidade a partir de imagens ruidosas. Além disso, as orientações locais de zonas ruidosas podem ser restauradas por processos de regularização. É por estes motivos que a maioria das técnicas de classificação existentes usam a imagem de orientação como característica.

linhas de fluxo de crista As linhas de fluxo das cristas são características discriminatórias de importância, apesar de não serem fáceis de extrair a partir de imagens ruidosas, mas ainda, são mais robustas que os pontos singulares. Elas são geralmente representadas como um conjunto de curvas correndo paralelamente sobre as linhas de crista; estas curvas não necessariamente coincidem com as cristas e vales das impressões digitais, mas elas exibem a mesma orientação local (ver Figura 2.10.b). As linhas de fluxo das cristas são tipicamente extraídas da imagem de orientação ou pela binarização da imagem e a realização da operação de afinamento das cristas binarizadas. Frequentemente, antes da extração das linhas de fluxo de cristas, a imagem é melhorada pela aplicação de técnicas de filtragem. Algumas técnicas [15][24] depois de detectar as linhas de crista, analisam suas formas e extraem uma ou mais curvas representativas (não necessariamente correspondentes a uma linha de crista), as quais são chamadas pseudo cristas ou traços de fluxo de linhas.

pontos singulares Achar os pontos singulares *core* e *delta* em uma impressão digital não é uma tarefa fácil, especialmente em imagens de pobre qualidade (ver Figura 2.10.c). A maioria de técnicas adotadas para a detecção das singularidades está baseada no índice de Poincaré [58], o qual é calculado sobre uma pequena curva fechada ao redor de um ponto particular; este é frequentemente realizado utilizando a imagem de orientação. Por outro lado, as posições das singularidades podem ser muito úteis para o alinhamento das impressões digitais com respeito a um ponto fixo (o ponto *core*).

resposta dos filtros de Gabor Os filtros de Gabor foram usados em alguns trabalhos para o melhoramento das imagens de impressões digitais [44], devido a suas propriedades seletivas de frequência e orientação. Mas em [50] apresentou-se uma representação das impressões digitais baseada na resposta dos filtros de Gabor chamada *Finger Code* (ver Figura 2.10.d), por

sua analogia com o conhecido *Iris Code* introduzido por Daugman [28]. O *Finger Code* pode ser usado tanto para a tarefa de classificação como de coincidência de impressões digitais.

A maioria das técnicas de classificação existentes fazem uso da imagem de orientação (ver Tabela 2.1). Isto não é surpreendente considerando que aquela característica, se é calculada com suficientes exatidão e detalhe, incluirá toda a informação requerida para a classificação. Geralmente, a imagem de orientação é registrada com respeito ao ponto *core* antes de ser processada. Mas, alguns autores [20], [15] propuseram técnicas específicas para o melhoramento da imagem de orientação, as quais permitem aumentar a exatidão da classificação.

2.5 Técnicas de classificação

A maioria das técnicas de classificação de impressões digitais existentes podem ser atribuídos a uma destas categorias:

- técnicas baseadas em regras,
- técnicas sintáticas,
- técnicas estruturais,
- técnicas estatísticas,
- técnicas basadas em redes neurais,
- técnicas de múltiplos classificadores.

A Tabela 2.1 ressalta as características usadas nas estratégias de classificação adotadas pelas técnicas de classificação de impressões digitais publicados nos últimos 30 anos. A classificação de impressões digitais é um dos mais representativos problemas de reconhecimento de padrões e pela observação da Tabela 2.1, pode-se avaliar a evolução nas últimas três décadas da investigação no campo do reconhecimento de padrões: o interesse nas técnicas sintáticas nos anos 1970/1980, e os êxitos das redes neurais nos anos 1990, e os sistemas de múltiplos classificadores nos anos atuais.

2.5.1 Técnicas baseadas em regras

Baseadas em técnicas usadas pelos especialistas forenses, utiliza a posição e o número de singularidades para efetuar a tarefa de classificação. Estas técnicas são atrativas por sua simplicidade, porém, alguns apresentam dificuldades

quando tratam-se impressões digitais parciais ou ruidosas, onde a detecção das singularidades pode ser extremadamente difícil. Outro problema comum que ocorre com a técnica baseada em singularidades é que, apesar de trabalhar razoavelmente bem sobre impressões digitais roladas ou escaneadas a partir de fichas papiloscópicas, elas não são adequadas quando trabalham com impressões digitais capturadas ao vivo (*live-scan*), pois os pontos *delta* são comumente perdidos neste tipo de imagens.

- Em [58] determina-se a posição e o tipo das singularidades através do cálculo do índice de Poincaré. Levando-se em conta essa informação, uma primeira classificação é realizada, depois uma classificação fina é obtida pelo seguimento das linhas de fluxo de cristas.
- Em [56] uma regularização iterativa (pela suavização da imagem de orientação com filtros 3x3) é realizada até que um número válido de pontos singulares são detetados; isto permite reduzir o ruído presente na imagem de orientação e conseqüentemente melhorar a exatidão na classificação.
- Em [43] uma técnica mais robusta é proposta; os autores introduzem um algoritmo de classificação baseado em regras que usa o número de singularidades junto com o número de cristas recorrentes achadas na imagem; a combinação destas duas características conduz a um melhor rendimento que o obtido em [56].
- Em [23] propõe-se uma técnica que usa unicamente os pontos *core* e classifica as impressões digitais de acordo com a curvatura e orientação da impressão digital nas áreas próximas ao *core*.
- Em [106] é desenvolvida uma nova técnica para o cálculo da imagem de orientação, e a classificação é realizada através de regras heurísticas vinculadas ao número de pontos singulares e ao número de giros das linhas de fluxo de cristas. Além disso, é implementada uma técnica de reforço para a melhora na classificação das impressões *Whorl*.

Em [24], [21], [49] e [27] foram propostas técnicas baseadas em regras que não realizam a classificação baseando-se nas singularidades: a classificação é baseada nas formas geométricas das linhas de crista. Vejamos:

- Em [24], curvas B-spline são usadas para modelar as linhas de crista das impressões digitais, curvas adjacentes são combinadas para limitar os efeitos do ruído, e a classificação é realizada pela detecção dos giros das curvas resultantes.

- Em [21] A classificação é realizada dependendo da distribuição global das cristas, considerando que tem-se dez padrões de linhas cristas básicos e pela análise da forma e seqüência destes padrões realiza-se a classificação.
- Em [49], para cada classe, um *kernel* (o qual modela a forma das impressões digitais nas classes) é definido; a classificação é então realizada pela determinação do *kernel* que melhor coincida com a imagem de orientação da impressões digitais analisadas.
- Em [27] é usada como característica de entrada as linhas de fluxo de cristas. As características geométricas das linhas são analisadas estudando os mapas isométricos de planos tangentes a um ponto transversal ao longo da curva desde um início à outro. A rota traçada pelo mapa isométrico contém várias e importantes características tais como pontos de troca de sinal e valores de extremos locais, que identificam a unicidade das características geométricas inerentes de cada linha de fluxo.

2.5.2 Técnicas sintáticas

Uma técnica sintática descreve os padrões através de símbolos terminais e regras de produção; uma gramática é definida para cada classe e um processo de análises (gramatical) é responsável pela classificação de cada novo padrão.

- Em [76] desenvolveu-se uma técnica sintática onde os símbolos terminais são associados com pequenos grupos de elementos da imagem de orientação; uma classe de gramática livre de contexto é usada para descrever os padrões de impressões digitais, os quais são divididos dentro de sete classes. Os mesmos autores também experimentaram com outros tipos de gramáticas: gramáticas estocásticas [74] e gramáticas de árvore [75].
- Em [85] a técnica introduzida está baseada na análise do fluxo de linhas de crista, a qual é representada por um conjunto de linhas conectadas. Estas linhas são etiquetadas tomando em conta as trocas de direção, assim obtêm-se um conjunto de cadeias que são processadas através de uma gramática *ad-hoc* ou técnicas de coincidência de cadeias para derivar a classificação final.

Em geral, devido à grande diversidade de padrões de impressões digitais, são necessárias gramáticas complexas cuja inferência requer complicadas e

instáveis técnicas; por esta razão, o uso de técnicas sintáticas para a classificação de impressões digitais foi quase abandonado, com pequenas exceções [21].

2.5.3 Técnicas Estruturais

As técnicas estruturais são baseadas na organização relacional das características de nível inferior dentro de estruturas de nível superior. Esta organização relacional é representada através de estruturas de dados simbólicas, tal como árvores e grafos, os quais permitem uma organização hierárquica da informação [67].

A imagem de orientação é adequada para a representação estrutural: de fato, ela pode ser particionada dentro de regiões conectadas que são caracterizadas por orientações homogêneas; estas regiões e as relações entre elas contêm informação útil para a classificação.

Esta é a ideia básica do técnica proposta por [65][64]: a imagem de orientação é particionada dentro de regiões pela minimização de uma função de custo que leva em conta a variação das orientações dentro de cada região. Uma técnica de coincidência de grafos inexatos é então usada para comparar os grafos relacionais com os grafos dos protótipos de classe (ver Figura 2.11).

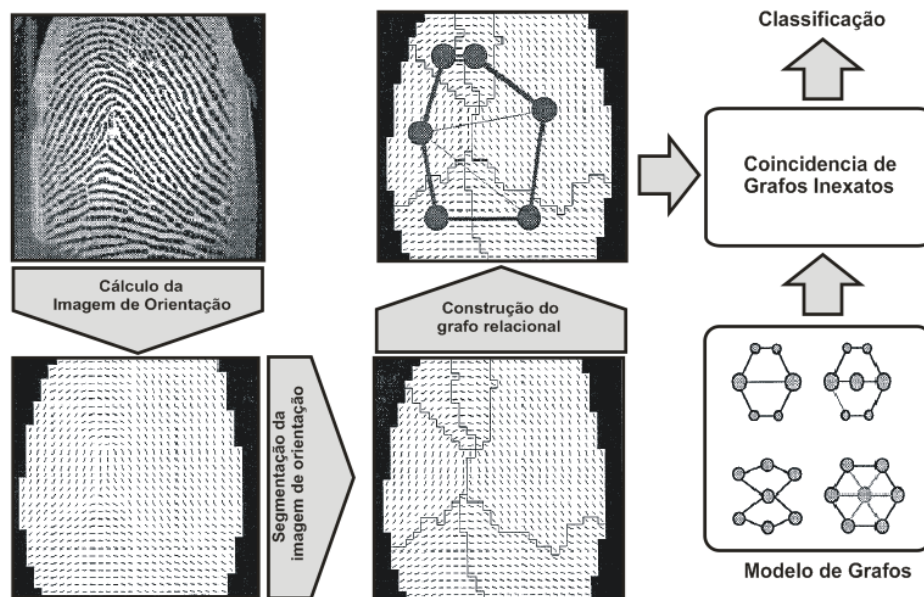


Figura 2.11: Técnica de classificação baseada em grafos inexatos [65].

Ainda que as técnicas de grafos relacionais tenham propriedades interessantes (tal como invariância à rotação e translação e a possibilidade de

manipular impressões digitais parciais), não é fácil particionar robustamente a imagem de orientação em regiões homogêneas, especialmente em impressões digitais de pobre qualidade.

- Em [20], uma técnica de coincidência baseada em máscaras é realizada para guiar o particionamento da imagem de orientação; a principal vantagem da técnica é que, por só considerar a informação estrutural global, ela é capaz de tratar com impressões digitais parciais, quando alguns pontos singulares não são disponíveis e pode também trabalhar com imagens ruidosas.
- Em [89] adota-se um classificador por cadeias ocultas de Markov (*Hidden Markov Models* - HMM), para a tarefa de classificação de impressões digitais; as características de entrada são as medidas tomadas nos pontos de intersecção entre algumas linhas fiduciárias horizontais e verticais e as linhas de crista. Em cada intersecção as características extraídas são: a orientação local das linhas de crista, a distância e a variação na orientação desde a última intersecção e a curvatura da linha de crista na intersecção. Posto que os modelos HMM utilizados neste trabalho são unidimensionais, os padrões bidimensionais de impressões digitais são linearizados pelo aninhamento das filas, obtendo-se assim uma seqüência única de observações.

2.5.4 Técnicas Estatísticas

Nas técnicas estatísticas, um vetor de características numérico de tamanho fixo é derivado de cada impressão digital e um classificador estatístico de propósito geral é usado para a classificação.

Um dos classificadores estatísticos mais comumente adotados é o k -NN ([46]), exemplos da aplicação do k -NN na classificação de impressões digitais pode ser achado nos seguintes trabalhos:

- Em [33], são usadas como características de entrada para o k -NN a representação do anel conificado (*wedge-ring*) obtidas a partir da transformada de Fourier Hexagonal.
- Em [50] o k -NN é o primeiro passo da técnica de classificação em duas fases (ver a seção técnicas baseadas em múltiplos classificadores).

Muitas técnicas usam a imagem de orientação como vetor de características, ao representar cada elemento da imagem de orientação através de seus

componentes vetoriais, $\beta [\cos(\theta), \sin(\theta)]$ e aninhar as filas das matrizes conformantes de ditos componentes (ver por exemplo, [16] e [15]). Uma típica imagem de orientação de 30x30 elementos resulta num vetor de 1800 (=30x30x2) elementos. O treinamento dos classificadores com vetores de tão alta dimensão requer grande quantidade de dados de treinamento, memória e tempo de cálculo. Por esta razão, técnicas estatísticas de redução dimensional são comumente usadas para reduzir a dimensionalidade do vetor de características. A técnica de análise da componente principal (*Principal Component Analysis*-PCA) ou transformada *Karhunen - Loève* (KL) [53] é geralmente adotada para esta tarefa, posto que ela garante uma boa preservação da distância Euclidiana entre vetores (ver por exemplo, [103] e [41]). O PCA além de ser utilizada para a redução dimensional, também pode ser adotada para a classificação em se.

- Em [18] foi usada para a representação dos vetores de características derivados a partir da imagem de orientação. A idéia desta técnica foi achar para cada classe um ou mais sub-espacos KL que estejam bem adequados para representar as impressões digitais pertencentes a cada classe. Estes sub-espacos foram criados de acordo com um critério de otimização que procura minimizar o erro de reconstrução quadrado médio sobre um conjunto de treinamento representativo. O número de sub-espacos para cada classe foi fixado *a priori* tomando-se em conta a complexidade da classe, em particular, mais sub-espacos foram criados para classes complexas (por exemplo, *whorl*).
- Em [16] a classificação era simplesmente realizada com respeito à mínima distância a partir dos sub-espacos KL, enquanto que, em [17] a classificação é realizada utilizando a mínima distância Euclidiana e um classificador *k*-NN.

2.5.5 Técnicas baseadas em redes neurais

Das técnicas propostas baseadas em redes neurais a maioria utiliza a rede perceptron multi-camada e os elementos da imagem de orientação como características de entrada.

- Em [54] apresenta-se uma interessante arquitetura piramidal constituída de vários perceptrons multi-camada, cada um dos quais é treinado para reconhecer impressões digitais pertencentes a uma determinada classe.

- Em [12] a posição das singularidades é usada junto com a imagem de orientação (20x20) elementos para o treinamento de duas redes neurais disjuntas, cujas saídas são passadas a uma terceira rede, a qual produz o resultado final da classificação.
- Em [50] treina-se dez redes neurais perceptron multi-camada para distinguir entre cada possível par de classes.
- Em [36] são usadas como características de entrada uma representação modificada do anel conificado e utiliza como técnica de classificação uma rede neural perceptron multi-camada. Os resultados não são conclusivos para a etapa de classificação, mas valida a representação usada para a classificação.
- Em [83] são usadas como características de entrada medidas de energia provenientes de uma setorização do domínio espectral filtrado com um filtro passa-banda para a atenuação de ruído e utiliza como técnica de classificação uma rede neural probabilística de 12, 16 e 20 neurônios de entrada e 7 neurônios de saída, correspondentes às sete classes. Aqui é feita a suposição que um aumento no número de classes implica num aumento na taxa de acerto do classificador.

Uma das técnicas baseadas em redes neurais mais conhecida foi proposta pelos pesquisadores do NIST [103], e foi denominada a técnica PCASYS (*Pattern-level Classification Automation SYStem*). Vejamos a continuação em detalhe o desenvolvimento desta técnica.

A técnica PCASYS

A técnica PCASYS é uma das técnicas mais referenciadas em classificação de impressões digitais, uma vez que foi um dos primeiros estudos que reporta resultados reproduzíveis e precisos sobre um banco de dados publicamente disponível. Esta técnica foi o resultado de estudos anteriores orientados a:

1. comparar o desempenho de diferentes tipos de classificadores, [14], [10].
2. avaliação de diferentes técnicas de melhoramento de impressões digitais, [98].
3. melhoramento na exatidão da classificação [102].

A técnica PCASYS atribui às impressões digitais cinco classes e é, em nossa opinião, a mais promissora técnica para a classificação exclusiva. Esta técnica está composta das seguintes partes:

Cálculo da Imagem de Regiões Detecta-se a área que tem presença de cristas.

Melhora da imagem Aplica-se uma melhora na imagem no domínio da frequência.

Cálculo da imagem de orientação A imagem de orientação é determinada gerando-se uma matriz de 28×30 elementos.

Registro da imagem de orientação A imagem de orientação é registrada com respeito à posição do ponto *core*, o qual corresponde ao centro da imagem.

Redução dimensional da imagem de orientação Cada elemento da imagem de orientação é representada como um vetor de dois componentes, o qual gera um vetor de 1680 elementos ($=28 \times 30 \times 2$), este vetor é reduzido a 64 elementos pelo uso de técnica PCA.

Classificação Uma rede neural probabilística é usada para atribuir cada vetor de 64 elementos a uma das classes do esquema de classificação. Para melhorar a confiabilidade, os autores também implementaram um módulo auxiliar (chamado seguidor de pseudo cristas), o qual determina o fluxo de cristas na parte superior da impressão digital; este módulo está especificamente projetado para detectar impressões digitais *whorl*. A Figura 2.12 mostra o esquema funcional da técnica de classificação do NIST.

algumas melhoras foram propostas sobre o PCASYS:

- Em [100] propuseram uma técnica baseada em algoritmos genéticos para uma seleção automática dos melhores parâmetros de entrada do PCASYS e alcançaram um bom rendimento na classificação.
- Em [82] uma técnica de extração de características foi adotada para melhorar o rendimento do PCASYS.
- Uma variação do PCASYS para classificação contínua é reportado em [63]. Em particular, os vetores obtidos depois do passo de redução dimensional são diretamente usados para a indexação das impressões digitais do banco de dados, através de uma estrutura de dados espacial. Nos experimentos, os autores afirmam que o melhor resultado pode ser obtido usando um espaço 5-dimensional (no lugar de 64) e reescalando linearmente o vetor de componentes no intervalo de $[0,1]$.

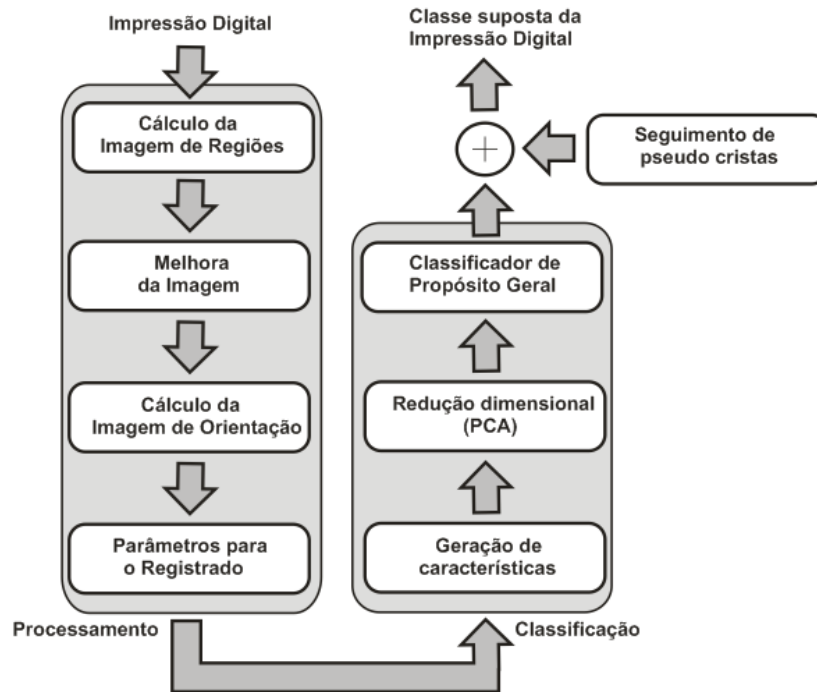


Figura 2.12: Esquema funcional do PCASYS para a classificação exclusiva.

2.5.6 Técnicas Baseadas em vários classificadores

Diferentes classificadores potencialmente oferecem informação complementar a respeito dos padrões a serem classificados, o qual pode ser utilizado para melhorar o rendimento. De fato, em alguns estudos de classificação de padrões, observou-se que diferentes classificadores geralmente classificam mal diferentes padrões. Tal fato motivou o interesse na combinação de diferentes técnicas para a tarefa de classificação de impressões digitais.

Várias opções são possíveis para a seleção dos classificadores (por exemplo, diferentes classificadores treinados sobre os mesmos dados, o mesmo classificador treinado com diferentes dados, diferentes características de entrada) e pela combinação de estratégias (por um simples critério heurístico ou pela regra do voto majoritário ou técnicas mais complexas que utilizam um classificador adicional para a decisão final).

- Em [90] mesclou-se a técnica do PCASYS com outros dois classificadores, um classificador HMM introduzido em [89] e uma técnica baseada nas formas características das cristas classificadas através de uma árvore de decisão.
- Em [50] Adotou-se uma estratégia de classificação em duas fases: um

classificador k -NN é usado para achar as duas classes mais prováveis ao vetor de características de entrada representado pelo *Finger Code*; então uma rede neural treinada para distinguir entre duas classes, é utilizada para obter a decisão final. Um total de 10 redes neurais são treinados para distinguir entre cada possível par de classes.

- Em [17] treinou-se dois diferentes tipos de classificadores sobre três conjuntos de treinamento disjuntos, obtendo-se um total de seis classificadores que os autores combinaram usando a regra de voto majoritário (ver também [19]). Muitos bons resultados foram reportados sobre um banco de dados público.

Outras tentativas de combinação de classificadores para impressões digitais foram realizados por [69], [105].

- Em [69] uma técnica estrutural similar à apresentada por [65] é combinada com um classificador neural que usa como vetor de características o *Finger Code*.
- Em [105] os autores misturaram múltiplos SVM (*Support Vector Machines*) para classificar o vetor de características definido pelo *Finger Code*.

2.6 Resumo

Deste capítulo pode-se concluir que:

- A classificação de impressões digitais foi objeto de vários estudos de reconhecimento de padrões nas três décadas passadas.
- A maioria dos AFIS usam uma classificação exclusiva. Ou seja, para toda impressão de entrada ou pertencente ao banco de dados é atribuída uma das classes pré-definidas (a maioria destes sistemas automáticos consideram cinco classes: *arch*, *tented arch*, *left loop*, *right loop* e *whorl*).
- Diferentes soluções foram propostas para enfrentar o problema da classificação exclusiva, os quais na maioria de casos utiliza a imagem de orientação como característica para realizar a classificação.
- A técnica mais promissora foi a proposta pelos pesquisadores do NIST, denominada PCASYS. A arquitetura desta técnica pode-se dividir em duas partes, aquela encarregada de processar a impressão digital e obter

seus componentes representativos (como são a imagem de orientação) e aquela encarregada de realizar a tarefa de classificação (ver figura 2.12).

O que segue deste trabalho são implementadas as partes constitutivas do sistema de classificação desenvolvido, tal que se tomou como referência a arquitetura do PCASYS. Portanto, os seguintes capítulos estão dedicados especificamente às etapas de processamento e classificação.

Referência	Características				Classificador					
	O	S	R	G	Rb	Sy	Str	Sta	Nn	Mc
Moayer e Fu 1975 [76]	X				X					
Moayer e Fu 1976 [74]	X				X					
Rao e Balk 1980 [85]	X				X					
Kawagoe e Tojo 1984 [58]		X	X		X					
Hughes e Green 1991 [45]	X								X	
Bowen 1992 [12]	X	X							X	
Kamijo, Mieno e Kojima 1992 [55]	X								X	
Moscinska e Tyma 1993 [77]	X				X				X	
Kamijo 1993 [54]	X								X	
Wilson, Candela e Watson 1994 [103]	X								X	
Omidvar, Blue e Wilson 1995 [79]	X								X	
Candela et al. 1995 [15]	X		X		X				X	X
Maio e Maltoni 1996 [65]	X						X			
Halici e Ongun 1996 [41]	X								X	
Karu e Jain 1996 [56]		X			X					
Chong et. al. 1997 [24]			X		X					
Ballan, Sakarya e Evans 1997 [3]		X			X					
Senior 1997 [89]			X				X			
Wei, Yuan e Jie 1998 [100]	X				X				X	X
Capelli et al. 1999 [20]	X					X				
Lumini, Maio e Maltoni 1999 [13]	X					X				
Jain, Prabhakar e Hong 1999 [50]				X				X	X	X
Hong e Jain 1999 [43]		X	X		X					X
Capelli, Maio e Maltoni 1999 [16]	X							X		
Capelli, Maio e Maltoni 2000 [17]	X							X		X
Cho et al. 2000 [23]		X			X					
Bartesaghi, Fernández e Gomes 2001 [4]		X			X					
Bernard et al. 2001 [9]	X								X	
Pattichis et al. 2001 [82]	X				X				X	X
Senior 2001 [90]	X		X		X		X		X	X
Yao, Frasconi e Pontil 2001 [105]				X				X		X
Capelli, Maio e Maltoi 2002 [19]	X							X		X
Jain e Minut 2002 [49]			X		X					

Tabela 2.1: Cronologia das diferentes técnicas de classificação de impressões digitais, cada trabalho é etiquetado de acordo as características usadas (**O** = imagem de orientação, **S** = singularidades, **R** = linhas de fluxo das cristas, **G** = operadores de Gabor) pelas técnicas de classificação propriamente ditas (**Rb** = uso de base de regras, **Sy** = técnicas sintáticas, **Str** = estruturais, **Sta** = estatísticas, **Nn** = redes neurais, **Mc** = múltiplos classificadores) [67].

Parte III

Implementação

Capítulo 3

Processamento da impressão digital

3.1 Introdução

Neste capítulo detalham-se os passos prévios para a etapa de classificação de impressões digitais. O capítulo inicia com o cálculo da imagem de regiões através de uma técnica original, a qual permite separar com êxito a área que tem presença de cristas/vales do fundo da imagem. Logo, explica-se a técnica usada para melhorar a imagem de entrada que consta de duas partes: a primeira, encarregada de normalizar o contraste da imagem, através de uma nova técnica fundamentada na transformada morfológica *Tophat* multiescala, a segunda, encarregada de melhorar a imagem normalizada no domínio da frequência. Em seguida é descrito o cálculo da imagem de orientação, parte chave para todo sistema de classificação. Neste ponto, desenvolvem-se três técnicas propostas na literatura. E, por último, detalha-se as técnicas que permitem registrar a imagem de orientação para eliminar as variações devidas a uma possível translação da imagem de entrada. Neste ponto, uma nova técnica é proposta, fundamentada numa análise realizada em múltiplos níveis de representação da imagem de orientação.

Antes de realizar uma explicação detalhada, é necessário estabelecer certas pautas. Ao longo deste trabalho uma imagem digital será representada matematicamente por uma função bidimensional discreta ou matriz de pontos I de tamanho $N_{Fil}^I \times N_{Col}^I$ cujo valor no ponto (u, v) denominado pixel (*picture element*), representa a iluminação ou intensidade da luz do pixel da u -ésima fila e a v -ésima coluna (ver Figura 3.1). Este é um dos possíveis valores quantizados a oito bits, que na escala de cinza vai a partir de 0 (preto) até 255 (branco).

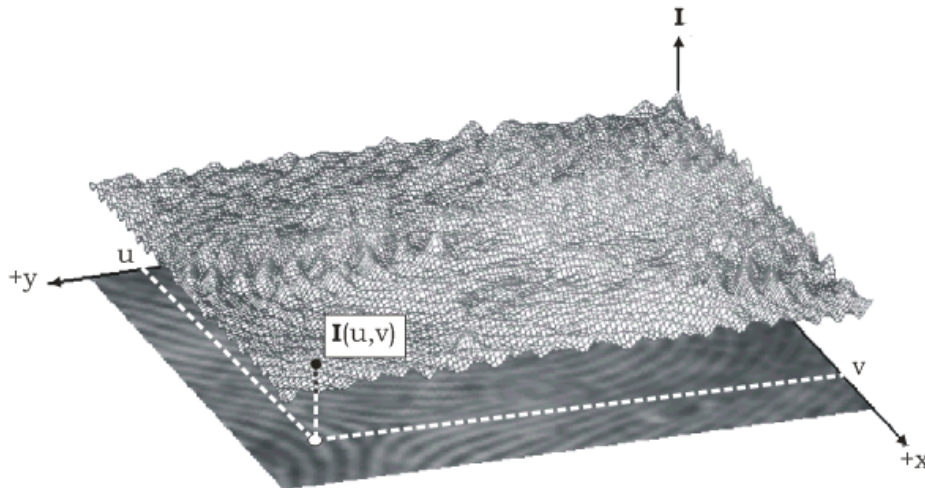


Figura 3.1: Representação da imagem como uma superfície tridimensional ou como uma matriz de pontos (píxels) que variam no intervalo de 0 a 255 (escala de cinza).

3.2 Cálculo da imagem de regiões

Uma impressão digital geralmente está constituída de duas regiões (ver Figura 3.2): a região bem definida ou primeiro plano e a região corrompida ou segundo plano. Na primeira região, as estruturas de cristas e vales são claramente diferenciadas; esta região é usada na etapa de reconhecimento, formando-se com o primeiro contato da polpa dos dedos com a superfície do sensor. Na segunda região, as estruturas de cristas e vales não são claramente diferenciáveis devido à presença de ruído e à distorção na imagem como bordas, manchas, rugas entre outras; esta região é ignorada na etapa de reconhecimento e se forma naquelas áreas em que não se teve um contato adequado da polpa do dedo com a superfície do sensor.

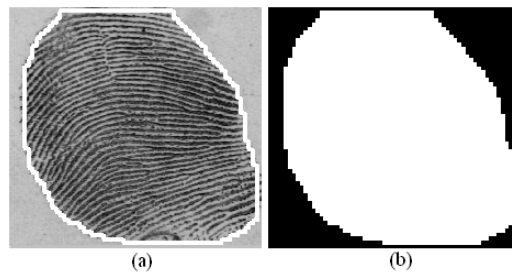


Figura 3.2: Regiões nas impressões digitais, (a) imagem original; (b) imagem de regiões (zonas de cor preto são definidas como segundo plano e as brancas como primeiro plano).

A tarefa da técnica de segmentação é decidir qual parte da imagem corre-

sponde ao primeiro ou segundo plano. Uma correta segmentação é especialmente importante para a confiável extração das características. A maioria de técnicas de extração de características extraem falsas características quando são aplicados ao segundo plano. O principal objetivo da técnica de segmentação é detectar o segundo plano e portanto reduzir o número de falsas características, aumentando a robustez do sistema. Vários abordagens sobre segmentação de impressões digitais são dadas na literatura.

- Em [73] a impressão digital é setorizada em blocos de 16x16 pixels, calculando-se para cada bloco o histograma das orientações das cristas. A presença de um pico significativo num histograma denota um padrão orientado (primeiro plano), um histograma aproximadamente plano representa um padrão uniforme (segundo plano).
- Em [72] propuseram uma melhora sobre a técnica proposta em [73] utilizando um limiar para a segmentação de acordo com a variância de cada bloco.
- Em [86] a impressão digital é dividida em blocos de 16x16 pixels por bloco, calculando-se para cada bloco a variância dos níveis de cinza na direção normal à orientação da crista, derivando-se deste um índice de qualidade. A idéia principal é que nas regiões ruidosas não existirá uma dependência direcional, pois o primeiro plano mostra uma alta variância na direção normal à orientação das cristas e uma baixa variância na direção destas.
- Em [66] a impressão digital é dividida em blocos de 16x16 pixels, determinando-se a média da magnitude do vetor gradiente de cada bloco; posto que uma impressão digital está composta por cristas e vales semi-paralelas (orientadas), o vetor gradiente tem uma alta magnitude no primeiro plano e um valor baixo no segundo plano.
- Em [107] a impressão digital é dividida em blocos de 15x15 pixels e, para cada bloco, calculam-se onze características. Através de uma rede neural perceptron multi-camadas, classifica-se cada bloco como pertencente ao primeiro ou segundo plano. A imagem resultante é processada, para eliminar os blocos mal classificados, através do uso de quatro características baseadas no contraste e a amplitude tanto global como local dos blocos. Por último, inclui-se uma etapa de revisão da imagem segmentada baseada num procedimento heurístico.
- Em [70] particiona-se a imagem em blocos de 32x32 pixels, extraindo-se descritores de Fourier para cada bloco e usa-se uma rede neural

perceptron multi-camada para classificar cada bloco como pertencente ao primeiro ou segundo plano em função de seus descritores.

- Em [6] propõe-se uma técnica de segmentação por pixel, usando a coerência. Quando o nível de coerência para cada pixel é menor que um limiar considera-se que pertence ao segundo plano, caso contrário, pertence ao primeiro plano, incluindo-se ao final uma etapa de pós-processamento baseado em morfologia matemática que reduz as regiões mal classificadas.
- Em [7] melhora-se a técnica anterior incluindo-se duas características mais, a variância e o valor médio de cada pixel utilizando-se como classificador um perceptron, posto que como a classificação é realizada ao nível de pixels o custo computacional de usar classificadores mais sofisticados aumentaria o tempo de processamento.

3.2.1 Técnica de segmentação baseada no mínimo valor de intensidade

Como foi exposto, duas são as vertentes ao momento de obter as características para a segmentação da imagem, baseados em blocos e pixels. Neste trabalho, segue-se a primeira opção, pela capacidade de poder diminuir o tempo de processamento e assim poder incrementar a complexidade algorítmica. Por tanto, define-se formalmente a imagem de regiões Π vinculada à impressão digital de entrada I como uma matriz de tamanho $N_{Fil} \times N_{Col}$, em que $\Pi(i, j)$ indica uma categoria relacionada com o bloco (i, j) . Um bloco $\Pi(i, j)$ com valor 0 indica que os pixels correspondentes de I pertencem a uma região não recuperável (segundo plano), no caso contrário (valor 1) os pixels correspondentes de I pertenceram a uma região recuperável (primeiro plano).

A técnica proposta apresenta as seguintes partes:

Setorização retangular da imagem Especifica-se o número de blocos e o tamanho de cada bloco em que vai dividir-se a imagem.

Extração de características e Segmentação Para cada bloco define-se como característica para a segmentação o mínimo valor na escala de cinza que esta tenha. A matriz de características obtida é processada para eliminar os dados que geram uma má segmentação.

Vejamos cada um dos passos anteriormente indicados.

Setorização retangular da imagem

A divisão por blocos de uma imagem e a posterior obtenção de um valor representativo de cada bloco implica numa diminuição dimensional dos dados. Suponha uma imagem de tamanho $N \times N$ pixels, se ela é dividida em blocos de tamanho $N_k \times N_k$ pixels (em [62] e [61] recomenda-se $N_k = 2^k$) obtém-se $(N/N_k)^2$ valores representativos, ou seja se produz uma redução de N_k^2 dados acelerando assim todo processamento posterior. Para o caso de I , a divisão em blocos de tamanho $N_k \times N_k$ produzirá uma imagem de regiões Π de tamanho $N_{Fil} \times N_{Col}$ em que $N_{Fil} = \lfloor N_{Fil}^I / N_k \rfloor$ e $N_{Col} = \lfloor N_{Col}^I / N_k \rfloor$ (ver Figura 3.3). Se denotamos por S_{ij}^k , ao (i, j) -ésimo bloco vinculado a I de tamanho $N_k \times N_k$, as posições na imagem I correspondentes a dito bloco são:

$$S_{ij}^k = \{ (u, v) \mid (i-1)N_k \leq u \leq iN_k - 1, (j-1)N_k \leq v \leq jN_k - 1 \}. \quad (3.1)$$

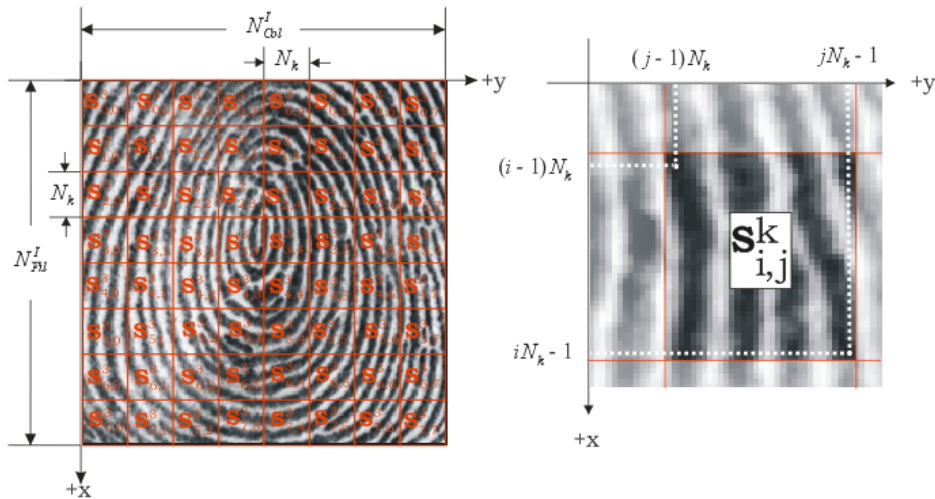


Figura 3.3: Setorização da imagem em blocos retangulares.

Extração de características e Segmentação

Como é indicado em [91], as técnicas de segmentação também podem ser classificadas pelo tipo de característica usada para a tarefa de segmentação. Tem-se aquelas baseadas nos níveis de intensidade [72], [86], [70] e aquelas baseadas na orientação das cristas [73], [66], [6]. Todas as técnicas têm um bom desempenho se o padrão de cristas está relativamente bem definido e tem uma boa orientação. Aqui, se utilizará uma característica da imagem baseada

nos valores de intensidade, o qual é o mínimo valor na escala de cinza de cada bloco em que é dividida a imagem. Esta escolha está fundamentada no fato de que aquelas zonas com uma mínima presença de cristas possivelmente terão um valor mínimo perto de zero e, dado que as cristas em uma imagem aparecem como um padrão e não de maneira isolada, estes valores mínimos estarão conectados. Com essa escolha, é possível detectar a mínima presença de padrões de cristas, fato que é necessário em imagens de baixa qualidade. Para tal tarefa, desenvolve-se o seguinte procedimento:

1. A imagem de entrada é equalizada adaptativamente utilizando a técnica CLAHE (*Contrast-Limited Adaptive Histogram Equalization*)[71], gerando-se a imagem I_{equ} .
2. Para cada bloco S_{ij}^k de I_{equ} é determinado o mínimo valor, obtendo-se a matriz de valores mínimos I_{\min} :

$$I_{\min}(i, j) = \min_{(u,v) \in S_{ij}^k} \{I_{equ}(u, v)\}. \quad (3.2)$$

Nesta matriz tem-se a presença de máximos isolados entre grupos de mínimos conectados (ver Figura 3.4.c), produto de regiões em que as cristas estão cortadas, apagadas ou são muito finas e pouco claras. Essas regiões dificultam a tarefa de segmentação e é necessária sua eliminação, para o qual, aplicamos um filtro de mediana, obtendo-se assim uma matriz em que a maioria destes pontos são retirados. (ver Figura 3.4.d)

3. A matriz I_{\min} já filtrada, é binarizada para um intervalo de limiares. Cada limiar gera uma imagem binária, em que, o primeiro plano é definido como os componentes conectados de 1 e o segundo plano é definido como os componentes conectados de 0. Para cada uma destas imagens binárias aplica-se o seguinte pós-processamento:
 - Eliminam-se as regiões espúrias do primeiro plano (porções isoladas do segundo plano presentes no primeiro plano), ou seja, todos os componentes conectados pertencentes ao segundo plano, menos o maior, são transformados em primeiro plano.
 - Eliminam-se as regiões espúrias do segundo plano (porções isoladas do primeiro plano presentes no segundo plano), ou seja, todos os componentes conectados pertencentes ao primeiro plano, menos o maior, são transformados em segundo plano.

- Extraímos uma vizinhança centrada no centróide do maior componente conectado pertencente ao primeiro plano de tamanho $P \times Q$ e determinamos o número de elementos vinculados às filas e colunas que são zero na vizinhança extraída, Nz . Se $Nz > 0,7PQ$ a imagem binária é descartada, caso contrario, calcula-se a razão $Ts = Arv/(PQ - Nz)$, Arv é a área do primeiro plano da vizinhança extraída.
 - Se $Ts < 0,75$ a imagem binária é dilatada, porque se supõe que existe a presença de regiões espúrias no primeiro plano. Se $0,75 \leq Ts \leq 0,90$ a imagem binária é erodida, porque se supõe que existem pontos duvidosos pertencentes ao contorno do primeiro plano.
 - Repete-se os passos 1 e 2 e calcula-se a área do primeiro plano da imagem binária resultante.
4. Define-se como Π aquela imagem binária que tem a maior área de primeiro plano.
 5. Suavizamos o contorno do primeiro plano de Π , fazendo uma varredura por linhas e colunas, detectando todas as trocas de $0 \rightarrow 1$ e de $1 \leftarrow 0$. Se a distância entre trocas é menor que um limiar, então os valores dos elementos entre as trocas respectivas são invertidos de valor. Este passo é repetido para diferentes limiares.

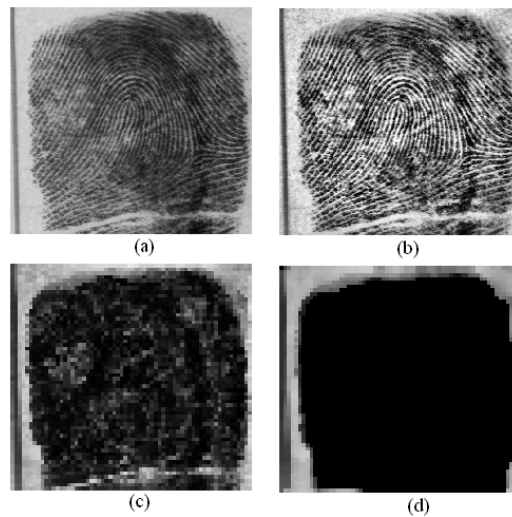


Figura 3.4: (a)Imagem original; (b)imagem equalizada localmente; (c)matriz de valores mínimos antes de filtrar; (d)matriz de valores mínimos depois de filtrar.

É evidente que esta técnica considera a hipótese de que o primeiro plano é um só componente conectado, não tendo buracos, o que na prática não necessariamente é assim. Isto não invalida a técnica posto que bastará incluir outra etapa de segmentação, mais fina, unicamente no primeiro plano, que permita detectar aquelas zonas que possam ser recuperadas por uma técnica de melhora e classificadas como pertencentes ao primeiro plano ou caso contrário considera-se como irrecuperáveis e portanto pertencentes ao segundo plano.

Através de uma inspeção visual pode-se ver que a técnica apresenta resultados satisfatórios para imagens de baixa qualidade. Na Figura 3.5, são apresentadas imagens segmentadas quando elas tem: buracos, cortes, variação de contraste e com anotações.

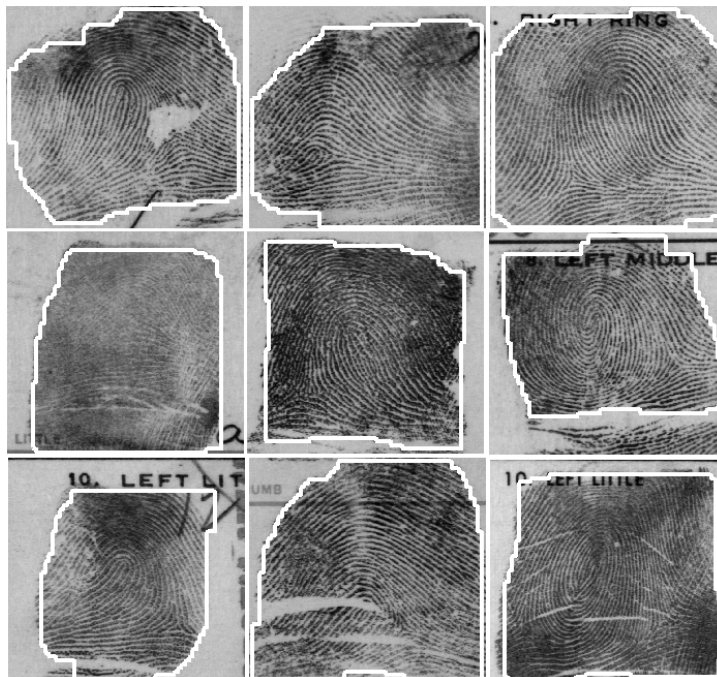


Figura 3.5: Resultados da segmentação. As linhas brancas indicam as fronteiras entre o primeiro e segundo plano definido por Π .

3.3 Melhora da imagem

O principal objetivo desta etapa é atenuar os efeitos devido ao ruído na imagem de entrada, tal que, gera-se uma imagem adequada para a extração das características utilizadas nas etapas subsequentes. Para o caso de classificação, se requer uma imagem que realce a informação direcional intrínseca da

impressão digital, tanto nas zonas de boa como de má qualidade. Dividimos o processo de melhora em duas fases:

Normalização A imagem de entrada I é normalizada, melhorando o contraste dela.

Melhora no domínio da Freqüência À imagem normalizada aplica-se a transformada discreta de Fourier bidimensional na sua forma local.

Vejamos cada um dos passos anteriormente indicados.

3.3.1 Normalização

O propósito de normalizar é reduzir as variações dos níveis de cinza de cristas e vales, além de uniformizar as variações de contraste entre as diferentes zonas da imagem, o que facilitará os passos subseqüentes.

- Em [44], toda a imagem é normalizada num valor médio e com variância estabelecida.
- Em [13], a técnica anterior é melhorada dividindo a imagem em blocos e os parâmetros para a normalização são determinados adaptativamente de acordo com as estatísticas de cada bloco.

Ambas técnicas geram imagens normalizadas que não conseguem, em muitos casos, diminuir a variação dos níveis de cinza de cristas e vales, e menos ainda, uniformizar o contraste em toda a imagem. Para que tal tarefa seja realizada, propõe-se o seguinte procedimento:

1. Para todo pixel (u, v) de I , extraímos uma vizinhança de tamanho $N_{Uf} \times N_{Uf}$, e realizamos:

$$I_{U_{nf}}(u, v) = \frac{I(u, v) - \min}{\max - \min}, \quad (3.3)$$

em que, \max e \min são o máximo e o mínimo relativo à vizinhança centrada em (u, v) .

2. Aplica-se à $I_{U_{nf}}$ a transformação morfológica *Tophat* multiescala [78], [29].

$$I_{Nor}(u, v) = I_{U_{nf}}(u, v) + 0.5 \sum_{i=n}^m HT_{brc}^{iB}(u, v) - 0.5 \sum_{i=n}^m HT_{prt}^{iB}(u, v) \quad (3.4)$$

onde:

$HT_{brc}^{iB}(u, v) = (I_{Unf} - I_{Unf} \circ iB)(u, v)$ e $HT_{prt}^{iB}(u, v) = (I_{Unf} \bullet iB - I_{Unf})(u, v)$ são as transformações morfológicas *Tophat* multiescala branca e preta, respectivamente, tal que B é um elemento estruturante que incrementa de tamanho segundo é indicado por i e os operadores " \circ " e " \bullet " denotam as operações morfológicas de abertura e fechamento em escala de cinza, respectivamente.

O passo (1) permite reduzir a variação dos níveis de cinza das cristas e vales e o passo (2) permite uniformizar o contraste entre diferentes zonas da imagem e sobretudo ressaltar as cristas sobre os vales. Na Figura 3.6 pode-se ver a imagem resultante e como o contraste das diferentes zonas de I são uniformizadas.

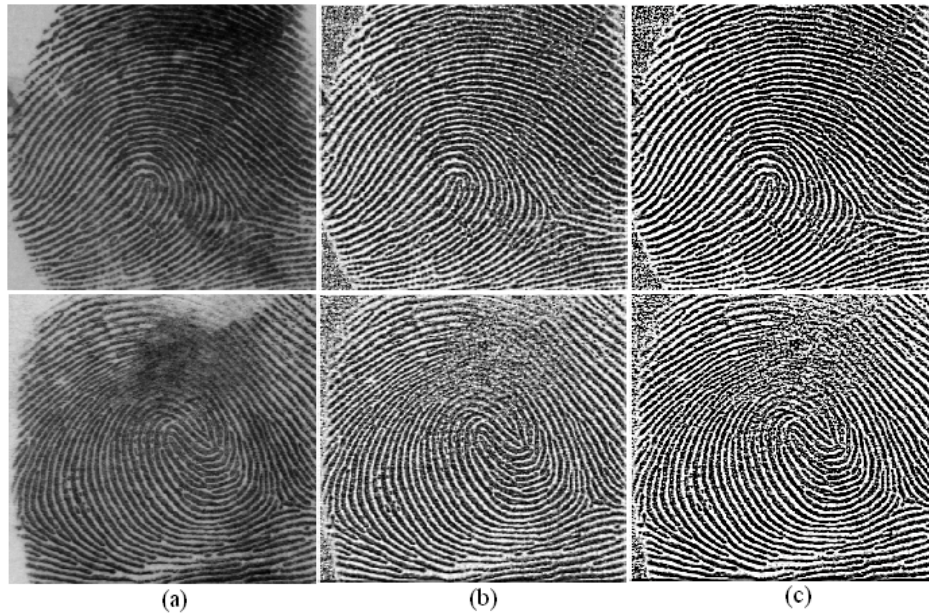


Figura 3.6: Resultados da Normalização: (a) I ; (b) I_{Unf} ; (c) I_{Nor} .

3.3.2 Melhora no domínio da Frequência

Posto que uma impressão digital é essencialmente uma textura orientada, sua mais importante característica intrínseca, sua orientação local dominante, está contida no módulo da transformada de Fourier bidimensional. As técnicas de melhora que aproveitam esta característica foram propostas por [98] e [101]. Em ambas, a imagem normalizada é setorizada em blocos, com um

determinado nível de sobreposição por lado (a sobreposição entre blocos da imagem de entrada reduz as descontinuidades entre blocos adjacentes na imagem melhorada) e para cada bloco sobreposto aplica-se, a função não linear:

$$I_{Mej}(u, v) = \left\| IDFT2 \left\{ \left\| DFT2 \{ I_{Nor}(u, v) \} \right\|^k DFT2 \{ I_{Nor}(u, v) \} \right\} \right\|, \quad (3.5)$$

tal que: $(u, v) \in S_{ij}^{N_B, N_{TR}}$, onde, $S_{ij}^{N_B, N_{TR}}$ é o bloco (i, j) -ésimo vinculado a I_{Nor} de tamanho $N_B \times N_B$ e com um sobreposição de N_{TR} pixels, $DFT2$ e $IDFT2$ são as transformadas bidimensionais de Fourier direta e inversa, k é o expoente ao qual é elevada a densidade espectral de energia de cada bloco da imagem e $\|\bullet\|$ toma a magnitude do operando.

Em [98] se estabeleceu $k = 0, 6$ e em [101] propõe-se $k = 1, 4$. O tempo de processamento está diretamente vinculado ao tamanho do bloco e o nível de sobreposição, portanto, estamos diante de um compromisso: reduzir a sobreposição gera distorções entre bordas de cada bloco e aumentá-la em grande medida incrementa o tempo de processamento.

A função não linear atua de maneira equivalente a aplicar um limiar (adaptativo) à densidade espectral de energia, tal que, a medida que k aumenta, a diferença entre o mínimo e máximo valor na densidade espectral de energia também aumenta, chegando um momento em que só se preservam os valores máximos. Estes valores máximos estão vinculados aos componentes sinusoidais fundamentais do espectro de Fourier, para o caso de uma vizinhança de uma impressão digital, são aqueles componentes sinusoidais fundamentais os que conformam a textura orientada, tal que ao retornar ao domínio espacial só se obterá a textura orientada reconstruída a partir de suas componentes sinusoidais fundamentais. Na Figura 3.7 tem-se as imagens melhoradas considerando $N_B = 32$, $N_{TR} = 20$ e $k = 1, 0$.

3.4 Cálculo da imagem de orientação

A estimação da imagem de orientação tem um importante papel num sistema de reconhecimento de impressões digitais, sendo esta muito utilizada para o projeto de filtros espaciais adaptativos na etapa de melhora [44], [38], [39], [93], alinhamento e coincidência na tarefa de identificação [104], [95], e mais importante ainda como base para a geração de características na tarefa de classificação de impressões digitais, [16], [15], [50]. Outra característica de importância, vinculada à orientação local das cristas, é a anisotropia ou coerência, que tem utilidade no cálculo da imagem de regiões [7], [88], na determinação da qualidade de uma impressão digital [22] e na detecção do

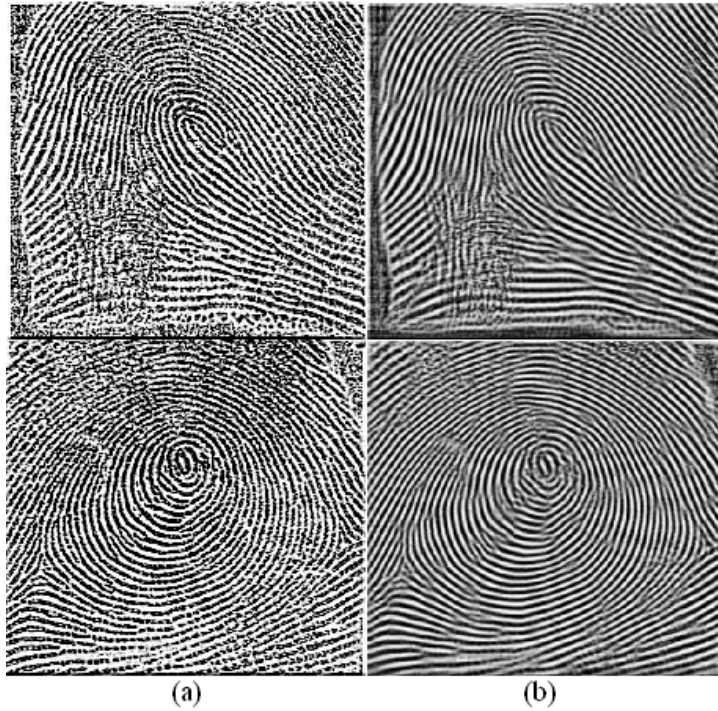


Figura 3.7: Melhora no domínio da frequência (a) imagem normalizada; (b) imagem melhorada.

ponto de referência da impressão digital [60]. Como se pode ver, a imagem de orientação e a coerência, são de importância na implementação de um sistema de reconhecimento de impressões digitais, portanto uma estimação robusta de ditos parâmetros é mais que necessária para uma melhora geral no rendimento do sistema.

- Define-se a orientação local de uma crista no píxel (u, v) de I como o ângulo que se forma entre a crista pertencente a uma pequena vizinhança centrada em (u, v) e o eixo horizontal da imagem (ver Figura 3.8). As orientações locais são, em princípio, perpendiculares aos vetores gradiente, pelo que podem ser derivados a partir destes. A imagem de orientação Θ também conhecida como imagem ou campo direcional, é definida como uma matriz de tamanho $N_{Fil} \times N_{Col}$ onde $\Theta(i, j)$ representa a orientação média normal das cristas do bloco (i, j) de I , tal que: $N_{Fil} \leq N_{Fil}^I$ e $N_{Col} \leq N_{Col}^I$.
- Define-se a coerência como um indicador de anisotropia normalizada que contém informação sobre a homogeneidade numa região da imagem. Se todos os vetores gradiente numa vizinhança apontam exatamente na

mesma direção, a coerência é igual a 1. Por outro lado, se os vetores gradiente numa vizinhança estão igualmente distribuídos em todas as direções, a coerência é igual a 0, ou seja aquelas regiões com orientação dominante tem um valor de coerência perto de 1, enquanto as regiões constantes(isotrópicas) tem uma coerência perto de 0.

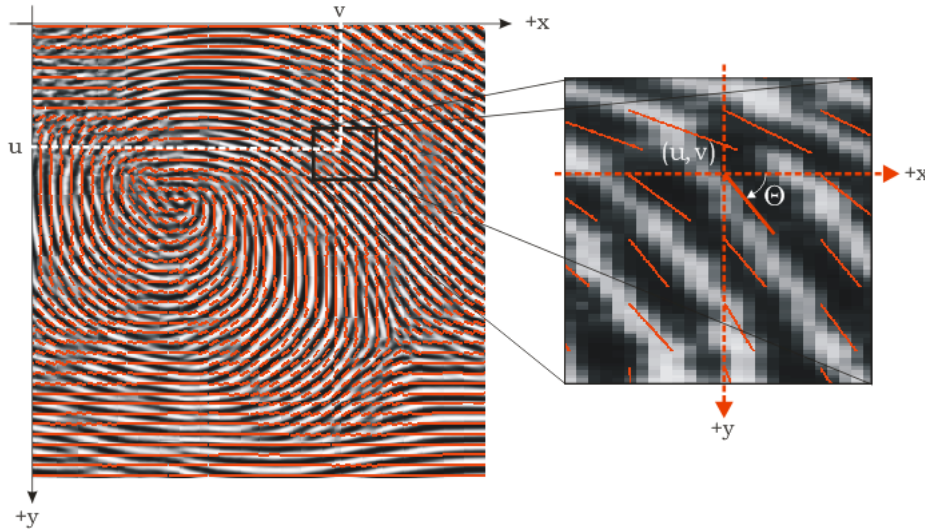


Figura 3.8: Imagem de orientação de uma impressão digital, cada linha representa a orientação média das cristas numa vizinhança de tamanho definido.

3.4.1 Obtenção da imagem de orientação

Várias técnicas para o cálculo da imagem de orientação foram propostas no passado. As técnicas mais consolidadas são baseadas na média quadrática dos gradientes ou na análise da componente principal da matriz de covariância dos vetores gradiente da imagem. Esta técnica foi proposta por [57], [30] e adaptada para a estimação da imagem de orientação por [86], [48], [5] e [40]. Ela consiste principalmente dos seguintes passos:

Cálculo do vetor gradiente Determina-se as componentes do vetor gradiente G_x e G_y , capturando assim a informação de orientação de todos os pixels da imagem.

Estimação da orientação local Calcula-se a orientação local média do vetor gradiente, permitindo assim atenuar o ruído contido nele.

Vejamos cada um dos passos anteriormente indicados.

Cálculo do vetor gradiente

O primeiro passo é determinar $G_x(u, v)$ e $G_y(u, v)$. Existem diferentes técnicas para realizar tal tarefa: baseadas em máscaras espaciais [37], utilizando morfologia matemática [32] entre outros. Por sua simplicidade e menor custo computacional nossa eleição foi a utilização de máscaras espaciais, especificamente a mascara de Sobel.

Estimação da orientação local

Nesta seção calculamos a orientação dominante de cada bloco utilizando a média do vetor gradiente ao quadrado proposto por [5] e [40] com as melhoras propostas por [52] e [97].

O primeiro passo é calcular o quadrado da representação complexa do vetor gradiente para cada pixel da imagem, definido como $G_p(u, v)$, o que equivale a duplicar os ângulos dos vetores gradiente. Esta operação é necessária, posto que os vetores gradiente não podem ser simplesmente ponderados em alguma vizinhança local, já que os vetores opostos se cancelariam ainda tendo a mesma orientação, devido a que as estruturas de cristas e vales locais não trocam quando são rotacionadas 180° .

$$\begin{aligned} G_d(u, v) &= (G_x(u, v) + jG_y(u, v))^2 \\ G_d(u, v) &= (G_x^2(u, v) - G_y^2(u, v)) + j(2G_x(u, v)G_y(u, v)) \end{aligned} \quad (3.6)$$

logo, gera-se a setorização S_{ij}^k e para cada bloco S_{ij}^k , calculamos o vetor gradiente quadrático médio $G_p(i, j)$:

$$\begin{aligned} G_p(i, j) &= \frac{1}{N_k^2} \sum_{(u,v) \in S_{ij}^k} G_d(u, v) \\ G_p(i, j) &= (\Gamma_{xx}(i, j) - \Gamma_{yy}(i, j)) + j(2\Gamma_{xy}(i, j)) \end{aligned} \quad (3.7)$$

onde:

$$\Gamma_{xx}(i, j) = \frac{1}{N_k^2} \sum_{(u,v) \in S_{ij}^k} G_x^2(u, v), \quad (3.8)$$

$$\Gamma_{yy}(i, j) = \frac{1}{N_k^2} \sum_{(u,v) \in S_{ij}^k} G_y^2(u, v), \quad (3.9)$$

$$\Gamma_{xy}(i, j) = \frac{1}{N_k^2} \sum_{(u,v) \in S_{ij}^k} G_x(u, v)G_y(u, v), \quad (3.10)$$

Finalmente, calcula-se a orientação média do vetor gradiente ao quadrado $\Phi(i, j)$ através da relação:

$$\Phi(i, j) = \frac{1}{2} \arctan \left(\frac{2\Gamma_{xy}(i, j)}{\Gamma_{xx}(i, j) - \Gamma_{yy}(i, j)} \right), \quad (3.11)$$

tal que: $\Phi(i, j) \in [-\pi/2, \pi/2]$ e $\arctan(y/x)$ é o arco tangente de y/x tendo em conta a sinal de x e y .

O indicador de coerência será definido pela relação:

$$Coh(i, j) = \frac{\|G_p(i, j)\|}{\frac{1}{N_k^2} \sum_{(u,v) \in S_{ij}^k} \|G_d(u, v)\|} \quad (3.12)$$

De (3.7) podemos concluir que tanto o módulo e o argumento de $G_d(u, v)$ são incorporados na operação de média, o qual afeta os resultados da estimação de $\Phi(i, j)$ e $Coh(i, j)$, por exemplo, os $G_d(u, v)$ com uma grande magnitude tem mais importância no processo de média que os $G_d(u, v)$ com uma magnitude pequena, o equivalente também se cumpre para a coerência. Em tal sentido, Jiang [52] propõe uma técnica que enfrenta este problema ao incluir uma segunda fase de média ponderada no cálculo da orientação local.

Primeiro, o setor retangular S_{ij}^k é particionado em blocos não sobrepostos Λ_r , tal que: $S_{ij}^k = \Lambda_1 \cup \Lambda_2 \cup \dots \cup \Lambda_r \cup \dots \cup \Lambda_M$, onde cada bloco Λ_r estará conformado por N_Λ pixels.

Logo, para cada Λ_r calcula-se o vetor gradiente quadrático médio e sua respectiva coerência:

$$G_p(r) = \frac{1}{N_\Lambda} \sum_{(m,n) \in \Lambda_r} G_d(m, n), \quad (3.13)$$

$$Coh(r) = \frac{\|G_p(r)\|}{\frac{1}{N_\Lambda} \sum_{(m,n) \in \Lambda_r} \|G_d(m, n)\|}. \quad (3.14)$$

Finalmente determinamos a média dos vetores gradientes quadrados médios, ponderados no bloco S_{ij}^k :

$$G_{pn}(i, j) = \frac{1}{M} \sum_{r=1}^M w(r)G_p(r), \quad (3.15)$$

onde o valor de ponderação vem definido por:

$$w(r) = \frac{Coh(r)}{\|G_p(r)\|}, \quad (3.16)$$

Substituindo-se (3.16) em (3.15) tem-se:

$$G_{pn}(i, j) = \frac{1}{M} \sum_{r=1}^M Coh(r) \frac{G_p(r)}{\|G_p(r)\|}, \quad (3.17)$$

e substituindo-se (3.14), (3.7) e (3.6) em (3.17) tem-se:

$$G_{pn}(i, j) = \frac{1}{M} \sum_{r=1}^M \frac{(\Gamma_{xx}(r) - \Gamma_{yy}(r)) + j(2\Gamma_{xy}(r))}{\Gamma_{xx}(r) + \Gamma_{yy}(r)}. \quad (3.18)$$

Finalmente, calculamos $\Phi(i, j)$ e $Coh(i, j)$, através das relações:

$$\Phi(i, j) = \frac{1}{2} \arctan \left(\frac{\text{Im} \{G_{pn}(i, j)\}}{\text{Re} \{G_{pn}(i, j)\}} \right) \quad (3.19)$$

$$Coh(i, j) = \frac{\|G_{pn}(i, j)\|}{\frac{1}{M} \sum_{r=1}^M Coh(r)} \approx \|G_{pn}(i, j)\| \quad (3.20)$$

Outra técnica, baseada nos vetores gradiente, é proposta em [97], neste trabalho, usa-se sobreposição de blocos e um esquema de votação para o cálculo de $\Phi(i, j)$. Em função de cada bloco S_{ij}^k definem-se os seguintes agrupamentos:

$$\begin{aligned} D_1 &= S_{i-1, j-1}^k \cup S_{i-1, j}^k \cup S_{i, j-1}^k \cup S_{i, j}^k \\ D_2 &= S_{i-1, j}^k \cup S_{i-1, j+1}^k \cup S_{i, j}^k \cup S_{i, j+1}^k \\ D_3 &= S_{i, j-1}^k \cup S_{i, j}^k \cup S_{i+1, j-1}^k \cup S_{i+1, j}^k \\ D_4 &= S_{i, j}^k \cup S_{i, j+1}^k \cup S_{i+1, j}^k \cup S_{i+1, j+1}^k \end{aligned} \quad (3.21)$$

tal que: $S_{i, j}^k = D_1 \cap D_2 \cap D_3 \cap D_4$, então para cada agrupamento D_r determina-se sua respectiva Φ_r e Coh_r , através das relações (3.11) e (3.12), definindo como o $\Phi(i, j)$ de S_{ij}^k ao Φ_r que esteja vinculado ao máximo Coh_r .

Após de calcular Φ , determina-se a imagem de orientação Θ . Esta é estimada com robustez ao aplicar-se uma operação de suavização sobre Φ . Por seu menor custo computacional utilizamos a técnica proposta em [56], onde, primeiro as componentes sinusoidais de Φ (ou seja $[\cos(2\Phi), \sin(2\Phi)]$) são filtradas com uma máscara de média w_{sv} :

$$\Upsilon_x(i, j) = \sum_{(m,n) \in S_{sv}} w_{sv}(m, n) \cos(2\Phi(i + m, j + n)) \quad (3.22)$$

$$\Upsilon_y(i, j) = \sum_{(m,n) \in S_{sv}} w_{sv}(m, n) \sin(2\Phi(i + m, j + n)) \quad (3.23)$$

para então calcular Θ através das relações:

$$\Phi_{sv}(i, j) = \frac{1}{2} \text{arctg}\left(\frac{\Upsilon_y(i, j)}{\Upsilon_x(i, j)}\right) \quad (3.24)$$

$$\Theta(i, j) = \begin{cases} \Phi_{sv}(i, j) & \Phi_{sv}(i, j) \geq 0 \\ \Phi_{sv}(i, j) + \pi & \Phi_{sv}(i, j) < 0 \end{cases} \quad (3.25)$$

Nas Figuras 3.9 e 3.10 pode-se ver as imagens de orientação calculadas para as três técnicas, quando tem-se impressões de boa e má qualidade, onde a técnica exposta em [52] é a que apresenta o melhor comportamento frente as impressões digitais de má qualidade, posto que, o padrão de crestas e vales é representado adequadamente pela imagem de orientação ainda com presença de ruído.

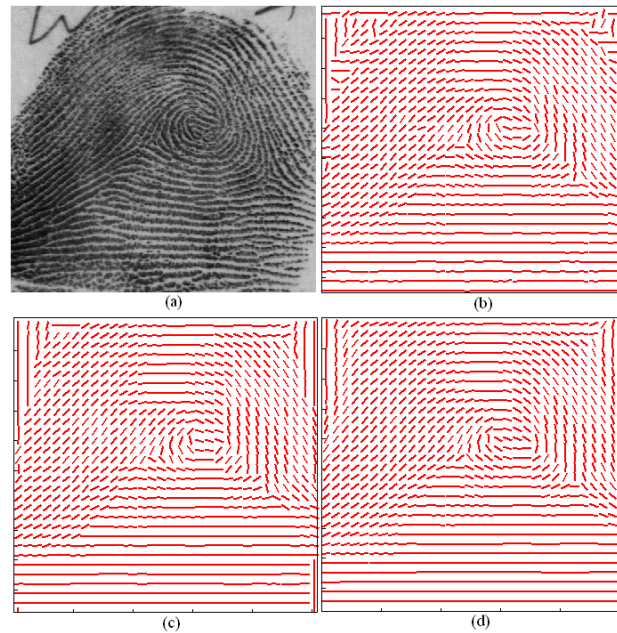


Figura 3.9: Imagem de orientação Θ para uma impressão de boa qualidade (a) impressão digital; (b) técnica padrão; (c) técnica proposta em [97]; (d) técnica proposta em [52].

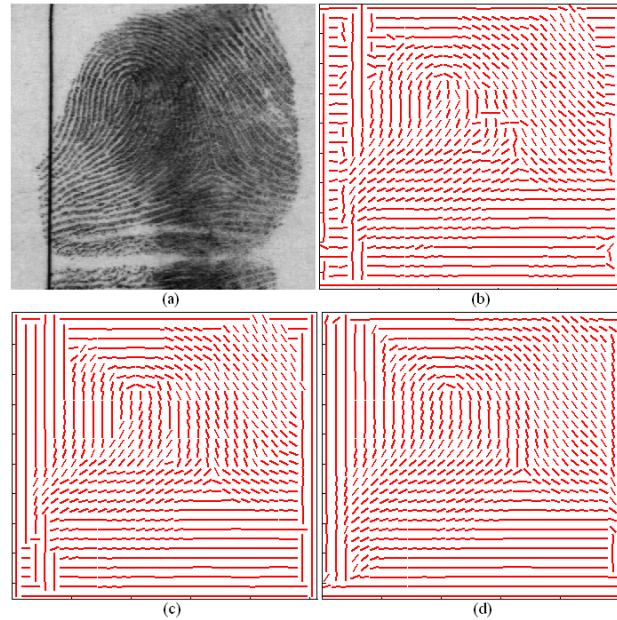


Figura 3.10: Imagem de orientação Θ para uma impressão de má qualidade (a) impressão digital; (b) técnica padrão; (c) técnica proposta em [97]; (d) técnica proposta em [52].

3.5 Registro de impressões digitais

Na maioria dos sistemas de reconhecimento de impressões digitais é estabelecida uma etapa que elimina a translação que apresentam as impressões digitais, permitindo uma maior robustez ao momento de extrair as características, tanto, para a etapa de coincidência como de classificação. Nesta seção, detalhamos as técnicas implementadas para realizar esta tarefa.

3.5.1 Determinação do ponto de referência de uma impressão digital

Um ponto de referência é estabelecido e usado para obter uma extração de características robusta à translação. Definimos como o ponto de referência de uma impressão digital aquele ponto que tem máxima curvatura na cristas mais interna (ver Figura 3.11).

Existe uma variedade de técnicas propostas para a detecção dos pontos singulares e a maioria deles opera na imagem de orientação, Θ , vejamos:

- Em [58] propuseram o uso do índice de Poincaré para a detecção das singularidades *delta* e *core*. Entretanto, apesar de ser eficiente, este índice é sensível ao ruído presente em Θ (zonas ruidosas na impressão

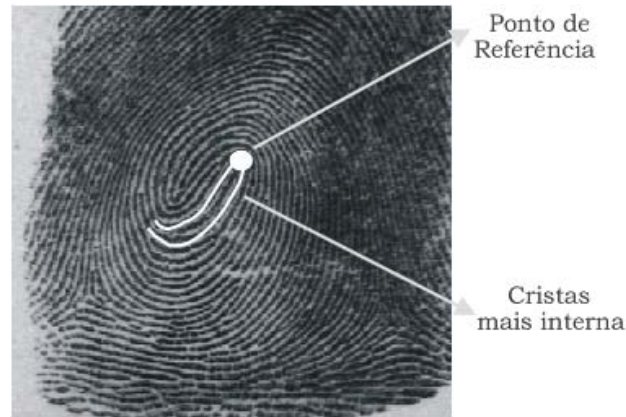


Figura 3.11: Ponto de referência numa impressão digital.

digital geram vizinhanças em Θ onde as orientações trocam rapidamente, gerando assim falsas singularidades). Além disso, esta técnica não detecta o ponto de referência para impressões digitais tipo *Arch* dado que o ponto com máxima curvatura não é o ponto *core* num sentido estrito.

- Em [51] propuseram uma técnica baseada na componente senoidal de Θ , que localiza o ponto de referência através da aplicação de uma máscara espacial, que permite obter as diferenças das componentes senoidais de duas regiões definidas sobre uma vizinhança centrada em todo bloco (i, j) de Θ . Esta técnica é robusta ao ruído, porém as duas regiões definidas são sensíveis à rotação da impressão digital. Além disso, esta técnica não detecta efetivamente o ponto de referência para impressões digitais tipo *Arch*.
- Em [81] propuseram uma técnica baseada no etiquetamento de Θ que permite a detecção do ponto de referência tipo *Arch*, porém é sensível ao ruído e à rotação que tenha a impressão digital.
- Em [60] propuseram uma técnica baseada numa análise multiescala do indicador de consistência da orientação (a consistência da orientação é equivalente à coerência) pela procura de um mínimo local. Esta técnica é robusta à rotação e ao depender unicamente do indicador da consistência da orientação torna-se sensível ao ruído da impressão digital.

Neste trabalho descrevem-se duas técnicas para a localização do ponto de referência, a primeira, baseada na relação que tem o ponto de referência

com o ponto singular *core* e a segunda, uma técnica proposta baseada numa análise de múltiplos níveis de representação da imagem de orientação.

Técnica baseada no ponto *Core*

Geralmente o ponto *core* é usado como ponto de referência. O número de pontos *core* difere de acordo com o tipo de impressão digital. Os tipos *Tented Arch* (TA), *Left Loop* (LL) e *Right Loop* (RL) têm um ponto *core*, enquanto o tipo *Whorl* (W) têm dois pontos *core*, o tipo *Arch* (A) não apresenta um ponto *core*. Por conseguinte, são necessárias duas técnicas de detecção, um para os tipos TA, LL, RL, W e outro para o tipo A.

Para o primeiro caso usamos a técnica baseada no índice de Poincaré, sendo uma das técnicas mais comumente usadas para a detecção dos pontos singulares *core* e *delta* [56], [106], [8], [96]. Esta técnica é simples e relativamente rápida comparada com outras técnicas [84], [8]. O problema de sua sensibilidade ao ruído presente em Θ é resolvido na etapa de estimação local da imagem de orientação, em que ao usar-se a técnica proposta por Jiang [52] as zonas ruidosas de Θ são atenuadas em grande medida. Para o segundo caso usamos a técnica baseada na transformação de covariância da imagem de orientação proposta em [50]. O procedimento em termos gerais pode-se descrever nas duas seguintes etapas:

Técnica baseada no índice de Poincaré Determina-se através do índice de Poincaré o ponto *core* da impressão digital. Se dois *core* são detectados, o ponto de referência é atribuído ao *core* superior. Se unicamente um *core* é detectado este é assumido como o ponto de referência.

Técnica baseada na transformação de covariância Se pontos *core* não são detectados definimos uma imagem característica formada pelos maiores autovalores da matriz de covariância de cada vizinhança local de Θ . Entretanto a matriz é limiarizada e atribui-se como o ponto de referência o centróide do maior componente conectado da matriz limiarizada.

Agora vejamos cada uma das etapas com maior detalhe.

Técnica baseada no índice de Poincaré

Se C é uma curva fechada imersa em Θ , o índice de Poincaré $P_{\Theta,C}$ é definido como a rotação total das linhas de fluxo de Θ ao longo de C (Ver Figura 3.12), ou seja, é a integração das variações de Θ ao redor da curva C . O índice de Poincaré $P_{\Theta,C}$ para cada posição (i, j) de Θ é calculado como segue:

1. Seja C_r uma curva fechada que define uma seqüência ordenada de alguns elementos de Θ , onde o índice r denota o ordem da curva; para nosso caso C_r será o contorno de um quadrado de $N_{C_r} \times N_{C_r}$ elementos, tal que: $N_{C_r} = 2r + 1$ para $r > 0$, e se $r = 0$, C_r será o contorno de um quadrado 2×2 . Tal contorno é percorrido em direção antihorária, como é indicado na Figura 3.12.

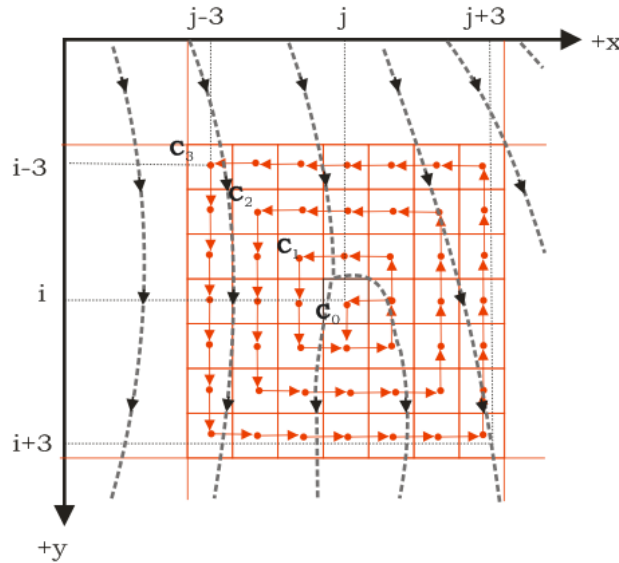


Figura 3.12: Curvas fechadas para a obtenção do índice de Poincaré.

O tamanho de C_r é crucial para o rendimento da técnica. Se é muito pequeno, então uma pequena perturbação na orientação pode resultar na detecção de pontos singulares espúrios. Por outro lado, se é muito grande, pontos singulares verdadeiros podem ser ignorados.

2. $P_{\Theta, C}(i, j)$ é calculado pela soma algébrica das diferenças das orientações entre elementos adjacentes da curva C_r . Se é considerado ϵ_x e ϵ_y como as coordenadas x e y de C_r , quando é o contorno de um quadrado centrado em $(0, 0)$, o índice de Poincaré poderá ser calculado através da relação:

$$P_{\Theta, C}(i, j) = \sum_{k=0}^{N_{C_r}^2 - 1} \Delta_{ij}(k), \quad (3.26)$$

onde:

$$\Delta_{ij}(k) = \begin{cases} \delta_{ij}(k) & |\delta_{ij}(k)| < \pi/2 \\ \pi + \delta_{ij}(k) & \delta_{ij}(k) \leq -\pi/2, \\ \pi - \delta_{ij}(k) & \text{outro caso} \end{cases}, \quad (3.27)$$

tal que:

$$\delta_{ij}(k) = \Theta(i + \epsilon_x(k+1), j + \epsilon_y(k+1)) - \Theta(i + \epsilon_x(k), j + \epsilon_y(k)). \quad (3.28)$$

$P_{\Theta,C}(i, j)$ toma unicamente valores próximos a $0^\circ, \pm 180^\circ, \pm 360^\circ$ e dependendo do valor tomado, o ponto (i, j) poderá ser catalogado como: um ponto comum se é próxima a 0° , um ponto singular *core* se é próximo a 180° e *delta* se é próximo a -180° , tal como é mostrado na Figura 3.13.

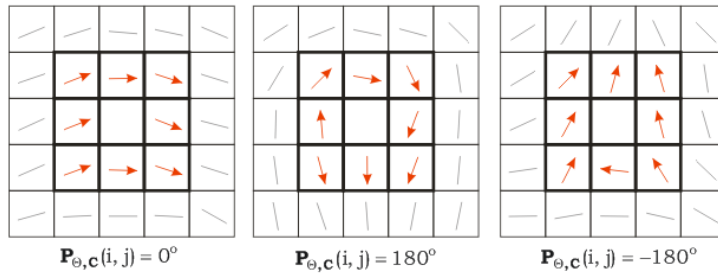


Figura 3.13: Cálculo do índice de Poincaré para $r = 1$.

3. Após calcular $P_{\Theta,C}$ para cada ponto de Θ realiza-se um pós-processamento para a eliminação daquelas falsas singularidades devidas ao ruído presente na impressão digital. Primeiro todas aquelas singularidades que pertencem ou estão muito perto ao segundo plano são descartadas. As singularidades restantes são agrupadas, os centróides gerados são considerados como pontos singulares válidos (Ver Figura 3.14).
4. Se dois pontos singulares *core* são detectados, o ponto de referência é atribuído ao *core* superior. Se unicamente um ponto singular *core* é detectado este é assumido como o ponto de referência, se pontos singulares *core* não são detectados, assume-se que estamos frente a uma impressão digital tipo A e o cálculo de seu ponto de referência será obtido pela técnica baseada na transformação de covariância.

Técnica baseada na transformação de covariância

A transformação de covariância é usada em [50] para determinar o centro de uma impressão digital tipo A. O cálculo da transformação de covariância

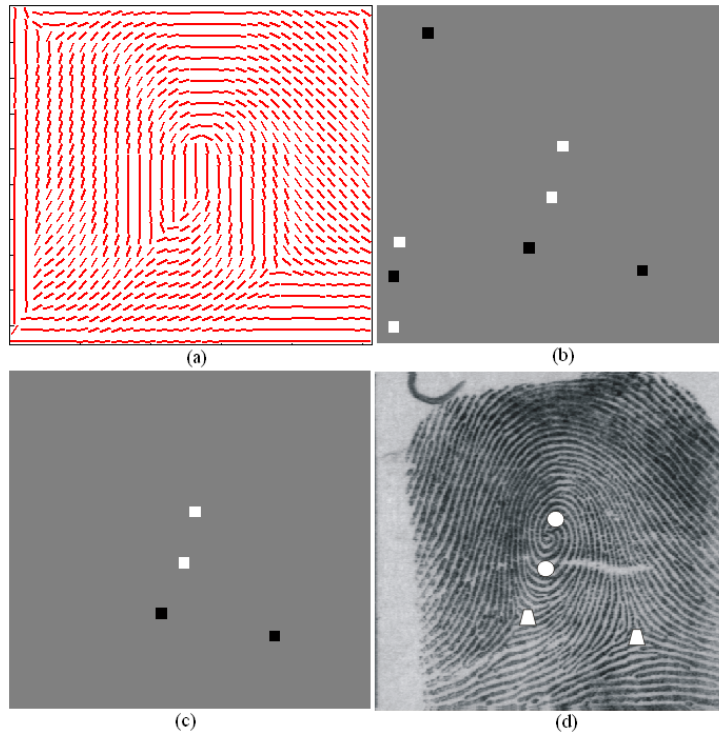


Figura 3.14: Detecção de singularidades através do índice de Poincaré (a) imagem de orientação; (b) índice de Poincaré; (c) singularidades válidas; (d); singularidades determinadas (pontos são *cores* e triângulos são *deltas*).

depende da determinação do maior autovalor da matriz de covariância associada a toda vizinhança centrada em cada elemento (i, j) de Θ . Os autovalores serão relativamente altos se a vizinhança tem uma orientação dominante e serão próximos a 0 se a vizinhança tem uma orientação uniforme. O procedimento para seu cálculo é dado por:

1. Para cada ponto (i, j) de Θ pertencente a uma região reconhecível da impressão digital, tomamos uma sub-matriz de dimensões qxq ($q = 3$) centrado em (i, j) denotada por $\Theta_{i,j}$. Todos os pontos da sub-matriz $\Theta_{i,j}$ devem pertencer ao primeiro plano de Π .
2. Estima-se a matriz de covariância de $\Theta_{i,j}$. Utilizando a relação:

$$\Sigma_{ij} = \frac{1}{q-1} \sum_{k=1}^q (\Theta_{ij}^k - U_{ij})(\Theta_{ij}^k - U_{ij})^T, \quad (3.29)$$

onde Θ_{ij}^k denota a k -ésima coluna da sub-matriz Θ_{ij} e U_{ij} denota o vetor médio das colunas da sub-matriz Θ_{ij} .

3. Em seguida determinamos o máximo autovalor de Σ_{ij} . O valor da transformação de covariância em (i, j) será o máximo autovalor correspondente à matriz de covariância Σ_{ij} , denotado por λ_{ij} .
4. Os passos de (1) a (3) são repetidos para todos aqueles (i, j) pertencentes ao primeiro plano de Π , obtendo assim a transformação de covariância, representada pela matriz Ψ (ver figura 3.15.b).
5. Ψ é limiarizada com um valor igual a $\max(\Psi)/6$ (onde $\max(\Psi)$ é o elemento máximo de Ψ), tal que, o valor de 6 foi determinado empiricamente [11]. Logo, procura-se as componentes conectadas na matriz limiarizada. Considera-se que o ponto de referência da impressão digital é o centro geométrico da componente conectada de maior área (ver figura 3.15.c).

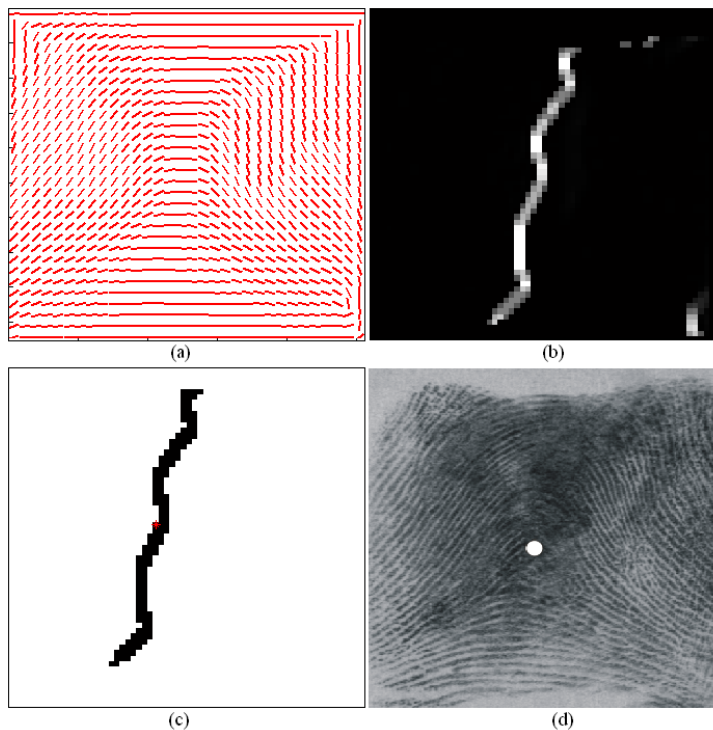


Figura 3.15: Detecção do ponto de referência para imagens tipo A (a) imagem de orientação; (b) transformação de covariância da imagem de orientação; (c) componente conectado de maior área da imagem binária da transformação de covariância; (d) ponto de referência da impressão digital tipo A.

Técnica baseada em múltiplos níveis de representação da imagem de orientação

A técnica anterior é eficiente na medida que a imagem de orientação seja suficientemente suavizada, de modo que as singularidades devidas ao ruído fiquem atenuadas e só se tenha a presença de singularidades devidas aos pontos *core* e *delta*. Além disso, como se verá no capítulo de resultados experimentais, a técnica apresenta uma baixa precisão no cálculo do ponto de referência para impressões tipo A, possivelmente devido a que o critério para a determinação do ponto de referência deste tipo de impressões está baseado numa característica geométrica que não necessariamente está vinculada a uma característica real que defina a posição do ponto de referência de uma impressão digital tipo A. Por tais motivos, tem-se a necessidade de elaborar uma técnica que não esteja fundamentada na detecção de singularidades da imagem de orientação e que tenha a capacidade de detectar o ponto de referência de qualquer tipo de impressão digital, tal que não apresente os inconvenientes da técnica baseada no ponto *core*. A seguir é explicada uma técnica proposta que cumpre com estes requisitos, baseada numa nova concepção a qual denominou-se "múltiplos níveis de representação da imagem de orientação".

Existe uma grande diferença entre as orientações das linhas de fluxo de cristas nas regiões superiores e inferiores ao ponto de referência. Pode observar, na Figura 3.16, que as linhas de fluxo de cristas na região superior ao ponto de referência são linhas quase retas, com uma inclinação de 0° , e as linhas de fluxo de cristas na região inferior não tem uma orientação constante. Portanto, pode-se supor que as cristas da região superior ao ponto de referência geralmente têm uma orientação quantizada de 0° . Em [81] é desenvolvida uma técnica de detecção do ponto de referência que utiliza esta característica ao considerar que o ponto de referência está localizado no elemento mais interno da parte superior de Θ cuja orientação quantizada seja 0° . Considerar unicamente esta informação gera uma perda de exatidão na estimação do ponto de referência (nem sempre o ponto de referência está em um elemento quantizado a 0°) e uma grande dificuldade em detectar o ponto de referência de uma impressão digital tipo A. Para superar estes problemas, propõe-se o análise de Θ em múltiplos níveis de representação segundo seus valores de quantização, tal que os blocos quantizados das orientações vizinhas de 0° sejam também utilizados, como se verá a seguir.

Quando Θ é quantizado em Q valores (ou seja, só é possível ter Q orientações espaçadas π/Q rad) ele pode representar-se como uma soma das matrizes $\Theta_0, \Theta_1, \dots, \Theta_{Q-1}$:



Figura 3.16: Distribuição das cristas sobre uma vizinhança do ponto de referência.

$$\Theta(i, j) = \sum_{k=0}^{Q-1} \frac{k\pi}{Q} \Theta_k(i, j), \quad (3.30)$$

tal que:

$$\Theta_k(i, j) = \begin{cases} 1 & k = \lfloor \Theta(i, j) \frac{Q}{\pi} \rfloor \\ 0 & \text{caso contrario} \end{cases}, \quad (3.31)$$

onde: $\lfloor \bullet \rfloor$ representa o operador de truncamento pelo valor mais baixo.

Cada Θ_k é uma imagem binária onde seus elementos iguais a 1 indicam a posição daqueles elementos em Θ que tem a orientação $k\pi/Q$, portanto pode-se dizer que Θ_k é a k -ésima representação (segundo seu valor de quantização) da imagem de orientação. Uma análise em múltiplos níveis de representação implicará a mistura das matrizes Θ_k (em algum nível de análise) de modo que alguma característica de nosso interesse de Θ seja ressaltada. Uma mistura que permite ressaltar as características de Θ vinculadas ao ponto de referência é a relação definida por:

$$\Theta_d^B(i, j) = \Theta_m(i, j) + \Theta_n(i, j), \quad (3.32)$$

tal que: $m = \text{mod}(-d, Q)$ e $n = \text{mod}(Q-1-d, Q)$, onde, d é um parâmetro de deslocamento que define o nível de análise considerado. Pode-se ver que Θ_d^B é uma imagem binária (ver Figura 3.17), onde unicamente os blocos com orientações $m\pi/Q$ e $n\pi/Q$ são 1, posto que tais orientações definem as representações da imagem de orientação usadas na formação da mistura binária Θ_d^B definida pela relação (3.32).

Para as impressões digitais tipo TA, RL, LL, W e TL o ponto de referência está situado na esquina mais interna do objeto binário localizado na parte

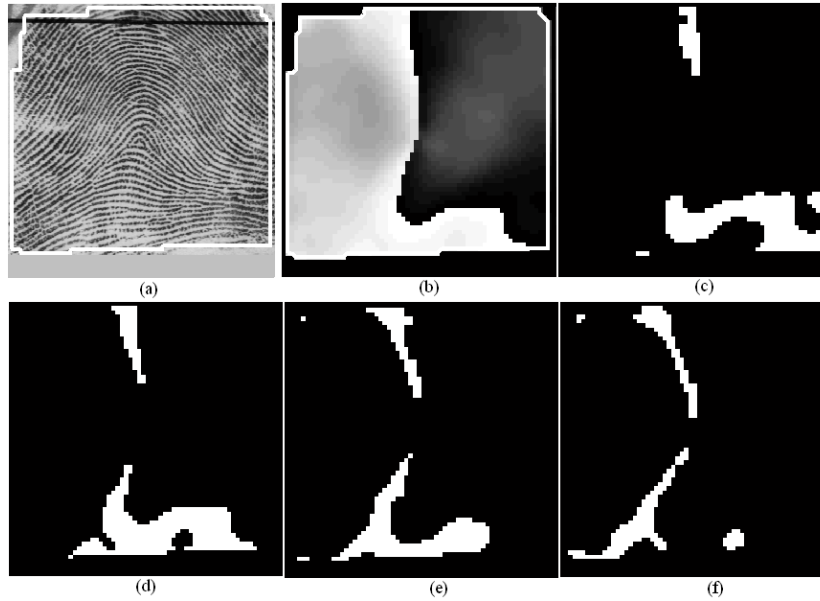


Figura 3.17: (a) I ; (b) Θ quantizado em 32 níveis; (c) Θ_0^B ; (d) Θ_1^B ; (e) Θ_2^B ; (f) Θ_3^B .

superior de Θ_d^B (ver Figura 3.18.b - 3.18.f) e para a impressão digital tipo A o ponto de referência está situado na parte mais fina do maior objeto binário de Θ_d^B (ver Figura 3.18.a)

Então, à medida em que obtemos os Θ_d^B , determina-se os possíveis pontos de referência, considerando como ponto de referência final ao mais interno deles. Uma variação em escala de cinza desta proposta [1] é definida pela relação:

$$\Theta_d^C(i, j) = \sum_{r=0}^{Q-1} \sum_{k=0}^{Q-1} w_d(k, r) \Theta_{\text{mod}(k-d, Q)}(i, j), \quad (3.33)$$

$$w_d(k, r) = \exp\left(-\frac{1}{2\sigma^2}(k - \text{mod}(r - d, Q))^2\right). \quad (3.34)$$

Onde, w_d permite ponderar a mistura dos diferentes Θ_k , tal que permita que os blocos com orientações $m\pi/Q$ e $n\pi/Q$ tenham os valores mínimos, esta representação em escala de cinza permite utilização do operador morfológico de abertura para eliminar as pequenas perturbações presentes em Θ_d^C , mantendo as características de importância que são as esquinas presentes no contornos do componente conectado dos blocos com orientações $m\pi/Q$ e $n\pi/Q$ (ver Figura 3.19).

Em função das representações em múltiplos níveis Θ_d^B e Θ_d^C , desenvolveu-se uma técnica original que permite a detecção do ponto de referência de

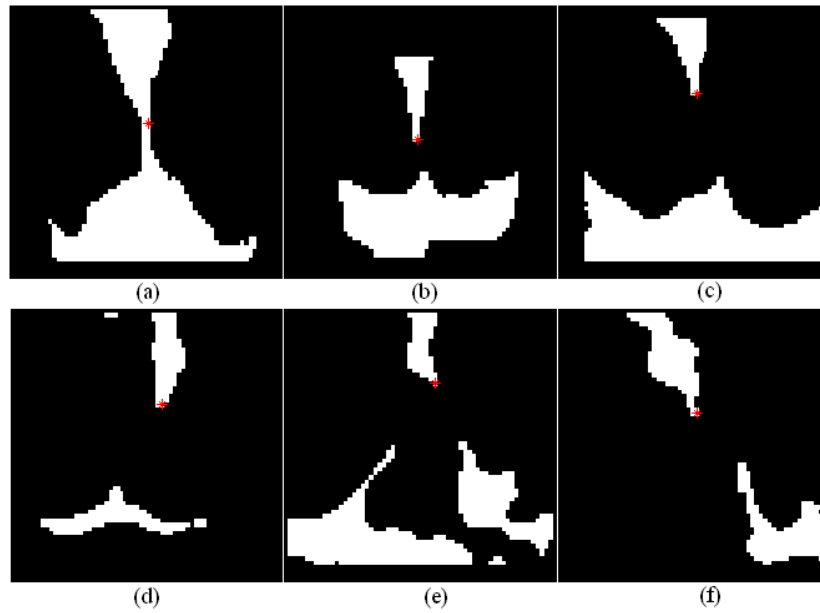


Figura 3.18: Imagens Θ_d^B características para as impressões digitais tipo (a)A; (b)TA; (b)RL; (c)LL; (d)W; (e)TL, quando $d = 1$ e $Q = 16$.

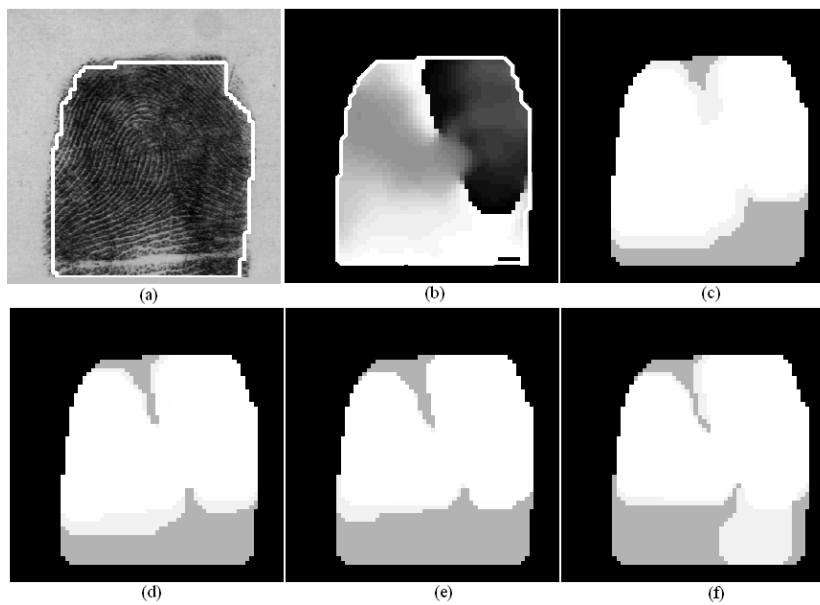


Figura 3.19: Imagens Θ_d^C para (a) $d = 0$; (b) $d = 1$; (c) $d = 2$; (d) $d = 3$;

uma impressão digital sem importar o tipo desta. Os passos são:

1. Quantizamos as orientações de Θ em Q valores angulares discretos

$$\{0, \pi/Q, \dots, (Q-1)\pi/Q\}.$$

2. Determinamos um índice de qualidade da impressão digital através da matriz de coerência Coh e a imagem de regiões Π . Primeiro, determina-se o número de elementos do primeiro plano de Π , ao qual denotamos por N_{Pri} , logo determinamos o número de elementos de Coh pertencentes ao primeiro plano que são inferiores a um limiar t_Q , ao qual denotamos por N_{Coh} . Definimos como um indicador de Qualidade $Q_c(t_Q) = N_{Coh}/N_{Pri}$, portanto, uma impressão digital será de boa qualidade se $Q_c(t_Q)$ está perto de 1, tal que, o limiar t_Q define o grau de sensibilidade de $Q_c(t_Q)$ (à medida que t_Q aproxima-se a 1 só as regiões com uma orientação muito bem definida são consideradas de boa qualidade).
3. para $d = 0, 1, \dots, Q/8 - 1$, realizar:
 - Se $Q_c(t_Q)$ é menor que 0.8, então calcula-se Θ_d^B , caso contrário calcula-se Θ_d^C . No caso que Θ_d^C seja calculado, aplica-se um operador morfológico de abertura, e limiariza-se a imagem resultante para unicamente ter os componentes conectados dos blocos com orientações $m\pi/Q$ e $n\pi/Q$, posto que a imagem resultante é binária ela também será denotada como Θ_d^B .
 - Multiplica-se elemento a elemento Θ_d^B por Π para eliminar os elementos de Θ vinculados ao segundo plano.
 - Determina-se a matriz B_m de pontos médios de Θ_d^B . Para cada linha de Θ_d^B é determinado o ponto médio de cada conjunto de 1 presentes na linha em questão. Tal que: $B_m(i, j) = 1$ se (i, j) é a posição de um ponto médio já detectado. (ver Figura 3.20)
 - Determina-se o vetor de pontos de finalização P_e da matriz B_m . Aplica-se a B_m a definição de Rutovitz do número de cruzamentos [47], o qual permite determinar as posições dos pontos de finalização numa imagem binária afinada(B_m é uma imagem afinada), onde seus pontos de finalização são as esquinas dos objetos binários presentes em Θ_d^B . Tais posições são armazenadas no vetor de pontos de finalização $P_e = \{(i_{p1}, j_{p1}), (i_{p2}, j_{p2}), \dots\}$.
 - Realiza-se uma filtragem em P_e para eliminar todas aquelas posições não vinculadas a uma esquina válida em Θ_d^B . Primeiro, eliminam-se os pontos que estejam perto do segundo plano; segundo, eliminam-se todos aqueles pontos cujo conjunto de 1 ao que pertençam tenha uma longitude superior a 2; terceiro, eliminam-se todos

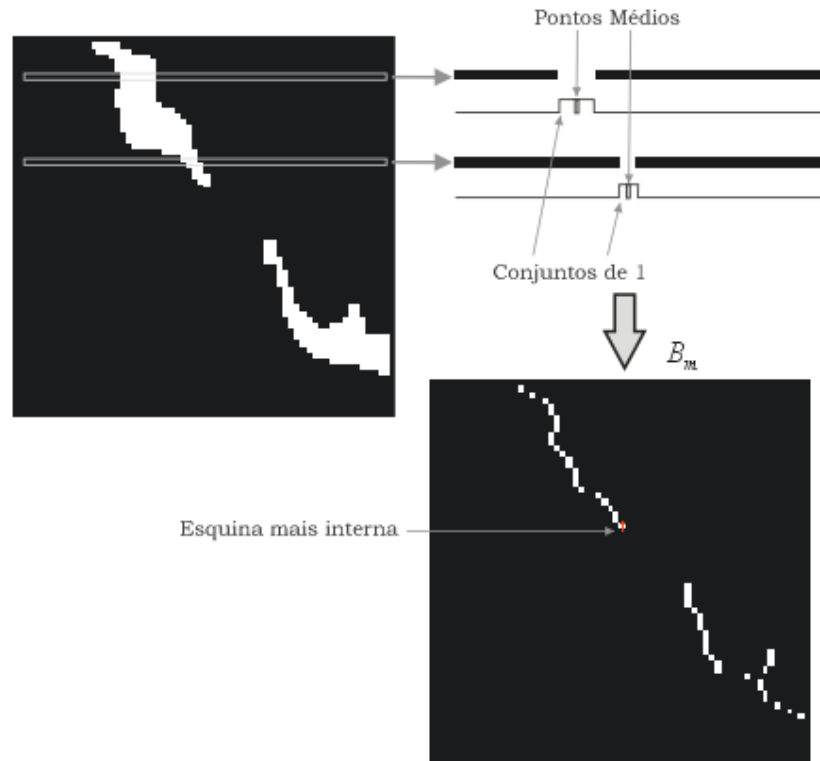


Figura 3.20: Determinação da matriz binária de pontos médios B_m e as longitudes dos conjuntos de 1 de Θ_d^B ao realizar um percorrido de suas linhas.

aqueles pontos que encontram-se na parte inferior de Θ_d^B (todo $i_p \geq N_{fil}/1.6$ é eliminado); quarto, extrai-se uma vizinhança 3×3 de Θ_d^B centrada em cada ponto (i_p, j_p) já validado e se a soma de elementos da linha inferior é diferente de 0, ou se a soma de elementos nas diagonais é superior a 2, então o ponto é eliminado.

4. Define-se a imagem binária B_T , tal que ela é a união de todas as Θ_d^B , ou seja, $B_T = \Theta_0^B \cup \Theta_1^B \cup \dots \cup \Theta_{Q/8-1}^B$, elimina-se de B_T as componentes conectadas isoladas, o qual é realizado através do seguinte procedimento: calcula-se a matriz binária $B_{Tp} = (B_T \cap \Pi) \cup (\bar{B}_T \cap \bar{\Pi})$, onde $\bar{\bullet}$ é o operador de negação, em seguida determina-se a área das componentes conectadas de B_{Tp} e eliminam-se as componentes conectadas em B_T cuja área seja menor que o maior componente conectado em B_{Tp} .
5. Para cada ponto válido de P_e , extrai-se sua linha correspondente em B_T , e determina-se a longitude do conjunto de 1 ao que pertence, se essa longitude é maior a um limiar (6 no nosso caso) então o ponto é eliminado de P_e . Também, o ponto é eliminado se não existe um conjunto de 1

(posto que o ponto em questão pode ter pertencido a uma ilha de B_T já eliminada).

6. Se o número de pontos válidos em P_e é maior que 0 e o número de componentes conectados de B_T é maior ou igual a 1, então, agrupamos os pontos válidos de P_e que pertençam a um mesmo componente conectado de B_T e consideramos como ponto representativo de tal componente conectado ao mais inferior deles (suponha-se que o ponto de referência está no elemento mais interno da componente conectada em questão). Logo, o ponto de referência será o superior de todos os pontos representativos (suponha-se que o ponto de referência não pode estar na parte inferior da impressão digital).
7. Se o número de pontos válidos em P_e é igual a 0 e o número de componentes conectados de B_T é maior a 1, define-se como ponto de referência o centróide de Π .
8. Se o número de pontos válidos em P_e é igual a 0 e o número de componentes conectados de B_T é igual a 1, pode-se supor que estamos frente a uma impressão digital tipo A, portanto assume-se como ponto de referência o centróide do único componente conectado de B_T , entretanto é necessário processar B_T para obter uma estimativa adequada do centróide. Então, primeiro, percorrem-se todas as linhas de B_T e eliminamos os conjuntos de 1 que estejam perto do segundo plano, para os conjuntos de 1 que estão no centro de B_T , calcula-se o ponto médio e a longitude de seu conjunto de 1; segundo, eliminam-se as ilhas de B_T com o mesmo procedimento do passo (4); terceiro, calcula-se o centróide do único componente conectado de B_T ; quarto, determina-se que pontos médios calculados no primeiro passo estão perto de centróide e define-se como ponto de referência o ponto médio que tenha a menor longitude de 1.
9. Se o ponto de referência determinado em (6) e (8) está perto do segundo plano, então é considerado como ponto de referência o centróide de Π , caso contrário, unicamente o ponto de referência determinado em (6) é recalculado numa vizinhança da matriz de coerência, Coh , para assim melhorar a exatidão no cálculo da posição. Primeiro extraímos uma vizinhança de Coh centrada no ponto de referência, de tamanho 11x11 elementos e nela definimos como ponto de referência o ponto da vizinhança que tenha o menor valor de coerência.

Na Figura 3.21 tem-se a seqüência de passos desta técnica, iniciando-se com a quantização de Θ em 32 valores (Figura 3.21.a), após o cálculo das

imagens Θ_d^B onde os pontos vermelhos indicam as esquinas mais internas do objeto binário superior de cada Θ_d^B e as rotas médias são indicadas pelas linhas azuis (figuras 3.21.b-e); na Figura 3.21.f tem-se a imagem binária total B_T onde os pontos vermelhos indicam todas as esquinas validas, considerando como possível ponto de referência ao mais inferior; na Figura 3.21.g temos o recálculo do ponto de referência através da matriz Coh e na Figura 3.21.h temos o ponto de referência final indicado em cor azul.

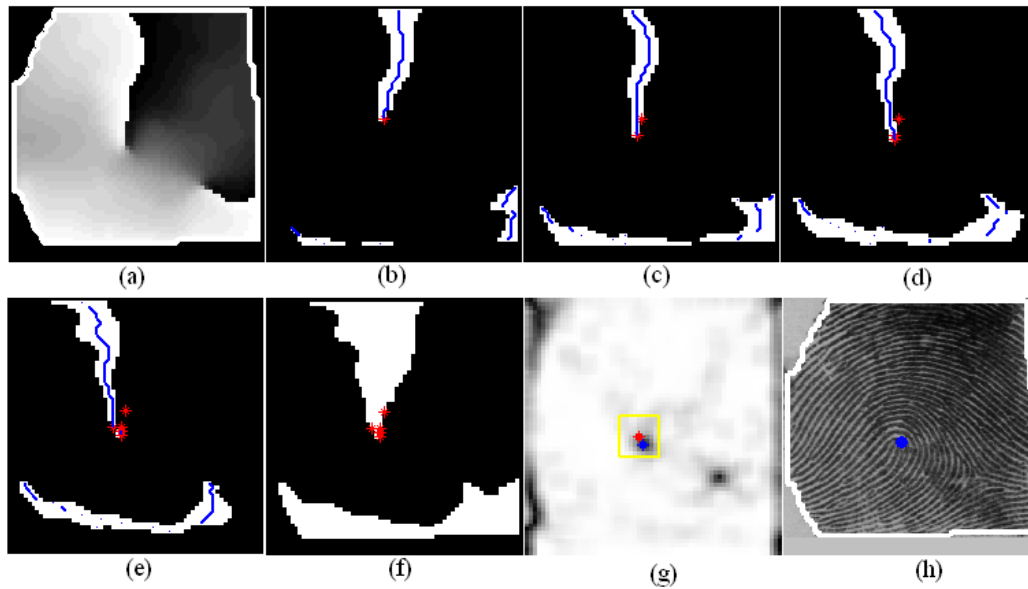


Figura 3.21: (a) Θ quantizado em 32 níveis; (b) Θ_0^B ; (c) Θ_1^B ; (d) Θ_2^B ; (e) Θ_3^B ; (f) B_T ; (g) recálculo do ponto de referência utilizando a matriz de Coh ; (h) ponto de referência final.

3.6 Resumo

Neste capítulo foram expostas as técnicas utilizadas para determinar as representações e características de importância de uma impressão digital que em conjunto se definiu como a etapa de processamento (ver Figura 3.22), como resultado desta etapa obteve-se a imagem de regiões que indica que parte de uma impressão digital é área válida ou fundo, a imagem de orientação que representa a informação direcional da impressão digital e a determinação do ponto de referência que permite realizar um registro da impressão digital dando a todas as impressões um ponto referencial comum.

No seguinte capítulo é vista a etapa de classificação, que realiza a tarefa da determinação da classe de uma impressão digital, considerando a informação obtida nesta etapa.

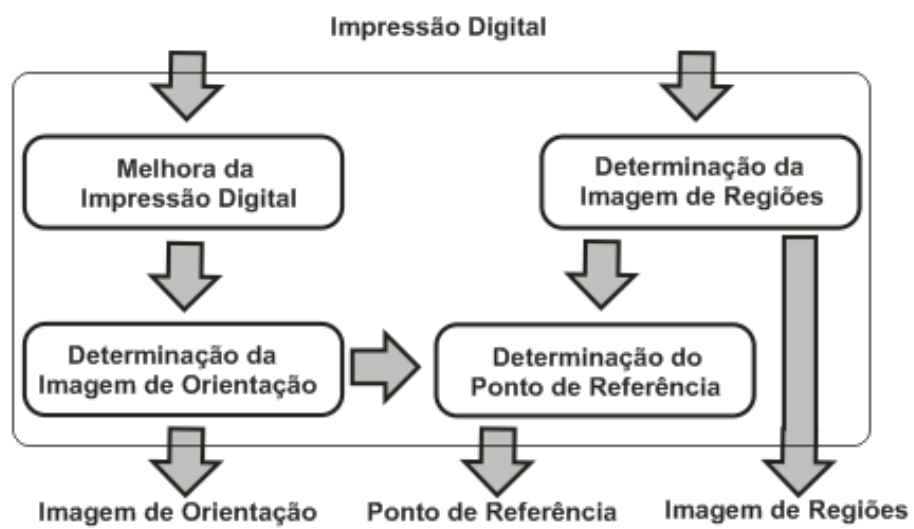


Figura 3.22: Arquitetura da etapa de processamento.

Capítulo 4

Classificação

4.1 Introdução

Considerando o PCASYS como referência, são utilizados os seguintes classificadores:

- Bayesiano com fdp normal¹,
- baseado em k -vizinhos mais próximos e
- baseados em redes neurais (o perceptron multi-camadas e uma rede neural probabilística).

A seleção dos classificadores neurais foi feita pois no PCASYS apresentaram as menores taxas de erro de classificação [10]. A seleção dos classificadores Bayesiano com fdp normal e k -vizinhos mais próximos foi feita com o propósito de avaliar a etapa de processamento da impressão digital explicada no capítulo anterior, posto que por serem classificadores simples o fato de gerar medidas de rendimento aceitáveis é fruto do bom desempenho da parte de processamento (ou seja, uma correta determinação da imagem de orientação, da imagem de regiões e do ponto de referência ajudará ao classificador a realizar melhor sua tarefa). Além disso, o PCASYS também foi testado com ditos classificadores, porém suas taxas de erro foram menores que os obtidos pelos classificadores neurais.

O Capítulo divide-se em duas seções, na primeira seção detalha-se as partes do classificador estatístico implementado, primeiro, é explicado a geração das características tomando como parâmetro de entrada a imagem de orientação. Segundo, é descrita a técnica de redução dimensional utilizada,

¹função de densidade de probabilidade normal multivariada

cujas saídas são os vetores de características a classificar e, finalmente, são descritos brevemente os classificadores estatísticos utilizados. Na última seção é apresentado um resumo do capítulo.

4.2 Classificação Estatística

Esta técnica utiliza uma representação da imagem de orientação para a geração do vetor de características e um classificador estatístico padrão para a determinação das classes (Normal, vizinhos mais próximos e baseado em redes neurais). Este tipo de abordagem permite utilizar técnicas consolidadas na literatura para a seleção das características e para a classificação, porém tem como inconveniente que os resultados do rendimento do classificador dependem do volume de amostras de treino que se tem. Como já se comentou, este trabalho segue a arquitetura do PCASYS, a qual realiza a classificação através dos seguintes passos:

Geração de Características A partir de Θ gera-se uma representação de maior dimensão que elimina as descontinuidades dos valores angulares, gerando-se o vetor de dados \mathbf{z} .

Redução dimensional Como o vetor de dados \mathbf{z} é de alta dimensão é mais que necessário sua redução dimensional. Para tal propósito utiliza-se a técnica de análise da componente principal (PCA), o qual gera o vetor de dados de menor dimensão, definido como o vetor de características \mathbf{x} .

Classificação O classificador é organizado como um banco de funções discriminativas D_i , tal que cada uma delas representa uma classe. O vetor de característica \mathbf{x} é classificado segundo a regra da máxima função discriminante [25].

Agora vejamos a seguir cada um destes passos.

4.2.1 Geração de Características

A proposta padrão usada em [102], [103], [17] e [90] é utilizar a representação vetorial de Θ , ou seja:

$$v(i, j) = \beta(i, j) \begin{bmatrix} \cos(\Theta_{reg}(i, j)) & \sin(\Theta_{reg}(i, j)) \end{bmatrix}, \quad (4.1)$$

onde, $\Theta_{reg}(i, j)$ é a imagem de orientação registrada e β define um valor de confiança. Este tipo de codificação permite que o problema induzido pela

descontinuidade da orientação ($180^\circ \leftrightarrow 0^\circ$) seja evitado. o valor de confiança utilizado neste trabalho é:

$$\beta(i, j) = 1, \quad (4.2)$$

tal que, todos os elementos de Θ tenham a mesma importância. O vetor de dados \mathbf{z} é definido como a representação em vetor coluna de v . Na Figura 4.1 pode-se ver a seqüência para a obtenção de \mathbf{z} , primeiro Θ e registrada segundo o ponto de referência, gerando-se Θ_{reg} (ver 4.1.b), segundo, a representação vetorial de Θ_{reg} é calculada, obtendo-se v (ver 4.1.c) e finalmente é obtida a representação coluna de v , obtendo-se \mathbf{z} (ver 4.1.d).

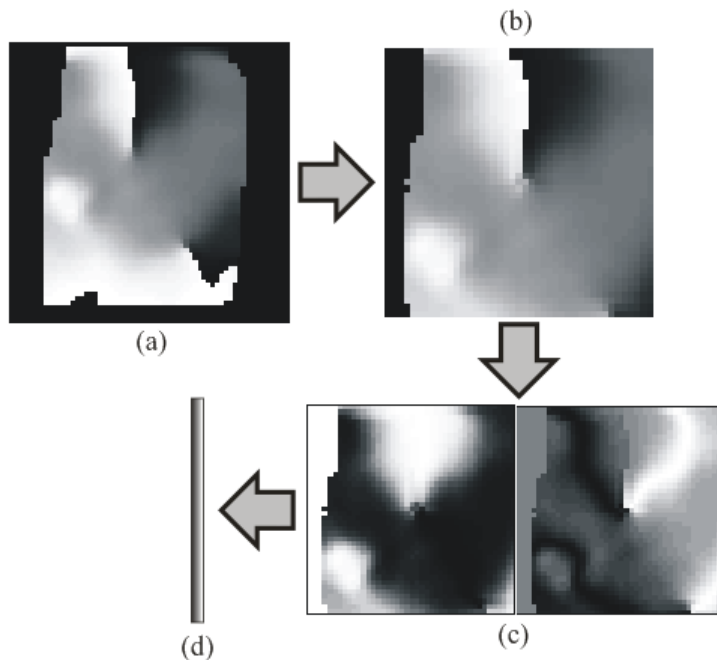


Figura 4.1: Representação Vetorial de Θ , (a) Θ ; (b) Θ_{reg} ; (c) v ; (d) \mathbf{z} .

4.2.2 Redução dimensional

Se temos um Θ de dimensão 64×64 elementos e utilizamos para a classificação a representação vetorial de Θ geraremos um vetor de dados de 8192 elementos ($=64 \times 64 \times 2$), o qual é um número muito alto de variáveis, dificultando o trabalho do classificador. Portanto, é mais que necessária a utilização de técnicas de redução dimensional. Uma das mais importantes e utilizada pelo PCASYS é a técnica PCA.

Análise da Componentes Principal (PCA)

Suponha que se tem um conjunto de observações conformado pelos vetores de dados de Θ : $\mathbf{z}_1, \dots, \mathbf{z}_p$ de m variáveis, e desejamos reduzir sua dimensão tal que cada observação possa ser descrita por um vetor \mathbf{x} de n variáveis, tal que, $n < m$. O problema pode ser definido como a obtenção de uma matriz de transformação linear \mathbf{W} que permite projetar todo vetor do espaço de dados \mathbf{z} como um vetor \mathbf{x} do espaço reduzido segundo a equação abaixo

$$\begin{aligned}\mathbf{x} &= \mathbf{W}^T \mathbf{z}, \\ E[\mathbf{z}] &= 0.\end{aligned}\tag{4.3}$$

Onde, $E[\bullet]$ é o operador de valor médio esperado.

A técnica PCA [53] obtém tal transformação ao utilizar um critério de otimização sobre a matriz de covariância do conjunto reduzido. A matriz de covariância amostral correspondente ao conjunto de observações e o conjunto de observações reduzidos é dado por:

$$\Sigma_z = \frac{1}{p-1} \sum_{j=1}^p \mathbf{z}_j \mathbf{z}_j^T,\tag{4.4}$$

$$\Sigma_x = \frac{1}{p-1} \sum_{j=1}^p \mathbf{x}_j \mathbf{x}_j^T,\tag{4.5}$$

onde, p é o número de observações. Sem perda de generalidade assume-se que $E[\mathbf{z}_j] = 0$ e $E[\mathbf{x}_j] = 0$. Σ_z quantifica as correlações² entre todos os possíveis pares de variáveis do conjunto de dados, pares de variáveis com uma grande covariância é um indicador que eles estão altamente correlacionados e pares de variáveis com uma baixa covariância é um indicador que eles estão pouco correlacionados. Como o objetivo é reduzir a dimensão dos dados, cada variável de nosso espaço reduzido tem que ter uma mínima covariância com respeito às outras variáveis. Com maior precisão a redução dimensional implicará determinar uma transformação \mathbf{W} , tal que, diagonalize Σ_x . Para a técnica PCA utiliza-se a solução da equação de autovalores de Σ_z para obter a diagonalização de Σ_x , segundo se observa na expressão abaixo.

$$\Sigma_z \mathbf{v} = \lambda \mathbf{v}.\tag{4.6}$$

Onde $\lambda \geq 0$ e $\mathbf{v} \in R^m - \{0\}$ são os autovalores e autovetores de Σ_z . A representação matricial de (4.6), definida como a decomposição espectral de Σ_z , é:

²No sentido estrito da palavra

$$\mathbf{D} = \mathbf{V}^T \Sigma_z \mathbf{V}, \quad (4.7)$$

onde as colunas de \mathbf{V} são os autovetores \mathbf{v} de Σ_z e a diagonal de \mathbf{D} são os autovalores λ de Σ_z . Agora substitui-se a (4.3) em (4.5):

$$\Sigma_x = \frac{1}{p-1} \mathbf{W}^T \Sigma_z \mathbf{W}. \quad (4.8)$$

Comparando (4.7) e (4.8), pode-se concluir que Σ_x será diagonalizável se \mathbf{W} é igual a matriz de autovetores \mathbf{V} .

A seguir detalhamos a descrição do algoritmo de PCA usando o técnica da covariância.

1. Concatenamos as observações $\mathbf{z}_1, \dots, \mathbf{z}_p$ gerando a matriz de observações $\mathbf{Z} \in R^{m \times p}$; $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_p]$ (ver Figura 4.2.a).
2. Calculamos o valor médio do \mathbf{Z} , obtendo-se a matriz $\mathbf{u} \in R^{m \times 1}$; $\mathbf{u} = \frac{1}{p} \mathbf{Z} \mathbf{1}_{p \times 1}$, onde $\mathbf{1}_{p \times 1} \in R^{p \times 1}$ e todos seus elementos são 1 (ver Figura 4.2.b).
3. Subtraímos \mathbf{u} de cada coluna de \mathbf{Z} , gerando-se a matriz $\mathbf{B} \in R^{m \times p}$; $\mathbf{B} = \mathbf{Z} - \mathbf{u} \mathbf{1}_{1 \times p}$, onde $\mathbf{1}_{1 \times p} \in R^{1 \times p}$ e todos seus elementos são 1.
4. Determina-se a matriz de covariância amostral $\Sigma_z \in R^{m \times m}$; $\Sigma_z = \frac{1}{p-1} \mathbf{B} \mathbf{B}^T$ e logo calculamos seus autovalores e autovetores; $\frac{1}{p-1} (\mathbf{B} \mathbf{B}^T) \mathbf{V} = \mathbf{V} \mathbf{D}$. A matriz $\mathbf{V} \in R^{m \times m}$, contém os autovetores $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ de Σ_z e a matriz $\mathbf{D} \in R^{m \times m}$, contém na sua diagonal os m autovalores $\lambda_1, \lambda_2, \dots, \lambda_m$ de Σ_z .
5. Reordenamos as colunas de \mathbf{V} conforme a ordem decrescente dos autovalores de Σ_z .
6. Separamos as primeiras n colunas de \mathbf{V} definindo a matriz de transformação \mathbf{W} como: $\mathbf{W} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n]$ (ver Figura 4.2.c).
7. Projetamos as observações no espaço reduzido definido pela matriz de transformação \mathbf{W} ; $\mathbf{x} = \mathbf{W}^T (\mathbf{z} - \mathbf{u})$ (ver Figura 4.2.d).

Se $m > p$ (número de variáveis é maior que o número de observações) acostuma-se realizar uma variação no passo (4), Neste caso, no lugar de calcular-se os autovalores e autovetores de $\mathbf{B} \mathbf{B}^T \in R^{m \times m}$, calcula-se os autovalores e autovetores da matriz $\mathbf{B}^T \mathbf{B} \in R^{p \times p}$, ou seja:

$$\frac{1}{p-1}\mathbf{B}^T\mathbf{B}\mathbf{V} = \mathbf{V}\mathbf{D}, \quad (4.9)$$

multiplicando-se (4.9) por \mathbf{B} , obtemos:

$$\frac{1}{p-1}(\mathbf{B}\mathbf{B}^T)(\mathbf{B}\mathbf{V}) = (\mathbf{B}\mathbf{V})\mathbf{D}. \quad (4.10)$$

Portanto, $\mathbf{B}\mathbf{V}$ e \mathbf{D} são os p autovetores e autovalores de $\mathbf{B}\mathbf{B}^T$, os outros autovalores $m-p$ são zero e seus correspondentes autovetores indefinidos. A matriz $\mathbf{B}\mathbf{V}$ representa os vetores ortogonais mas não ortonormais, portanto, é necessário dividir cada k -ésima coluna de $\mathbf{B}\mathbf{V}$ por $\sqrt{p\lambda_k}$.

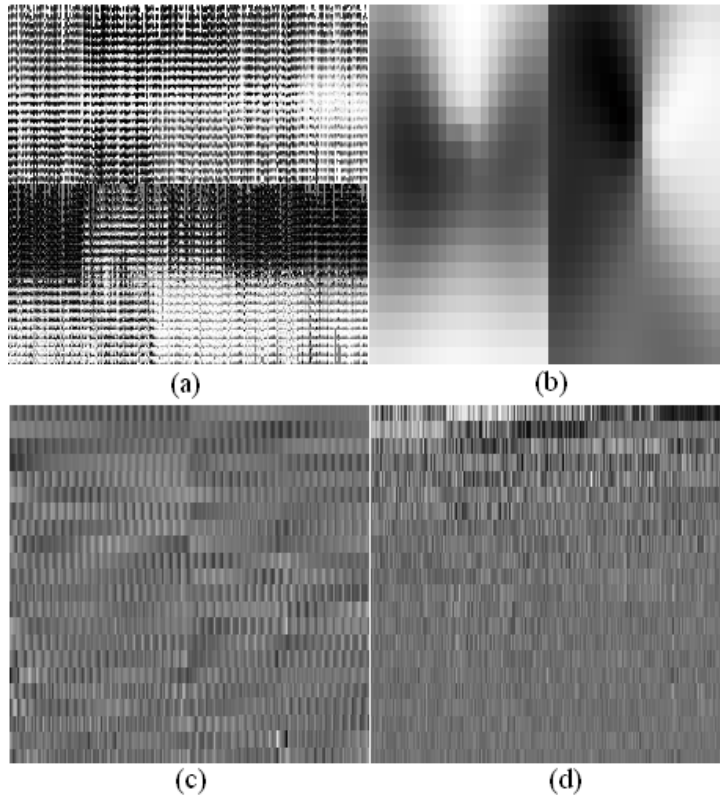


Figura 4.2: (a)Matriz de observações \mathbf{Z} ; (b)valor médio de \mathbf{Z} , \mathbf{u} ; (c)observações projetadas \mathbf{x} ; (d) matriz de transformação \mathbf{W} .

4.2.3 Classificação

Neste trabalho são utilizados quatro classificadores, um classificador Bayesiano com fdp normal, k -vizinhos mais próximos (k -NN) e dois classificadores neu-

rais o perceptron multi-camadas (MLP - *Multi Layer Perceptron*) e uma rede neural Probabilística (PNN - *Probabilistic Neural Network*), todos os classificadores tiveram como entradas as saídas do redutor dimensional baseado em PCA, ou seja o vetor de características \mathbf{x} . Cada classificador é representado como um banco de funções discriminantes em que o critério de classificação está baseado no valor do máximo discriminante. A seguir é realizada uma explicação dos classificadores utilizados, empregando-se a seguinte notação:

- L : número de classes; $L = 4, 5$ (no problema de 4 classes ($L = 4$) as impressões digitais são classificadas nos tipos RL, W, LL, TA-A, no problema de 5 classes ($L = 5$) as impressões digitais são classificadas nos tipos RL, W, LL, TA e A),
- Ω_i : rótulo da i -ésima classe,
- $P(\Omega_i)$: probabilidade *a priori* da classe Ω_i ,
- $\hat{P}(\Omega_i)$: estimação de $P(\Omega_i)$,
- n : dimensão do vetor de característica (dimensão do espaço reduzido),
- \mathbf{x} ; vetor de características (saída do PCA),
- $\mathbf{x}_j^{(i)}$: i -ésimo vetor de características pertencente ao j -ésima classe,
- \mathbf{u}_i : vetor média da classe Ω_i ,
- $\hat{\mathbf{u}}_i$: estimativa de \mathbf{u}_i ,
- Σ_i : matriz de covariância da classe Ω_i ,
- $\hat{\Sigma}_i$: estimativa de Σ_i ,
- $d(\mathbf{x}, \mathbf{y}) = (\mathbf{x} - \mathbf{y})^T(\mathbf{x} - \mathbf{y})$: distância Euclidiana entre \mathbf{x} e \mathbf{y} ; $\mathbf{x}, \mathbf{y} \in R^n$,
- $D_i(\mathbf{x})$: i -ésima função discriminante; $1 \leq i \leq L$, $\mathbf{x} \in R^n$.

Classificador Bayesiano com fdp normal (classificador normal)

Este classificador está baseado na aplicação da regra de Bayes supondo uma função de densidade normal multivariável para cada classe Ω_i . Primeiro, será útil mencionar alguns conceitos vinculados aos classificadores paramétricos, para os quais usaremos a seguinte terminologia:

- $p(\mathbf{x})$: função de densidade de probabilidade para $\mathbf{x} \in S \subseteq R^n$, onde, $\int_S p(\mathbf{x}) d\mathbf{x} = P(\mathbf{x} \in S)$,

- $p(\mathbf{x}|\Omega_i)$: função de densidade condicional para $\mathbf{x} \in S \subseteq R^n$ dado que a classe verdadeira é Ω_i , onde, $\int_S p(\mathbf{x}|\Omega_i) d\mathbf{x} = P(\mathbf{x} \in S|\Omega_i)$,
- $P(\Omega_i|\mathbf{x})$: probabilidade *a posteriori* para um \mathbf{x} particular, $P(\Omega_i|\mathbf{x}) = P(\mathbf{x} \text{ pertence a classe } \Omega_i)$.

O classificador Bayesiano ótimo é aquele que minimiza as perdas esperadas, isto significa que uma correta classificação não produz perdas e uma incorreta classificação produz perdas iguais a 1. Neste caso, o classificador Bayesiano classifica cada \mathbf{x} como pertencente à classe Ω_i para o qual a probabilidade *a posteriori* $P(\Omega_i|\mathbf{x})$ é a maior. De acordo com a regra de Bayes:

$$P(\Omega_i|\mathbf{x}) = \frac{P(\Omega_i)p(\mathbf{x}|\Omega_i)}{p(\mathbf{x})}. \quad (4.11)$$

O fim é classificar \mathbf{x} como pertencente à classe Ω_i se sua probabilidade *a posteriori* $P(\Omega_i|\mathbf{x})$ é máxima. Para classificadores Normais cada classe Ω_i é assumida que tem uma função de densidade dada por:

$$p(\mathbf{x}|\Omega_i) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma_i|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{u}_i)^T \Sigma_i^{-1}(\mathbf{x} - \mathbf{u}_i)\right), \quad (4.12)$$

onde, $|\Sigma_i|$ é o determinante de Σ_i . Na prática \mathbf{u}_i e Σ_i são desconhecidos e usa-se suas estimativas, ou seja: $\hat{\mathbf{u}}_i$ e $\hat{\Sigma}_i$. Substituindo (4.12) em (4.11) com as estimações $\hat{\mathbf{u}}_i$ e $\hat{\Sigma}_i$ obtemos a estimativa para a probabilidade *a posteriori* $\hat{P}(\Omega_i|\mathbf{x})$. Define-se como função discriminante do classificador normal, $D_i(\mathbf{x})$, como o logaritmo natural de $\hat{P}(\Omega_i|\mathbf{x})$:

$$D_i(\mathbf{x}) = \ln(\hat{P}(\Omega_i)) - \frac{1}{2} \ln(|\hat{\Sigma}_i|) - \frac{1}{2}(\mathbf{x} - \hat{\mathbf{u}}_i)^T \hat{\Sigma}_i^{-1}(\mathbf{x} - \hat{\mathbf{u}}_i), \quad (4.13)$$

$p(\mathbf{x})$ é omitido posto que não afeta na escolha de Ω_i que maximiza $P(\Omega_i|\mathbf{x})$ e o termo $(2\pi)^{-\frac{n}{2}}$ não é considerado por ser constante. Determina-se a classe vinculada à observação \mathbf{x} segundo a regra da máxima função discriminante [25].

$$c = \arg \max_{i=1,2,\dots,L} \{D_i(\mathbf{x})\}, \quad (4.14)$$

onde, c é o índice da classe atribuída a \mathbf{x} é *arg max* devolve o índice da máxima função discriminante.

Classificador k -vizinhos mais próximos

Os conceitos básicos sobre esta técnica aparecem pela primeira vez no trabalho de Fix e Hodges [34]. Este primeiro estudo definiu o problema da discriminação não paramétrica e sua solução como uma extensão aos desenvolvimentos prévios da discriminação paramétrica. Neste caso, conhecidas as funções de densidade, a decisão sobre a pertinência de um ponto a uma de duas classes depende só da razão entre suas funções de densidade. Para o caso não paramétrico, Fix e Hodges propuseram a estimativa da função de densidade de uma classe, num ponto \mathbf{x} , como a proporção de amostras de tal classe entre as k amostras próximas a \mathbf{x} . Eles também demonstraram que esta técnica era assintoticamente ótima para grandes conjuntos de dados. Cover e Hart [26] analisaram formalmente o procedimento proposto em [34], denominando o procedimento como a regra do 1-vizinho mais próximo para $k = 1$ (abreviado como 1-NN, por suas conhecidas siglas anglo-saxônicas), e em geral regra dos k -vizinhos mais próximos (k -NN). Nestes trabalhos obtiveram-se cotas relativamente ajustadas para sua taxa de erro em termos da probabilidade de erro de Bayes. Formalmente, a taxa de erro do k -NN com respeito à taxa de erro de Bayes, considerando um número infinito de dados de treino, vem definido pela desigualdade:

$$P^* \leq P \leq P^* \left(2 - \frac{c}{c-1} P^* \right), \quad (4.15)$$

onde, P^* é a taxa de erro de Bayes e P é taxa de erro do k -NN, para o problema de classificação em c classes (ver Figura 4.3).

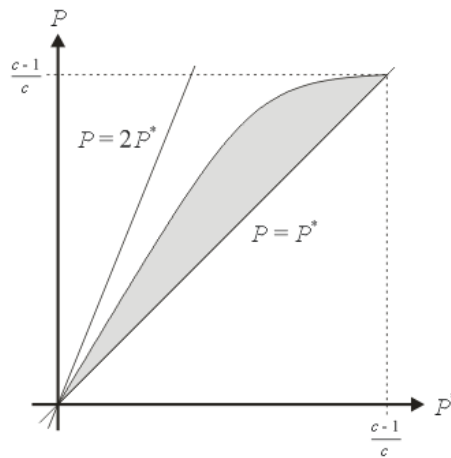


Figura 4.3: Limites da taxa de erro do k -NN num problema de c classes, dado um número infinito de dados de treino [31].

A regra de classificação do 1-vizinho mais próximo é explicado como segue, quando deseja-se classificar uma observação \mathbf{x} , que não pertence ao conjunto de treino, primeiro determina-se o vizinho mais próximo a \mathbf{x} dentre as amostras do conjunto de treino. Seja \mathbf{x}_{1NN} o vizinho mais próximo a \mathbf{x} , se a classe a que pertence \mathbf{x}_{1NN} é Ω_c , então \mathbf{x} é rotulado como da classe Ω_c .

Ao definir-se a função discriminante do 1-NN pela relação:

$$D_i(\mathbf{x}) = - \min_{j=1,2,\dots,L_i} \left\{ d^2 \left(\mathbf{x}, \mathbf{x}_j^{(i)} \right) \right\}, \quad (4.16)$$

a regra do 1-NN é definida pela relação (4.14). É evidente que ao considerar-se somente o vizinho mais próximo a \mathbf{x} não fazemos um uso eficiente de toda a informação contida no conjunto de treino. Uma extensão natural da regra 1-NN é a regra de classificação pelos k -vizinhos mais próximos, k -NN, que consiste em procurar os k -vizinhos mais próximos a \mathbf{x} no conjunto de treino e atribuir a pertinência de \mathbf{x} à classe com maior votação obtida nos k vizinhos selecionados. Ao definir-se a função discriminante do k -NN pela relação:

$$D_i(\mathbf{x}) = \text{card} \left\{ \mathbf{S}_x^{(i)} \right\}, \quad (4.17)$$

onde, $\mathbf{S}_x^{(i)}$ é um conjunto que contém todas as amostras de treino pertencentes à classe i que são parte dos k -vizinhos mais próximos a \mathbf{x} e card devolve o número de elementos de um conjunto. A regra do k -NN também é definida pela relação (4.14).

Perceptron Multi-camada

É uma classe de rede de propagação para frente, baseada no paradigma de aprendizagem supervisionada, o que implica que requer exemplos que representem o conhecimento sobre o problema e de um algoritmo de aprendizagem que corrija os parâmetros adaptáveis da rede, ou seja, os pesos. Esta rede é também conhecida como MLP (*Multilayer Perceptron*). Neste tipo de rede o fluxo de informação é unidirecional, quando se apresenta um padrão à rede (o vetor de características \mathbf{x} em nosso caso), este se propaga desde sua camada de entrada até sua camada de saída. Para explicar com detalhe este processo, é conveniente definir a seguinte notação, considerando um MLP com três camadas (contando a camada de entrada e a camada de saída):

- $N^{(i)}$: número de neurônios na camada i -ésima ($i = 0, 1, 2$), $N^{(0)} = n$ e $N^{(2)} = L$.
- $f(s)$: função de ativação não linear ($s \in R$).

- $b_i^{(k)}$: bias do i -ésimo neurônio da k -ésima camada ($k = 1, 2$).
- $w_{ij}^{(k)}$: peso do i -ésimo neurônio da k -ésima camada ao j -ésimo neurônio da $(k - 1)$ -ésima camada ($k = 1, 2$; $1 \leq i \leq N^{(k)}$; $1 \leq j \leq N^{(k-1)}$),
- $n_i^{(k)}(\mathbf{x})$: saída do i -ésimo neurônio da k -ésima camada ($k = 1, 2$; $1 \leq i \leq N^{(k)}$).

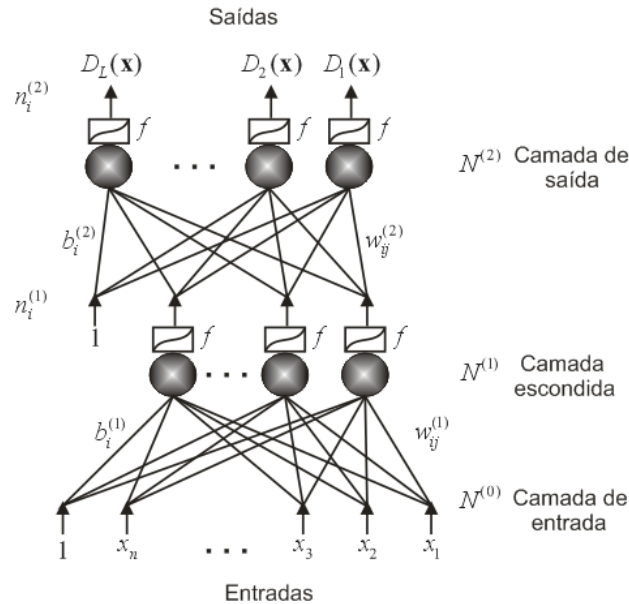


Figura 4.4: Topologia de um MLP de três camadas.

O processo de propagação a frente é explicado como segue:

- apresenta-se um vetor de características \mathbf{x} na camada de entrada da rede.
- Cada neurônio da camada escondida recebe as entradas da camada de entrada e calcula uma combinação linear deles, o resultado da combinação linear passa pela função de ativação f , gerando-se as saídas da camada escondida:

$$n_j^{(1)}(\mathbf{x}) = f \left(b_j^{(1)} + \sum_{k=1}^{N^{(0)}} w_{jk}^{(1)} x_k \right), \quad (4.18)$$

- o mesmo procedimento é aplicado desde as entradas da camada escondida até os neurônios da camada de saída, gerando-se as saídas da camada de saída:

$$n_i^{(2)}(\mathbf{x}) = f \left(b_i^{(2)} + \sum_{j=1}^{N^{(1)}} w_{ij}^{(2)} n_j^{(1)}(\mathbf{x}) \right). \quad (4.19)$$

Ao definir-se a função discriminante do MLP pela relação: $D_i(\mathbf{x}) = n_i^{(2)}(\mathbf{x})$, o critério de classificação vem estabelecido pela relação (4.14).

A função (4.19) constitui um instrumento de importância para realizar cálculos não lineares em muitas áreas de aplicação. Aqui é usada para a tarefa de classificação considerando que as classes podem ser não linearmente separáveis.

Para o treino dos pesos da rede, um procedimento razoável é usar um algoritmo de otimização para minimizar o erro quadrático médio sobre o conjunto de treino, entre os valores discriminantes produzidos pela propagação a frente da rede e os valores discriminantes objetivo, consistentes de seqüências de 1 e 0 definidos pelas classes atuais dos exemplos de treino.

O algoritmo de retropropagação do erro é a ferramenta fundamental para treinar o perceptron multi-camada. Existem vários tipos de algoritmos de retropropagação do erro, A idéia em geral é atualizar os pesos da rede na direção na qual a função do erro decresce mais rapidamente, ou seja a direção negativa do gradiente (algoritmo do gradiente descendente). Tem-se duas maneiras nas quais o algoritmo do gradiente descendente pode ser implementado: modo incremental (atualiza os pesos a cada apresentação de um novo padrão) e modo *batch* (somente ajusta os pesos após a apresentação de todos os padrões). Neste trabalho, foi utilizado o algoritmo do gradiente descendente, com momento, trabalhando no modo *batch* [42]. Foi usada a implementação de uma rede multi-camada (e os respectivo algoritmo de treino) do *toolbox* de Redes Neurais de MATLAB.

Rede Neural Probabilística

Também conhecida como PNN (*Probabilistic Neural Network*), foi proposta por Specht [92]. É interpretada como uma função que aproxima a densidade de probabilidade da subjacente distribuição das amostras do treino. A PNN consiste de neurônios localizados em três camadas depois das entradas, como é descrito abaixo:

camada padrão Para cada neurônio (i, j) -ésimo da camada padrão calcula-se a distância quadrática Euclidiana $d^2(\mathbf{x}, \mathbf{x}_i^{(j)})$ da observação de entrada \mathbf{x} ao vetor de treino j -ésimo pertencente à classe Ω_i , $\mathbf{x}_j^{(i)}$, para logo aplicar a função não linear:

$$f_{ij}(\mathbf{x}) = \exp\left(-\frac{1}{2\sigma^2}d^2(\mathbf{x}, \mathbf{x}_i^{(j)})\right), \quad (4.20)$$

onde σ é o parâmetro de suavização.

camada de adição Para cada i -ésimo neurônio da camada de adição calcula-se a densidade condicional de que \mathbf{x} pertença à classe Ω_i como a média das saídas de todos os neurônios da camada padrão que pertencem à classe Ω_i , como se observa em (4.21).

$$p(\mathbf{x}|\Omega_i) = \frac{1}{M_i} \sum_{j=1}^{M_i} f_{ij}(\mathbf{x}). \quad (4.21)$$

Portanto, definimos a i -ésima função discriminante do classificador PNN, como:

$$\begin{aligned} D_i(\mathbf{x}) &= \hat{P}(\Omega_i)p(\mathbf{x}|\Omega_i) \\ &= \frac{\hat{P}(\Omega_i)}{M_i} \sum_{j=1}^{M_i} f_{ij}(\mathbf{x}). \end{aligned} \quad (4.22)$$

camada de decisão Determina a classe vinculada à observação \mathbf{x} segundo (4.14).

$$c = \arg \max_{i=1,2,\dots,L} \{D_i(\mathbf{x})\}. \quad (4.23)$$

O único fator que necessita ser selecionado para o treino é σ (ver (4.20)), este é otimizado por tentativa e erro. Na literatura tem-se o desenvolvimento de arquiteturas mais sofisticadas para o PNN [35][68] porém, neste trabalho é utilizada a arquitetura original do PNN para poder ter resultados que sejam comparáveis aos obtidos pelo PCASYS.

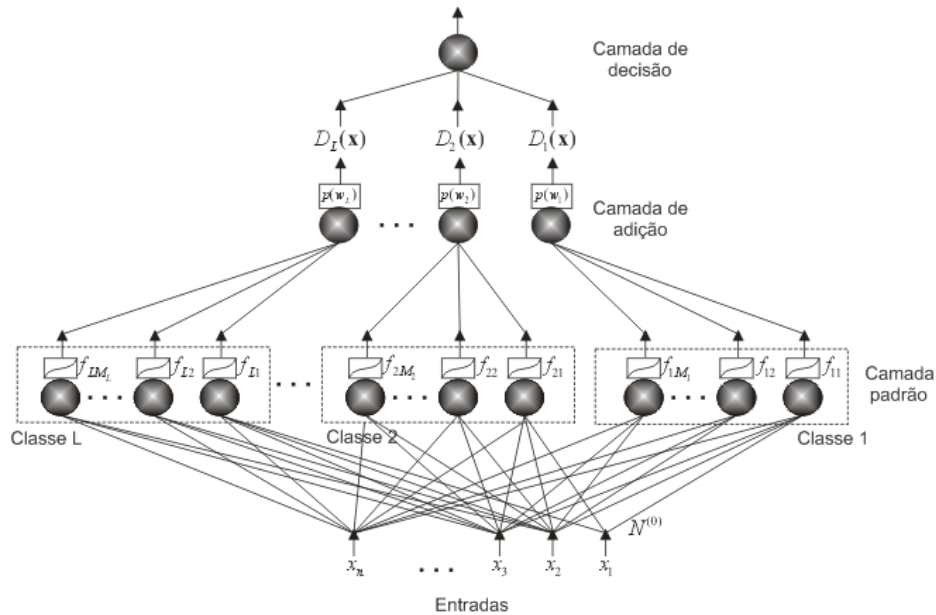


Figura 4.5: Topologia de um PNN.

4.3 Resumo

Neste capítulo foram apresentados os componentes do classificador implementado, explicando-se cada uma de suas partes constitutivas: geração de características, redução dimensional e classificação (como é indicado na Figura 4.6). O único por determinar é a classe suposta da impressão de entrada ao sistema, porém é necessário estabelecer algumas pautas: que banco de dados a usar, selecionar o melhor classificador em função de alguma medida de rendimento, o impacto da etapa de processamento entre outros. Tais pontos são pesquisado no capítulo final deste trabalho.

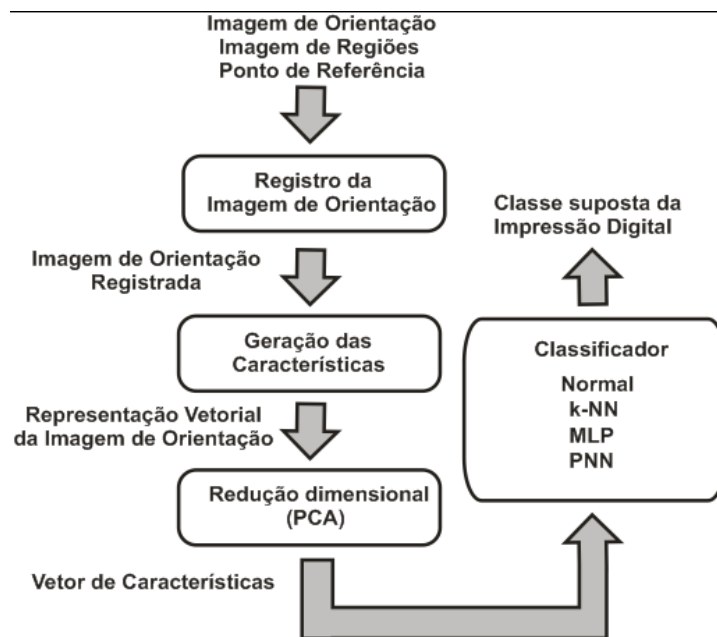


Figura 4.6: Arquitetura da etapa de classificação.

Parte IV
Parte Experimental

Capítulo 5

Testes e resultados

5.1 Introdução

Neste capítulo são apresentados os resultados experimentais obtidos para a avaliação do sistema de classificação implementado, usando um banco de dados especializado para testar sistemas de classificação de impressões digitais, de modo que, os resultados obtidos sejam comparáveis aos publicados por outros pesquisadores. O capítulo divide-se em seis seções. Nas quatro primeiras, realiza-se uma exposição de: o ambiente de desenvolvimento utilizado para a implementação computacional de cada uma das técnicas descritas nos capítulos anteriores, o banco de dados utilizado para testar o sistema, as medidas de rendimento que permitem estabelecer os margens de erro e exatidão do sistema de classificação e os resultados obtidos por outros pesquisadores utilizando o mesmo banco de dados com que o sistema foi testado. Na penúltima seção, são explicados os experimentos realizados e são mostradas as medidas de rendimento obtidas. Na última seção, é apresentado um resumo do capítulo.

5.2 O Ambiente Desenvolvido

Todo o sistema foi desenvolvido sobre o ambiente MATLAB, aproveitando-se as facilidades de implementação por ele proporcionadas. Este ambiente, permite diminuir o tempo de "prototipagem" de sistemas complexos. Hoje em dia, com a adição de diversas ferramentas, organizadas em caixas de trabalho denominadas *toolbox*, é possível realizar inclusive aplicações profissionais.

No que tange à maneira de escrever em MATLAB deve-se estar atento para evitar funções de execução lenta. Como tal linguagem é de propósito geral, é comum basear-se no estilo de programação de outros linguagens

como JAVA ou C o que acarreta numa diminuição de desempenho. Melhores resultados são obtidos programando-se na forma matricial evitando-se o uso de laços sobretudo para tais operações. Outra de suas vantagens refere-se à visualização de dados em geral, possível através de um conjunto de recursos gráficos de fácil utilização. O MATLAB também possui funções de leitura e escrita especiais para arquivos de imagens, permitindo a utilização de um banco de imagens. Por todas estas facilidades enumeradas acima, optou-se pelo desenvolvimento do sistema de classificação em tal ambiente.

5.3 Banco de Dados

Os experimentos foram feitos utilizando o banco de dados de impressões digitais número 4 (BD4) do NIST [99], tal banco é idôneo para o desenvolvimento e prova de um sistema de classificação de impressões digitais, tal que, a maioria de técnicas para classificação publicadas na última década foram provadas com este banco de dados.

A DB4 consiste de 4.000 imagens em níveis de cinza de 480x512 pixels, capturadas a 8 bits, com uma resolução de 500 dpi. Estas imagens são de impressões digitais roladas e escaneadas a partir de fichas papiloscópicas, cada impressão tem duas diferentes instâncias rotuladas com as letras F e S, indicando se corresponde à primeira (*First*) ou segunda (*Second*) instância de cada impressão digital. Além disso, cada uma delas foi analisada por um especialista forense que atribuiu um rótulo, definindo a classe a que pertence (A, LL, RL, TA ou W). Algumas impressões digitais ambíguas (perto de 17,5%) tem um rótulo adicional referente a uma classe secundária.

Na Tabela 5.1 mostra-se a distribuição das 4.000 impressões digitais na DB4, cada linha soma 800 de acordo com a classificação do primeiro rótulo de cada impressão, o que implica que a DB4 tem uma distribuição uniforme nas cinco classes. Posto que a predição da classe T é considerada a mais difícil e sua distribuição natural é relativamente pequena (2,9%) comparada com as distribuições das classes LL, RL e W (33,8%, 31,7% e 29,9%), a distribuição uniforme da DB4 faz com que o problema da classificação tenha uma maior complexidade. Também pode-se observar na Tabela 5.1 o número de imagens de impressões digitais que apresentam dois rótulos (elementos fora da diagonal).

Uma vez que as imagens da DB4 correspondem a impressões digitais roladas, uma parte importante deste banco de dados corresponde a impressões digitais de pobre qualidade, as quais são difíceis de classificar, até para um especialista forense. Para analisar o impacto da qualidade das impressões digitais na tarefa de classificação, cada impressão foi rotulada manualmente

	LL	W	RL	TA	A
LL	756	2	0	42	0
W	2	792	6	0	0
RL	0	4	746	50	0
TA	166	0	238	246	150
A	0	0	2	38	760

Tabela 5.1: Distribuição das impressões digitais na DB4.

como de pobre qualidade, média qualidade ou boa qualidade, em função de sua imagem de regiões e sua imagem de orientação, vejamos:

- Uma imagem é rotulada de pobre qualidade se ocorre uma super segmentação¹ na imagem de regiões ou se a imagem de orientação apresenta um padrão de linhas de fluxo das cristas que não corresponde ao padrão de uma das cinco classes (ver Figura 5.1).

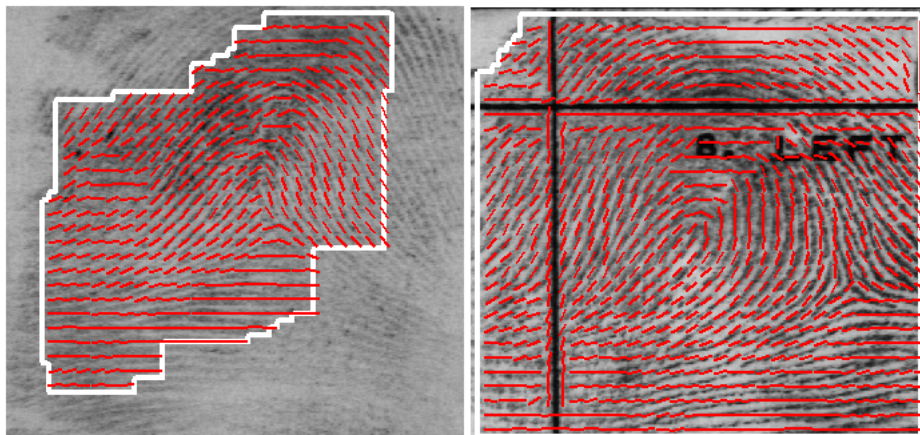


Figura 5.1: Imagens de má qualidade.

- Uma imagem é rotulada de média qualidade se não ocorre uma super segmentação e a imagem de orientação apresenta um padrão de fluxo de cristas que corresponde a uma das cinco classes, porém tem regiões onde as orientações não estão bem definidas (ver Figura 5.2).
- Uma imagem é rotulada de boa qualidade se não ocorre uma super segmentação e a imagem de orientação apresenta um padrão de fluxo de cristas que corresponde a uma das cinco classes (ver Figura 5.3).

¹é gerar uma segmentação extra, ou seja, que o segundo plano contenha regiões do

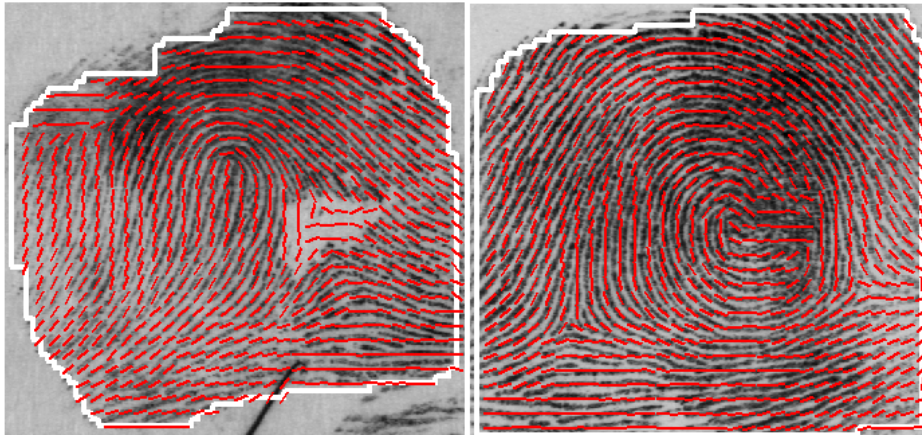


Figura 5.2: Imagens de média qualidade.

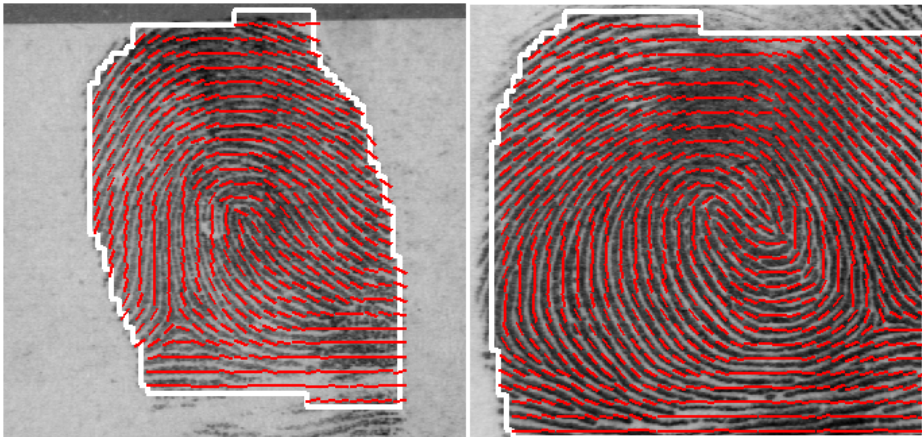


Figura 5.3: Imagens de boa qualidade.

A distribuição das impressões segundo sua qualidade é mostrada na Tabela 5.4, onde observa-se que o 22,35% do número total de imagens na DB4 corresponde a impressões de pobre qualidade, sendo a classe TA a que tem o maior número de impressões de pobre qualidade. Somado ao fato de que a classe TA apresenta-se em menor quantidade na natureza, o problema de classificação fica ainda mais complexo.

primeiro plano

Qualidade	Classes					
	LL	W	RL	TA	A	Total(%)
Pobre qualidade	193	162	185	222	132	22,350
Média qualidade	91	73	71	94	66	9,875
Boa qualidade	516	565	544	484	602	67,775

Tabela 5.2: Distribuição das impressões digitais da DB4 segundo sua qualidade.

5.4 Rendimento das técnicas de classificação de impressões digitais

O rendimento das técnicas de classificação de impressões digitais é usualmente medido em termos da taxa de erro ou pela exatidão. A taxa de erro é calculada como o taxa entre o número de impressões digitais mal classificadas e o número total de amostras no conjunto de prova; a exatidão é calculada como o porcentagem das impressões digitais classificadas corretamente.

$$taxa\ erro = \frac{número\ de\ impressões\ digitais\ mal\ classificadas}{número\ total\ de\ impressões\ digitais}, \quad (5.1)$$

$$exatidão = 1 - taxa\ erro. \quad (5.2)$$

Um análise mais detalhada do comportamento de um classificador pode ser obtido pelo exame da matriz de confusão. Esta matriz possui uma linha para cada classe verdadeira e uma coluna para cada classe hipotética; cada célula da linha r e coluna c indicam quantas impressões digitais pertencentes à classe r são (in)corretamente atribuídas à classe c . A Tabela 5.3 mostra exemplos da matriz de confusão.

5.5 Resultados prévios obtidos sobre a DB4

A Tabela 5.4 reporta as taxas de erro sobre a DB4 de 14 técnicas de classificação: a maioria delas foi obtida usando as 2.000 imagens das primeiras 1.000 impressões (de F0001 a F1000 e S0001 a S1000) para o treinamento e as restantes 2.000 imagens para a prova do sistema (de F1001 a F2000 e S1001 a S2000). Alguns métodos baseados em regras (tais como [56], [43], [49] e [27]) são testados sobre todo o banco de dados. Em [89] reporta-se resultados sobre 542 impressões digitais selecionadas aleatoriamente. Todos os resultados são reportados com uma taxa de rejeição do 0%, com a exceção das técnicas

Classes Verdadeiras	Classes Atribuidas			
	LL	W	RL	A/TA
LL	370	2	4	10
W	4	384	14	0
RL	2	2	380	8
A/TA	66	5	112	637

(a)

Classes Verdadeiras	Classes Atribuidas				
	LL	W	RL	TA	A
LL	375	5	3	14	2
W	6	378	10	1	3
RL	2	1	394	17	5
TA	12	1	8	292	32
A	5	0	0	12	422

(b)

Tabela 5.3: Matriz de confusão das técnicas propostas por : (a) Senior [90] para o problema de classificação em 4 classes e por (b) Capelli, Maio e Maltoni [87] para o problema de classificação em 5 classes, ambas técnicas foram testadas sobre o banco de dados DB4.

[50], [69] e [105] onde 1.8% das impressões digitais foram rejeitadas durante a fase de extração de características. Para as impressões digitais com duas classes, o resultado é usualmente assumido correto se a classe suposta pelo classificador coincide com qualquer um dos rótulos.

Alguns autores notaram que a maioria dos erros é devido a uma incorreta classificação de algumas impressões digitais do tipo TA como se fossem do tipo A (por exemplo, ver a matriz de confusão da Tabela 5.3.b), considerando que estas classes são pouco comuns na natureza (2,9% e 3,7% respectivamente) propôs-se misturar as duas classes numa única classe (problema de quatro classes). Quando estão disponíveis, os resultados para ambos problemas (de quatro classes e cinco classes) são reportados na Tabela 5.4.

Posto que a DB4 tem um número igual de impressões digitais para cada classe, alguns autores preferem ponderar os resultados de acordo com a distribuição natural das classes (segundo o procedimento especificado em [10][90]), para obter uma estimativa da classificação em condições reais. Por exemplo, para o problema de 4 classes, um classificador que reconhece muito bem as impressões tipo A-TA, apresentará uma mínima taxa de erro ao ser testado com as 2.000 imagens de teste da DB4, pois as imagens do tipo

A-TA de tal conjunto são 800 imagens. Na realidade esta qualidade do classificador não é muito útil uma vez que as impressões A-TA são as que menos aparecem na natureza. Os resultados ponderados, publicados pelos autores ou derivados das matrizes de confusão, são reportados na Tabela 5.4, para ambos problemas (quatro classes e cinco classes).

Técnica	Conjunto de teste	4 classes		5 classes	
		N	B	N	B
Blue et al. 1993 [10]	2 ^a metade			7,2	
Candela et al. 1995 [15]	2 ^a metade	6,1	11,4		
Karu e Jain 1996 [56]	Toda a DB4	9,4	8,6		
Senior 1997 [89]	542 aleatórias		8,4		
Capelli, Maio e Maltoni 1999 [16]	2 ^a metade		5,5	6,5	7,9
Jain, Prabhakar e Hong 1999 [50]	2 ^a metade	5,2		7,0	10,0
Hong e Jain 1999 [43]	Toda a DB4		7,7	10,6	12,5
Marcialis, Roli e Frasconi 2001 [69]	2 ^a metade			9,6	12,1
Yao, Frasconi e Pontil 2001 [105]	2 ^a metade		6,9	9,0	10,7
Senior 2001 [90]	2 ^a metade	5,1			
Jain e Minut 2002 [49]	Toda a DB4	9,3	8,8		
Dass e Jain 2004 [27]	Toda a DB4		5,6		
Park e Park 2005 [80]	2 ^a metade		6,0		9,3
Tan, Bhanu e Lin 2005 [94]	2 ^a metade		6,1		7,5

Tabela 5.4: Taxas de erro de 14 técnicas testadas com a DB4 para os problemas de 4 e 5 classes, em que o rótulo **N** significa que os resultados da classificação são ponderados de acordo a distribuição 'natural' das classes e o rótulo **B** significa que os resultados da classificação não são ponderados, posto que assume-se uma distribuição 'balanceada' de cada classe.

5.6 Resultados

Nesta seção são apresentados os resultados das provas realizadas sobre o sistema de classificação desenvolvido, seguindo a metodologia usada na maioria de trabalhos vinculados à classificação de impressões digitais. Portanto, o banco de dados utilizado foi o DB4 do NIST, usando as 2.000 imagens das primeiras 1.000 impressões para o treinamento e as restantes 2000 imagens para a prova do sistema. Os resultados da classificação são apresentados para os problemas de 4 e 5 classes e todas as taxas de erro são apresentadas considerando uma distribuição uniforme e natural. as técnicas de melhora e segmentação foram validadas num contexto diferente, utilizando um banco

de dados especializado para a tarefa de reconhecimento (ver [2]) posto que nesta problemática pode-se testar a robustez destas técnicas. Porém, para o caso de segmentação, foi feita uma comprovação baseada na observação das imagens de regiões calculadas pela técnica desenvolvida, encontrando-se o problema da super segmentação em 52 imagens num total de 4.000 imagens. Na Figura 5.4 pode-se ver alguns exemplos destes casos. Basicamente o erro na segmentação ocorre quando a imagem de entrada apresenta regiões onde a diferença entre seu mínimo e máximo valor de escala de cinza é pequena. Isto acontece porque a técnica de segmentação utiliza o mínimo valor da escala de cinza de cada setor em que foi dividida a imagem para realizar a segmentação, portanto, considera como fundo aquelas regiões que têm um valor mínimo próximo ao máximo, produzindo uma super segmentação. Estas imagens não foram descartadas pois representam as impressões digitais parciais, as quais são comuns no âmbito forense, tendo a técnica de classificação a tarefa de atribuir uma classe, apesar de ter informação incompleta.

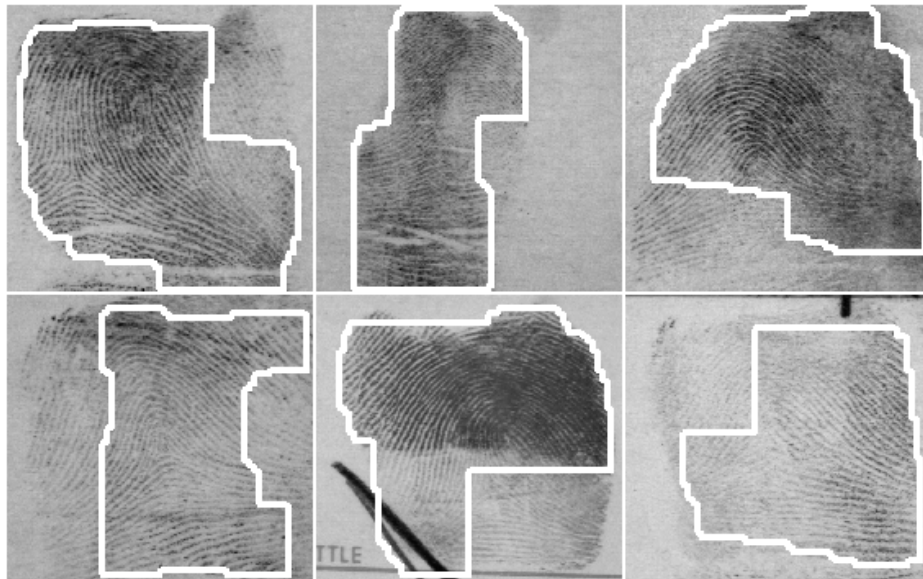


Figura 5.4: Imagens que apresentam a super segmentação.

O rendimento das técnicas de detecção do ponto referência são validados observando-se a capacidade que estes têm de detectar sua posição com precisão. Vista desta maneira, a análise da robustez destas técnicas fica resumida em avaliar a influência que tem a técnica utilizada para o cálculo da imagem de orientação e a influência da qualidade da impressão digital, para o qual desenvolveu-se a seguinte experiência:

1. Para cada uma das 4.000 imagens da DB4:

- Um especialista determina manualmente o ponto de referência da imagem, considerando a definição básica do ponto de referência (aquele ponto que tem a máxima curvatura na crista mais interna), tal que para aquelas imagens onde o ponto de referência não é detectável pela má qualidade da imagem ou porque não cumpre com a definição (algumas impressões tipo A), assume-se como ponto de referência o ponto central da imagem.
 - Calcula-se o ponto de referência da imagem utilizando uma das técnicas implementadas (baseada na detecção do ponto *core* ou baseada em múltiplos níveis de representação da imagem de orientação), tendo estas técnicas como entrada a imagem de orientação determinada por alguma das técnicas de estimação local da orientação.
 - Define-se uma vizinhança de 61x61 pixels centrada no ponto de referência detectado manualmente, se o ponto de referência detectado pela técnica em questão está dentro da vizinhança, então considera-se como uma correta estimação do ponto de referência, caso contrário, supõe-se que a técnica não foi precisa na determinação do ponto de referência.
2. Calcula-se a taxa de erro de precisão, definida como o número de vezes que não teve-se uma correta determinação do ponto de referência dividido pelo número total de imagens no banco de dados. Da mesma forma, define-se a taxa de erro de precisão segundo o tipo de impressão digital e segundo a qualidade da impressão digital.

Os resultados da realização desta experiência são resumidos nas tabelas 5.5, 5.6 e 5.7. Pode-se observar a grande influência que tem a seleção da técnica de estimação local da orientação. Neste caso, a técnica de estimação local que gera a menor taxa de erro de precisão, para ambas técnicas de detecção do ponto de referência, é aquela proposta em [52], aqui obteve-se uma taxa de erro de precisão para a técnica baseada no ponto *core* de 21,675% e para a técnica proposta, baseada em múltiplos níveis de representação da imagem de orientação, de 8,325%. O desempenho superior é devido à capacidade da técnica proposta de detectar com precisão o ponto de referência das impressões digitais tipo TA e A, porém, a exatidão na detecção do ponto de referência para as impressões digitais tipo LL, RL e W é comparável à técnica baseada no índice de Poincaré. Para o caso das impressões digitais tipo W, a técnica baseada no índice de Poincaré é superior à técnica proposta (5,87% versus 8,125%, respectivamente). Este fato deve-se a que as

impressões digitais tipo W apresentam dois pontos *core* um ponto *core* inferior e outro superior e, algumas vezes eles estão muito próximos, fazendo que a técnica proposta considere o ponto *core* inferior como ponto de referência (a suposição principal desta técnica é que o ponto de referência encontra-se nas proximidades do elemento mais interno da parte superior de Θ , cuja orientação quantizada seja 0°). Em geral, a técnica baseada no ponto *core* para qualquer das técnicas de estimação local tem uma taxa de erro médio de precisão de 21,8% a diferença da técnica baseada em múltiplos níveis, que para o caso da técnica de estimação local proposta em [97] tem uma taxa de erro de precisão de 23,2%, deve-se este resultado a que a matriz de coerência que produz esta técnica está muito suavizada. Na Figura 5.5) pode-se ver um exemplo, onde o ponto de referência é estimado com precisão (o quadrado), já que está dentro da vizinhança centrada no ponto de referência determinado manualmente (o círculo), porém a reestimação de sua posição (o triângulo) fica fora da vizinhança centrada no ponto de referência determinado manualmente, uma vez que ao estar a matriz de coerência muito suavizada seus valores mínimos (que definem as possíveis posições das singularidades) são eliminados e portanto o ponto reestimado não necessariamente ficará perto do ponto de referência real da impressão digital.

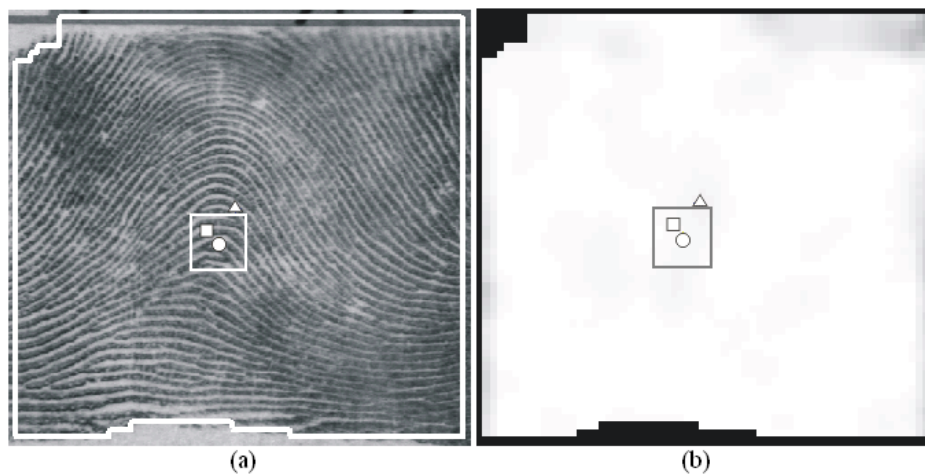


Figura 5.5: (a) impressão digital, (b) matriz de coerência.

Com respeito à qualidade das imagens, considerando unicamente a técnica de estimação local proposta por [52], a técnica baseada em múltiplos níveis é superior à técnica baseada no ponto *core* nos três grupos de qualidade. Em que a maior taxa de erro de precisão corresponde as imagens de pobre qualidade (20,18% do número total de imagens de pobre qualidade), este resultado é interessante já que é preferível que o erro na estimação do

ponto de referência esteja concentrado nas impressões da baixa qualidade do que nas de média e boa qualidade posto que elas têm maior probabilidade de serem classificadas corretamente, ou o contrário, ou seja, que o erro na estimação do ponto de referência esteja concentrado nas imagens de média ou boa qualidade (como no caso da técnica do ponto *core*) somaria a má classificação das impressões de pobre qualidade com a má classificação das impressões cujo ponto de referência foi mal estimado, gerando uma taxa de erro de classificação maior, e portanto uma menor precisão na classificação.

Técnica	Erro (%)								
	Total	Por Classes					Por qualidade		
		LL	W	RL	TA	A	MC	MD	BC
Poincaré	22,05	8,5	7,25	6,75	25,62	62,12	28,93	18,98	19,40
Proposta	14,42	9,87	10,37	8,00	15,37	28,50	31,35	14,93	9,36

Tabela 5.5: Taxas de erro de precisão quando utiliza-se a técnica de estimação local da orientação padrão, em que o rótulo (**MC**) faz referência as impressões de má qualidade, o rótulo (**MD**) faz referência às impressões de média qualidade e o rótulo (**BC**) faz referência às impressões de boa qualidade.

Técnica	Erro (%)								
	Total	Por Classes					Por qualidade		
		LL	W	RL	TA	A	MC	MD	BC
Poincaré	21,8	8,50	7,00	10,5	21,5	61,50	32,30	19,74	19,25
Proposta	23,2	14,87	15,37	19,5	29,2	37,00	40,96	26,58	17,63

Tabela 5.6: Taxas de erro de precisão quando utiliza-se a técnica de estimação local da orientação proposta por [97].

Técnica	Erro (%)								
	Total	Por Classes					Por qualidade		
		LL	W	RL	TA	A	MC	MD	BC
Poincaré	21,67	7,50	5,87	7,35	27,25	60,37	28,74	17,46	20,51
Proposta	8,32	4,75	8,12	5,00	7,00	16,50	20,18	8,86	4,72

Tabela 5.7: Taxas de erro de precisão quando utiliza-se a técnica de estimação local da orientação proposta por [52].

A seguir são explicados os experimentos realizados para a determinação da taxa de erro dos classificadores. Aqui utilizou-se a técnica de estimação

local da orientação proposta em [52] para o cálculo da imagem de orientação e a técnica baseada em múltiplos níveis de representação da imagem de orientação para a determinação do ponto de referência, por seu bom desempenho mostrado. O protocolo seguido para o estabelecimento das imagens usadas para treino e teste dos classificadores é: as 2.000 primeiras imagens da DB4 vinculadas às primeiras 1.000 impressões são para o treino e as restantes 2000 imagens são para teste. Ao utilizar este protocolo os resultados obtidos podem ser comparados com os resultados publicados na literatura (ver Tabela 5.4).

Para cada um dos classificadores utilizados são realizados dois experimentos:

- o primeiro permite analisar a resposta do sistema de classificação trabalhando de maneira automatizada, ou seja, não tem a intervenção de um especialista em nenhuma parte do processo de classificação,
- o segundo permite analisar a resposta do sistema de classificação trabalhando de maneira semi-automatizada, ou seja, considera-se a intervenção de um especialista numa única parte do processo de classificação, na correção do ponto de referência das impressões digitais pertencentes ao grupo de teste, tal como trabalha um sistema real de classificação de impressões digitais.

A metodologia seguida para cada experimento é explicado na continuação:

1. Define-se o problema de classificação ao analisar (o problema de 4 classes ou o problema de 5 classes)
2. Para cada uma das 2.000 imagens de treino:
 - Determina-se sua imagem de regiões, sua imagem de orientação, o possível ponto de referência que esta tenha segundo as técnicas explicadas no capítulo 2.
 - Realiza-se o registro da imagem de orientação e determina-se sua representação vetorial definida pelas relações (4.1) e (4.2).
 - Aplica-se a representação vetorial da imagem de orientação ao PCA obtendo-se, um vetor de características de n elementos.
3. Obtêm-se as características dos classificadores, segundo o problema de classificação a analisar. Para o classificador normal calcula-se as estimações das matrizes de covariância, as estimações dos vetores médios de cada classe e estabelece-se as probabilidades *a priori* de cada classe

segundo a distribuição natural das impressões digitais, para o MLP determina-se os valores dos pesos através de uma técnica de treino, para o k -NN define-se o número de vizinhos a considerar, e para o PNN define-se o fator de suavizamento σ .

4. Para cada uma das 2.000 imagens de teste:
 - Determina-se sua imagem de regiões, sua imagem de orientação, o possível ponto de referência que esta tenha segundo as técnicas explicadas no capítulo 2. Se é realizado o experimento em que o sistema trabalha de maneira semi-automatizada, o ponto de referência determinado é validado por um especialista, tal que, caso ele esteja mal calculado, ele é corrigido.
 - Realiza-se o registro da imagem de orientação e determina-se sua representação vetorial definida pelas relações (4.1) e (4.2).
 - Aplica-se a representação vetorial da imagem de orientação ao PCA, obtendo-se, o vetor de características de n elementos.
5. Apresenta-se na entrada do classificador cada um dos vetores de características e determina-se a classe suposta da impressão digital vinculada ao vetor de características em questão.
6. Calcula-se a taxa de erro de classificação considerando uma distribuição balanceada e natural das impressões digitais.

Seguindo o procedimento descrito, as tabelas 5.8, 5.9, 5.10 e 5.11 são obtidas, onde observa-se a taxa de erro para os problemas de 4 e 5 classes quando utiliza-se os classificadores Normal, k -NN, MLP e PNN, respectivamente. Para os classificadores Normal, k -NN e PNN são determinadas as taxas de erro para diferentes valores de n , para o caso do classificador MLP determina-se a taxa de erro médio de classificação para $n = 112$, dado que com este número de elementos do vetor de características o MLP teve a melhor resposta na classificação. Um fato de interesse é que o classificador PNN é sintonizado para um $\sigma = 6$, este valor foi o resultado de um processo de tentativa e erro, onde o valor de σ foi modificado entre 1 e 20 com incrementos de 0.1, tal que para $\sigma = 6$ teve-se as menores taxas de erro de classificação.

Destes resultados podemos observar que para o experimento onde o sistema trabalha de maneira totalmente automatizada o classificador que apresenta a melhor resposta é:

- para o problema de 4 classes, o classificador Normal apresenta a melhor resposta quando $n = 38$, tendo uma taxa de erro de classificação

	Automático (%)				Semi-automático (%)			
	4 classes		5 classes		4 classes		5 classes	
PCA	N	B	N	B	N	B	N	B
30	7,53	7,30	9,45	13,00	6,05	6,00	7,86	11,45
32	7,34	7,00	9,25	13,05	5,96	5,75	7,82	11,55
34	7,60	7,00	8,72	12,25	6,29	5,75	7,38	10,85
36	7,68	6,90	8,66	12,35	6,28	5,60	7,29	10,80
38	7,19	6,75	8,34	12,30	6,02	5,55	7,00	11,00
40	7,98	7,30	9,29	12,90	6,33	5,75	7,74	11,55

Tabela 5.8: Taxa de erro para os problemas de classificação de 4 e 5 classes, quando é usado um classificador normal, para diferentes valores de n (número de características).

	Automático (%)				Semi-automático (%)			
	4 classes		5 classes		4 classes		5 classes	
k	7		5		7		5	
PCA	N	B	N	B	N	B	N	B
26	13,36	9,85	14,36	17,05	12,29	8,90	13,30	16,05
30	13,72	10,10	13,97	16,70	12,65	9,20	13,06	15,90
32	13,78	10,10	14,28	16,40	12,56	9,10	13,29	16,10
34	14,01	10,40	14,30	16,85	12,71	9,30	13,33	16,00
36	14,13	10,45	13,92	16,60	12,83	9,35	12,92	15,70

Tabela 5.9: Taxa de erro para os problemas de classificação de 4 e 5 classes, quando é usado um classificador k -NN, para diferentes valores de n .

	Automático (%)				Semi-automático (%)			
	4 classes		5 classes		4 classes		5 classes	
PCA	N	B	N	B	N	B	N	B
112	8,94	7,15	9,27	10,70	7,57	5,70	9,02	10,02
	9,07	7,30	9,39	9,75	7,79	5,95	8,84	9,95
	9,03	7,35	9,50	10,70	7,54	5,70	8,37	9,95
	9,01	7,30	9,15	10,60	7,79	5,95	9,02	10,08
Média	9,24	7,27	9,33	10,68	7,67	5,82	8,81	10,00

Tabela 5.10: Taxa de erro para os problemas de classificação de 4 e 5 classes, quando é usado um classificador MLP, para diferentes valores de n .

de 7,19%, para uma distribuição natural e uma taxa de erro de classi-

	Automático (%)				Semi-automático (%)			
	4 classes		5 classes		4 classes		5 classes	
σ	6		6		6		6	
PCA	N	B	N	B	N	B	N	B
42	9,73	15,95	10,14	21,10	8,33	14,65	10,14	21,10
44	9,58	15,10	10,00	21,30	8,42	14,85	8,88	20,35
48	9,56	15,7	10,14	21,10	8,33	14,65	10,14	21,10
50	9,70	15,65	10,29	21,15	8,35	14,60	9,01	20,15

Tabela 5.11: Taxa de erro para os problemas de classificação de 4 e 5 classes, quando é usado um classificador PNN, para diferentes valores de n .

figação de 6, 75% para uma distribuição balanceada.

- Para o problema de 5 classes, considerando uma distribuição natural, o classificador Normal apresenta a melhor resposta quando $n = 38$, tendo uma taxa de erro de classificação de 8,34%, agora considerando uma distribuição balanceada o MLP apresenta a melhor resposta quando $n = 112$, tendo uma taxa de erro de classificação média de 10,98%.

Para o experimento onde o sistema trabalha de maneira semi-automatizada o classificador que apresenta a melhor resposta é:

- Para o problema de 4 classes, considerando uma distribuição natural, o classificador Normal apresenta a melhor resposta quando $n = 38$, tendo uma taxa de erro de classificação de 5,55%. Agora, considerando uma distribuição balanceada, também o classificador Normal apresenta a melhor resposta quando $n = 32$, tendo uma taxa de erro de classificação de 5,95%.
- Para o problema de 5 classes, considerando uma distribuição natural o classificador Normal apresenta a melhor resposta quando $n = 38$, tendo uma taxa de erro de classificação de 7,00%. Agora considerando uma distribuição balanceada o MLP apresenta a melhor resposta quando $n = 112$, tendo uma taxa de erro de classificação média de 10,00%.

Pode-se ver que para ambos experimentos e considerando uma distribuição natural, o classificador Normal é o que melhor resultados apresenta, tanto para o problema de classificação em 4 como em 5 classes, o que é um resultado de suma importância, posto que o classificador normal é o mais simples dos implementados. As possíveis explicações deste resultado são:

- Que os vetores de características vinculados a cada classe tenham a tendência a apresentar uma distribuição unimodal normal multivariada, como consequência da etapa de processamento implementada, posto que, na literatura, não se tem resultados equivalentes considerando unicamente um classificador normal² (o PCASYS treinado com um classificador normal tem uma taxa de erro de 11,3% para o problema de 5 classes e considerando uma distribuição natural [10]).
- que a inclusão das probabilidades *a priori* nas funções discriminantes de cada classe seja um fator determinante ao momento de definir a pertinência dos vetores de características a uma classe, considerando que elas foram estimadas sobre um banco de dados de 222 milhões de pessoas [87].

Considerando os resultados obtidos para o PNN implementado, que inclui em suas funções discriminantes as probabilidades *a priori* de cada classe, a segunda suposição não gera uma explicação completa. Portanto, a melhor explicação está fundamentada na primeira suposição, somado ao fato de incluir a informação *a priori*, em tal sentido, assumimos que de maneira restrita para o problema de classificação em 4 classes (posto que para este problema o classificador normal teve as taxas de erro mais baixas) os vetores de características aparentemente apresentaram uma distribuição unimodal e razoavelmente bem representada por uma normal multivariada. Para o MLP o fato de utilizar-se 112 neurônios na camada oculta pode ter produzido uma inadequação da ordem da função discriminante com o ordem da função necessária para separar as classes mesmo que o erro de treinamento obtido tenha sido alcançado. Este fato é reforçado pois a normal multivariável empregada deu bons resultados, tal que as fronteiras de decisão geradas por este classificador concordam melhor com a natureza das distribuições de probabilidade dos vetores de características de cada classe mitigando a inclusão de vetores de uma classe em outra (ver Figura 5.6). Para o classificador k -NN seus resultados cumprem com o esperado segundo a relação definida por (4.15).

²Em [94] é utilizado também um classificador Normal porém os vetores de características são produto de operadores de composição aplicados sobre 16 imagens vinculadas à imagem original (a imagem de orientação, os componentes seno e cosseno da imagem de orientação, a média, a desvio padrão, o máximo, o mínimo e o valor mediana valor de cada vizinhança da 3x3 e 5x5 da imagem de orientação, imagens de bordas obtidas ao se aplicar filtros de Sobel ao longo das direções verticais e horizontais e a imagem de orientação limiarizada em 90°.), os operadores de composição são determinados através de programação genética

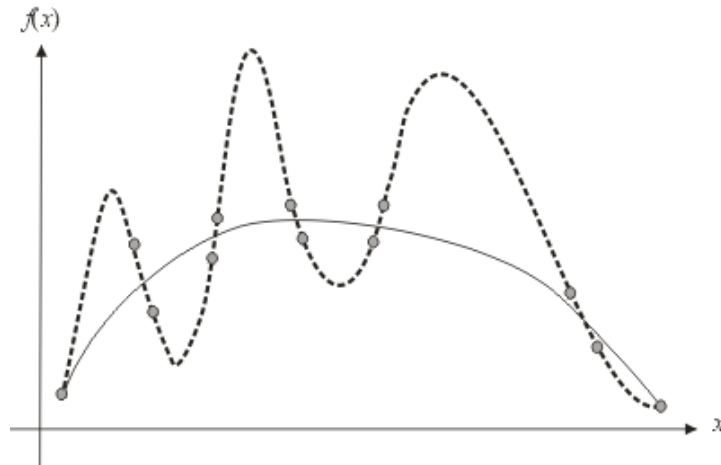


Figura 5.6: A função discriminante do MLP (linha pontuada) passa por todos os pontos de treinamento (círculos cinza), porém a função discriminante do classificador normal (linha fina) permite uma melhor previsão para as novas amostras [31].

Na tabela 5.12 pode-se observar as matrizes de confusão para o classificador Normal, quando têm a melhor resposta em cada um dos experimentos realizados. Para o caso da classificação em 5 classes, a maior taxa de erro na classificação encontra-se nos tipos TA e A, o que é o esperado, dado que entre estes tipos existe uma pequena variabilidade inter-classes. Para o caso da classificação em 4 classes, a maior taxa de erro na classificação encontra-se nos tipos RL e TA-A. Isto deve-se a que existem algumas impressões tipo RL que apresentam um ponto *core* que fica próximo de seu correspondente ponto *delta*, fazendo que a característica principal do tipo RL (as linhas de crista que ingressam e saem pelo mesmo lado por onde ingressaram) seja atenuada no momento de obter sua imagem de orientação. Para os dois problemas pode-se observar que também existe um erro de classificação entre os tipos LL-RL e W, estes erros devem-se a que o tipo W apresenta uma sub classe denominada *Central Pocket* ela tem dois pontos *core* e dois pontos *delta*, porém um dos pontos delta está invertido, o que faz com que algumas linhas de fluxo da imagem de orientação saiam pelo mesmo lado por onde ingressaram, a qual é a característica principal dos tipos LL e RL. Este comportamento fica ressaltado segundo o grau de suavizamento que gera a técnica de estimação local da orientação empregada.

Na tabela 5.13 observa-se as matrizes de confusão considerando a qualidade das imagens e a exatidão no cálculo do ponto de referência.

- Em 5.13.a tem-se a matriz de confusão considerando unicamente as impressões com super segmentação, nela pode-se ver que só tem-se 4

	LL	W	RL	A/TA
LL	354	17	1	14
W	5	389	4	4
RL	0	6	354	32
A/TA	23	10	19	768

(a)

	LL	W	RL	TA	A
LL	353	17	1	12	3
W	7	389	4	1	1
RL	0	6	351	30	5
TA	13	10	16	294	57
A	8	0	2	43	367

(b)

	LL	W	RL	A/TA
LL	356	16	0	14
W	2	392	6	2
RL	0	5	364	23
A/TA	22	7	18	773

(c)

	LL	W	RL	TA	A
LL	356	16	0	12	2
W	2	394	4	0	2
RL	0	5	359	27	1
TA	11	9	15	299	66
A	4	0	2	42	372

(d)

Tabela 5.12: Matrizes de confusão quando é usado um classificador normal para os problemas de: (a) 4 classes com $n = 38$ e registro automático; (b) 5 classes com $n = 38$ e registro automático; (c) 4 classes com $n = 32$ e registro semi-automático; (d) 5 classes com $n = 38$ e registro semi-automático.

erros na classificação, indicando assim que o classificador normal tem a capacidade de classificar as impressões digitais parciais.

- Em 5.13.b tem-se a matriz de confusão das impressões de pobre qualidade, onde 14,93% sobre o total delas são mal classificadas, o erro de classificação concentra-se principalmente entre os tipos A e TA.
- Em 5.13.c tem-se a matriz de confusão das impressões de média qualidade, onde 16,57% sobre o total delas são mal classificadas, o erro de classificação também concentra-se entre os tipos A e TA.
- Em 5.13.d tem-se a matriz de confusão das impressões de boa qualidade, onde 15,38% sobre o total delas são mal classificadas, o erro de classificação encontra-se entre os pares de classes A e TA, LL e W, RL e TA, lembrando que a atribuição de rótulos de qualidade foi realizado analisando a imagem de orientação e a imagem de regiões. Pode-se explicar este resultado como produto da técnica de estimação local da orientação, posto que ela gerará imagens de orientação bem comportadas, porém para alguns casos, ela atenuará ou deslocará as posições das singularidades e portanto gerará a diminuição da variabilidade inter-classes. Por isso é que os erros acontecem entre os tipos

A e TA, cuja variabilidade inter-classes é sempre pequena, LL e W, que devido à sub classe *Central pocket* elas ficam misturadas quando a imagem de orientação é muito suavizada. Agora, para os tipos RL e TA, a explicação se deve a que para alguns casos os pontos *core* e *delta* de uma impressão tipo RL estão na mesma linha vertical traçada desde o ponto *core*, portanto suas linhas de fluxo de cristas tomam uma forma acampanada clássica dos tipos TA, fazendo-se com que para alguns casos esta variação do tipo RL seja classificada como do tipo TA.

- Em 5.13.e tem-se a matriz de confusão das impressões mal registradas é classificadas no experimento semi-automático; em 5.13.f tem-se a matriz de confusão das mesmas impressões mal registradas porém corrigidas e classificadas no experimento semi-automático. Pode-se ver que para os tipos LL, W e RL o fato de corrigir seus pontos de referência melhora a classificação, porém para os tipos TA-A os erros de classificação quase mantêm-se iguais, devido a que um registro correto não atenua em grande medida a pequena variabilidade inter-classes que existe entre elas.

5.7 Resumo

Neste capítulo foram realizadas as provas para a avaliação do sistema de classificação implementado onde, a menor taxa de erro de classificação considerando uma distribuição natural para o problema de 4 e 5 classes foi obtida pelo classificador normal (7,19% e 8,34% para o experimento onde o sistema trabalha de maneira automatizada e 6,02% e 7,00% para o experimento onde o sistema trabalha de maneira semi-automatizada). Comparando estes resultados com obtidos na literatura (ver tabela 5.4) estas taxas de erro são equivalentes às obtidas por técnicas de classificação mais sofisticadas, tais como a combinação de SVM (*support vector machine*) propostas em [105] (com uma taxa de erro de 9,0% para o problema de 5 classes considerando uma distribuição natural), um classificador em duas fases (um k -NN e 10 redes neurais) apresentado em [50] (com uma taxa de erro de 9,0% para o problema de 5 classes segundo uma distribuição natural e considerando uma taxa de rejeição de 1,8%), pela combinação de três classificadores (PCASYS, HMM e árvores de decisão) apresentado em [90] (com uma taxa de erro de 5,1% para o problema de 4 classes considerando uma distribuição natural) e por uma técnica de representação baseada numa generalização da transformada KL proposta [16] (com uma taxa de erro de 6,5% para o problema

	LL	W	RL	TA	A
LL	5	0	0	0	1
W	0	1	0	0	0
RL	0	0	7	0	0
TA	0	0	0	8	3
A	0	0	0	0	1

(a)

	LL	W	RL	TA	A
LL	65	4	1	2	0
W	5	72	0	1	0
RL	0	2	51	7	3
TA	4	3	5	75	13
A	2	0	2	7	44

(b)

	LL	W	RL	TA	A
LL	50	1	0	6	0
W	0	41	0	0	1
RL	0	0	38	5	0
TA	0	3	2	41	7
A	2	0	0	9	35

(c)

	LL	W	RL	TA	A
LL	233	12	0	4	2
W	2	275	4	0	0
RL	0	4	255	18	2
TA	9	4	9	170	44
A	4	0	0	27	287

(d)

	LL	W	RL	TA	A
LL	15	1	1	1	1
W	5	26	0	1	1
RL	0	1	7	7	4
TA	4	2	2	14	10
A	5	0	0	12	50

(e)

	LL	W	RL	TA	A
LL	18	0	0	1	0
W	0	31	0	0	2
RL	0	0	15	4	0
TA	2	1	1	19	9
A	1	0	0	11	55

(f)

Tabela 5.13: Matrizes de confusão quando é usado um classificador normal para o problemas de 5 classes com $n = 38$, em que cada matriz corresponde: (a) as impressões mal segmentadas; (b) as impressões de má qualidade; (c) as impressões de média qualidade; (d) as impressões de boa qualidade; (e) as impressões mal registradas; (f) as impressões mal registradas porém corrigidas manualmente.

de 5 classes considerando uma distribuição natural). Este resultado também valida a etapa de processamento dado que o bom desempenho do sistema de classificação não só se deve ao classificador (nosso caso um classificador normal) como também às etapas anteriores.

Capítulo 6

Coclusões e Projetos Futuros

6.1 Conclusões

Quanto à determinação da imagem de regiões propõe-se um nova técnica para a segmentação de imagens de impressões digitais, que utiliza como característica para a segmentação o mínimo valor na escala de cinza de cada bloco em que é dividida a imagem, observando-se que: (a) o uso do mínimo valor de cada bloco e sua posterior filtragem através de um filtro de mediana, gera uma característica que apresenta uma boa capacidade de discriminação entre os valores pertencentes ao primeiro e segundo plano, (b) a inclusão de uma etapa de pós-processamento e o uso de múltiplos limiares assegura que a eleição da imagem de segmentação final seja aquela que abarca a maior área vinculada ao primeiro plano representado pelos mínimos valores de intensidade dos blocos. Em conjunto, a técnica é robusta, posto que, apresentou o problema de super segmentação em 52 imagens de um total de 4000. Além disso, ela é rápida, já que a segmentação é realizada no nível de blocos (64x64 blocos) utilizando uma única característica para realizar a segmentação.

Quanto à melhora da impressão digital propõe-se um nova técnica de normalização baseada na transformada *Tophat* multi-escala. Esta técnica uniformiza com êxito o contraste na imagem, gerando em conjunto com a técnica de melhora no domínio da frequência uma imagem em que o efeito do contraste como o de ruído são atenuados em grande medida permitindo realizar uma estimação precisa da imagem de orientação.

Quanto à determinação da imagem de orientação foram testadas três técnicas para a estimação local da orientação, dos resultados experimentais

e tomando como referência a resposta da técnica de detecção do ponto de referência, a selecionada foi a técnica proposta por [52]. Esta técnica utiliza a informação dada pela matriz de coerência para realizar uma estimação local da orientação de maneira ponderada, realizando um suavizamento adaptativo da imagem de orientação.

Quanto à determinação do ponto de referência propõe-se uma nova técnica baseada em múltiplos níveis de representação da imagem de orientação. Esta é uma nova concepção que permite utilizar as características morfológicas da imagem de orientação limiarizada para cada um de seus valores de quantização. A técnica apresenta as seguintes características: (a) adapta-se segundo a qualidade da imagem, (b) detecta os pontos de referência das impressões digitais A e TA obtendo-se taxas de erro aceitáveis para estes tipos de impressões, (c) o uso da matriz de coerência aumenta sua exatidão na determinação do ponto de referência, porém faz a técnica depender da qualidade e do nível de suavizamento com que foi estimada a matriz de coerência, (d) sempre tem uma resposta caso não detecte um ponto de referência válido, permitindo misturar esta técnica com outras, (e) ela foi testada em todo o banco de dados da DB4 obtendo um erro de estimação do ponto de referência de 8,32% sobre um total de 4000 imagens.

Quanto à tarefa de classificação foram implementados todos componentes de um classificador de impressões digitais baseada na arquitetura proposta no PCASYS. Tal arquitetura foi testada com o banco de dados DB4 do NIST, utilizando-se os classificadores: Normal, k -NN, MLP e PNN, onde determinou-se as taxas de erro para os problemas de 4 e 5 classes considerando uma distribuição natural e balanceada. Dos resultados obtidos e tendo em conta uma distribuição balanceada, o melhor classificador foi o Normal, o qual é um resultado de importância, posto que permite fazer a suposição que: somado ao fato de incluir a informação *a priori* e de maneira restrita para o problema de quatro classes, os vetores de características vinculados a cada classe aparentemente apresentaram uma distribuição unimodal e razoavelmente bem representada por uma normal multivariada, como consequência de ter realizado uma etapa de processamento que atenua em grande medida os erros que costumam gerar uma má classificação. Este resultado também confirmam que a premissa inicialmente proposta é válida: elaborar técnicas robustas na etapa de processamento da impressão digital de maneira que o rendimento do classificador seja maximizado.

6.2 Temas a serem pesquisados

Neste trabalho foi exposta uma técnica de classificação supervisionada, baseada na imagem de orientação como característica a utilizar para realizar a classificação. Na literatura tem-se um abordagem diferente baseado nas linhas de fluxo de cristas [21][49][27], estas técnicas de classificação tem como conveniência não requerer de dados de treino posto que a decisão da classificação está fundamentada nas formas geométricas das linhas de crista. Acreditamos que seria de interesse a realização de um classificador que aproveite estas características. Porém, o cálculo das linhas de crista deve ser realizado de maneira global, tal que permita obter a maior quantidade de informação útil para a classificação.

Validar a suposição de que os vetores de características vinculados às classes LL, W, RL e TA-A aparentemente apresentaram uma distribuição unimodal, através da utilização de outras distribuições unimodais. Se ela é comprovada, então aproveitar esta característica para enfrentar outros problemas vinculados à classificação de impressões digitais como é o problema das 8 classes, sendo este o esquema de classificação usado pelo FBI.

Utilizar outras técnicas de representação de dados e analisar suas consequências, entre elas temos KPCA (*Kernel Principal Component Analysis*), LPCA (*Local Principal Component Analysis*), ICA (*Independent Component Analysis*), LLA (*Locally Linear Embedding*) entre outros. Um dos pontos a considerar seria a procura de uma representação que afrente o problema da translação, dado que pela comparação dos experimentos realizados (automático e semi-automático) fica evidente a dependência da representação do PCA com a precisão na estimação do ponto de referência, o qual afeta de maneira direta a taxa de erro obtida pelo classificador.

Referências Bibliográficas

Referências Bibliográficas

- [1] ACHING, J. L., AND SALLES, E. O. T. Múltiplos níveis de representação da imagem de orientação na escala de cinza. Tech. rep., Laboratorio CISNE-UFES, ago 2007.
- [2] ACHING, J. L., AND SALLES, E. O. T. Um método de segmentação de impressões digitais de baixa qualidade. Tech. rep., Laboratorio CISNE-UFES, feb 2007.
- [3] BALLAN, M., SAKARYA, F. A., AND EVANS, B. L. A fingerprint classification technique using directional images. In *Conference on signals systems and computers* (1997).
- [4] BARTESAGHI, A., FERNÁNDEZ, A., AND GÓMEZ, A. Performance evaluation of an automatic fingerprint classification algorithm adapted to a vucetich based classification system. In *Conference on Audio and video-based biometric person authentication (3rd)* (2001), pp. 259–265.
- [5] BAZEN, A., AND GEREZ, S. Directional field computation for fingerprints based on the principal component analysis of local gradients. In *11th Annual Workshop on Circuits, Systems and Signal Processing, Veldhoven* (Netherlands, 2000).
- [6] BAZEN, A., AND GEREZ, S. Directional field computation for fingerprints based on the principal component analysis of local gradients. In *11th Annual Workshop on Circuits, Systems and Signal Processing* (2001).
- [7] BAZEN, A., AND GEREZ, S. Segmentation of fingerprint images. In *12th Annual Workshop on Circuits, Systems and Signal Processing* (2001).
- [8] BAZEN, A., AND GEREZ, S. Systematic methods for the computation of the directional fields and singular points of fingerprints. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, 7 (may 2002), 905–919.

- [9] BERNAND, S., BOUJEMAA, N., VITALE, D., AND BRICOT, C. fingerprint classification using kohonen topologic map. In *International Conference on Image Processing* (2001), vol. 3, pp. 230–233.
- [10] BLUE, J. L., CANDELA, G. T., GROTHOR, P. J., CHELLAPPA, R., AND WILSON, C. Evaluation of pattern classifiers for fingerprint and ocr applications. Tech. Rep. 3162, National Institute of Standards and Technology (NIST), sep 1993.
- [11] BOBUNG, S. *Evaluation of Transformation Methods to Detect Structures in Fingerprints*. Bachelor thesis, Technical University of Hamburg, Hamburg, 2002.
- [12] BOWEN, J. The home office automatic fingerprint pattern classification project. In *IEEE Colloquium on neural Networks for image processing Applications* (1992).
- [13] BYUNG-GYU, K., AND DONG-JO, P. Adaptive image normalisation based on block processing forenhancement of fingerprint image. *Electronics Letters* 38, 14 (Jul 2002).
- [14] CANDELA, G. T., AND CHELLAPPA, R. Comparative performance of classification methods for fingerprints. Tech. Rep. 5163, National Institute of Standards and Technology (NIST), Apr 1993.
- [15] CANDELA, G. T., GROTHOR, P. J., WATSON, C. I., WILKINSON, R. A., AND WILSON, C. L. Pcasys: A pattern-level classification automation system for fingerprints. Tech. Rep. 5647, National Institute of Standards and Technology (NIST) Intelligent Systems Division, Aug 1995.
- [16] CAPELLI, R., MAIO, D., AND MALTONI, D. Fingerprint classification based on multi-space kl. In *Workshop on Automatic Identification Advanced Technologies* (1999), pp. 117–120.
- [17] CAPELLI, R., MAIO, D., AND MALTONI, D. Combining fingerprint classifiers. In *International Workshop on Multiple Classifier Systems (1st)* (2000), pp. 351–361.
- [18] CAPELLI, R., MAIO, D., AND MALTONI, D. Multi-space kl for pattern representation and classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23, 9 (sept 2001), 977–996.
- [19] CAPELLI, R., MAIO, D., AND MALTONI, D. A multi-classifier approach to fingerprint classification. *IEEE Transactions on Pattern Analysis and Applications* 5, 2 (2002), 136–144.

- [20] CAPPELLI, R., LUMIMI, A., MAIO, D., AND MALTONI, D. Fingerprint classification by directional image partitioning. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21, 5 (may 1999), 402–421.
- [21] CHANG, J., AND FAN, K. A new model for fingerprint classification by ridge distribution sequences. *Pattern Recognition* 35, 6 (2002), 1209–1223.
- [22] CHEN, Y., DASS, S. C., AND JAIN, A. K. Fingerprint quality indices for predicting authentication performance. In *Audio- and Video-Based Biometric Person Authentication* (2005), pp. 160–169.
- [23] CHO, B. H., KIM, J. S., BAE, J. H., BAE, L. G., AND YOO, K. Y. Core based fingerprint image classification. In *International Conference on Pattern Recognition (15th)* (2000), pp. 863–866.
- [24] CHONG, M. M. S., NGEE, T. H., JUN, L., AND GAY, R. K. L. Geometric framework for fingerprint image classification. *Pattern Recognition* 30, 9 (1997), 1475–1488.
- [25] CORTIJO, F. J. *Estudio Comparativo de Métodos de Clasificación de Imágenes Multibanda*. Doutorado em informática, Universidade de Granada, Granada, jun 1995. 15 f.
- [26] COVER, T. M., AND HART, P. E. Nearest neighbor pattern classification. *Transactions on Information Theory* 13, 1 (jan 1967), 7–21.
- [27] DASS, S. C., AND JAIN, A. K. Fingerprint classification using orientation field flow curves. In *ICVGIP 2004, Proceedings of the Fourth Indian Conference on Computer Vision, Graphics Image Processing* (Kolkata, India, Dec 2004), pp. 650–655.
- [28] DAUGMAN, J. High confidence visual recognition of persons by a test of statistical independence. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15, 11 (1993), 1148–1161.
- [29] DERONG, Y., YUANYUAN, Z., AND DONGGUO, L. Fast computation of multiscale morphological operations for local contrast enhancement. In *27th Annual International Conference of the 2005* (2005), pp. 3090–3092.
- [30] DONAHUE, M. J., AND ROKHLIN, S. I. On the use of level curves in image analysis. *CVGIP: Image Understanding* 57, 2 (mar 1993), 185–203.

- [31] DUDA, R. O., HART, P. E., AND STORK, D. G. *Pattern Classification*. John Wiley and Sons, 2000.
- [32] FACON, J. *Morfologia matemática*. Universidade Católica do Paraná, 1996.
- [33] FITZ, A. P., AND GREEN, R. J. Fingerprint classification using hexagonal fast fourier transform. *Pattern Recognition* 29, 10 (1996), 1587–1597.
- [34] FIX, E., AND HODGES, J. L. Discriminatory analysis, non-parametric discrimination. Tech. Rep. 4, United States Air Force School of Aviation Medicine, Randolph Field, Texas, feb 1951.
- [35] GANCHEV, T. D., TASOULIS, D. K., VRAHATIS, M. N., AND FAKOTAKIS, N. D. Generalized locally recurrent probabilistic neural networks with application to text-independent speaker verification. *Neurocomputing* 70, 7-9 (2007), 1424–1438.
- [36] GANOTRA, D., JOSEPH, J., AND SINGH, K. Modified geometry of ring-wedge detector for sampling fourier transform of fingerprints for classification using neural networks. *Optics and Lasers in Engineering* 42, 2 (2004), 167–177.
- [37] GONZALES, R. C., AND WOODS, R. E. *Digital Image Processing*. Prentice Hall, 2002.
- [38] GREENBERG, S., ALADJEM, M., AND KOGAN, D. Fingerprint image enhancement using filtering techniques. *Real-Time Imaging* 8, 3 (2002), 227–236.
- [39] GREENBERG, S., AND KOGAN, D. Improved structure-adaptive anisotropic filter. *Pattern Recognition Letters* 27, 1 (jan 2006), 59–65.
- [40] GUANG, X., AND MILANFAR, P. Multiscale principal components analysis for image local orientation estimation. In *In Proceedings of the 36 Asilomar Conference on Signals Systems and Computers* (2002), pp. 478–482.
- [41] HALICI, U., AND ONGUN, G. Fingerprint classification through self-organizing feature maps modified to treat uncertainties. *Proceedings of the IEEE* 84, 10 (1996), 1497–512.
- [42] HAYKIN, S. *Redes Neurais: Princípios e prática*, segunda ed. Bookman, 2004.
- [43] HONG, L., AND JAIN, A. K. Classification of fingerprint images. In *Scandinavian Conference on Image Analysis (11th)* (1999).

- [44] HONG, L., WAN, Y., AND JAIN, A. K. Fingerprint image enhancement: Algorithm and performance evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20, 8 (aug 1998), 777–789.
- [45] HUGHES, P., AND GREEN, A. The use of neural networks for fingerprint classification. In *Conference on Neural Networks* (1991), vol. 2, pp. 79–81.
- [46] JAIN, A. K., DUIN, P. W., AND MAO, J. Statistical pattern recognition: a review. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22, 1 (2000).
- [47] JAIN, A. K., HONG, L., AND BOLLE, R. M. On-line fingerprint verification. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19, 4 (sep 1997), 302–314.
- [48] JAIN, A. K., HONG, L., PANKANTI, S., AND BOLLE, R. M. An identity-authentication system using fingerprints. *Proceedings of IEEE* 85, 9 (sep 1997), 1365–1388.
- [49] JAIN, A. K., AND MINUT, S. Hierarchical kernel fitting for fingerprint classification and alignment. In *International Conference on Pattern Recognition (16th)* (2002), vol. 2, pp. 469–473.
- [50] JAIN, A. K., PRABHAKAR, S., AND HONG, L. A multichannel approach to fingerprint classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21, 4 (1999).
- [51] JAIN, A. K., PRABHAKAR, S., HONG, L., AND PANKANTI, S. Filterbank-based fingerprint matching. *IEEE Transactions Image Processing* 9, 5 (2000), 846–859.
- [52] JIANG, X. On orientation and anisotropy estimation for online fingerprint authentication. *IEEE Transactions on Signal Processing* 53, 10 (Oct 2005), 4038–4049.
- [53] JOLLIFFE, I. T. *Principal Component Analysis*. Springer-Verlag, New York, 1986.
- [54] KAMIJO, M. Classifying fingerprint images using neural network: Deriving the classification state. In *Proceedings of 1993 IEEE International Conference on Neural Networks* (San Francisco, California, mar-apr 1993), vol. 3, IEEE/INNS, pp. 1932–1937.
- [55] KAMIJO, M., MIENO, H., AND KOJIMA, K. Classification of fingerprint images using a neural network. *Systems and Computers in Japan* 23 (1992), 89–101.

- [56] KARU, K., AND JAIN, A. K. Fingerprint classification. *Pattern Recognition* 29, 3 (1996), 389–404.
- [57] KASS, M., AND WITKIN, A. Analyzing oriented patterns. *Comput. Vision Graph. Image Process* 37, 3 (1987), 362–385.
- [58] KAWAGOE, M., AND TOJO, A. Fingerprint pattern classification. *Pattern Recognition* 17 (1984).
- [59] KOMARINSKY, P. *Automated Fingerprint Identification System (AFIS)*. Elsevier Academic Press, 2005.
- [60] LIU, M., JIANG, X., AND KOT, A. C. Fingerprint reference point detection. *Journal of Applied Signal Processing* 4 (2005), 498–509.
- [61] LOBATO, E., AND SEARA, R. Classificação e identificação de impressões digitais em tempo real de análise multi-resolução da informação direcional. In *SBrT 2000-XVIII Simpósio Brasileiro de Telecomunicações* (Gramado, Brasil, 2000), pp. 612–617.
- [62] LOBATO, E., SEARA, R., AND ROCHA, C. Identificação de impressões digitais via transformada wavelet. In *CBA 2000-XIII Congresso Brasileiro de Automática* (Florianópolis, Brasil, 2000), pp. 1415–1420.
- [63] LUMINI, A., MAIO, D., AND MALTONI, D. Continuous vs. exclusive classification for fingerprint retrieval. *Pattern Recognition* 18, 10 (1997).
- [64] LUMINI, A., MAIO, D., AND MALTONI, D. Inexact graph matching for fingerprint classification. *Machine Graphics & Vision (Special Issue on Graph Transformations in Pattern Generation and CAD)* 8, 2 (1999).
- [65] MAIO, D., AND MALTONI, D. Structural approach to fingerprint classification. In *International Conference on Pattern Recognition (13th)* (San Francisco, California, 1996), vol. 3, IEEE/INNS, pp. 578–585.
- [66] MAIO, D., AND MALTONI, D. Direct gray-scale minutiae detection in fingerprints. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19, 1 (1997).
- [67] MALTONI, D., MAIO, D., JAIN, A. K., AND PRABHAKAR, S. *Handbook of fingerprint recognition*. Springer-Verlag, 2003.
- [68] MAO, K. Z., TAN, K. C., AND SER, W. Probabilistic neural-network structure determination for patternclassification. *IEEE Transactions on Neural Networks* 11 (2000), 1009–1016.

- [69] MARCIALES, G. L., ROLI, F., AND FRASCONI, P. Fingerprint classification by combination of flat and structural approaches. In *International Conference On Audio and Video-Based Biometric Person Authentication (3rd)* (2001), pp. 241–246.
- [70] MARQUES, A. C. P. B., AND THOMÉ, A. C. G. A neural network fingerprint segmentation method. In *5th International Conference on Hybrid Intelligent Systems* (Rio de Janeiro, Brazil, Nov 2005), pp. 385–392.
- [71] MATHWORKS. *Image Processing Toolbox 5, User's Guide*. 2007.
- [72] MEHTRE, B. M., AND CHATTERJEE, B. Segmentation of fingerprint images a composite method. *Pattern Recognition* 22, 4 (1989).
- [73] MEHTRE, B. M., MURTHY, N. N., KAPOOR, S., AND CHATTERJEE, B. Segmentation of fingerprint images using the directional image. *Pattern Recognition* 20 (2000), 429–435.
- [74] MOAYER, B., AND FU, K. An application of stochastic languages to fingerprint pattern recognition. *Pattern Recognition* 8, 3 (1976).
- [75] MOAYER, B., AND FU, K. A tree system approach for fingerprint pattern recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8, 3 (1986).
- [76] MOAYER, B., AND FU, K. S. A syntactic approach to fingerprint pattern recognition. *Pattern Recognition* 7, 1-2 (jun 1975).
- [77] MOSCINSKA, K., AND TYMA, G. Neural network based fingerprint classification. In *Conference on Artificial Neural Networks* (1993), no. 2.
- [78] MUKHOPADHYAY, S., AND CHANDA, B. A multiscale morphological approach to local contrast enhancement. *Signal Processing* 80, 4 (Abr 2000).
- [79] OMIDVAR, O. M., BLUE, J. L., AND WILSON, C. L. Improving neural network performance for character and fingerprint classification by altering network dynamics. In *World Congress on Neural Networks* (1995), pp. 151–158.
- [80] PARK, C., AND PARK, H. Fingerprint classification using fast fourier transform and nonlinear discriminant analysis. *Pattern Recognition* 38, 4 (Abr 2005), 495–503.
- [81] PARK, C.-H., OH, S.-K., KWAK, D.-M., KIM, B.-S., SONG, Y.-C., AND PARK, K.-H. A new reference point detection algorithm

- based on orientation pattern labeling in fingerprint images. In *Pattern Recognition and Image Analysis, First Iberian Conference, IbPRIA* (Puerto de Andratx, Mallorca, Spain, June 4-6 2003), vol. 2652 of *Lecture Notes in Computer Science*, pp. 697–703.
- [82] PATTICHIS, M. S., PANAYI, G., BOVIK, A. C., AND HSU, S. P. Fingerprint classification using an am-fm model. *IEEE Transactions on Image Processing* 10, 6 (2001).
- [83] POURGHASSEM, H., AND GHASSEMIAN, H. Fingerprint classification with probabilistic neural network. In *Signal Processing and Its Applications, 2005. Proceedings of the Eighth International Symposium* (Aug 28-31 2005), vol. 2, pp. 663–666.
- [84] RAMØ, P., TICO, M., ONNIA, V., SAARINEN, J., AND KORKEAKOULU, T. T. Optimized singular point detection algorithm for fingerprint images. In *International Conference on Image Processing (ICIP01)* (Greece, 2001), pp. 242–245.
- [85] RAO, C. V. K., AND BALCK, K. Type classification of fingerprint: A syntactic approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2, 3 (1980).
- [86] RATHA, N., CHEN, S., AND JAIN, A. K. Adaptive flow orientation-based feature extraction in fingerprint images. *Pattern Recognition* 28, 11 (Nov 1995).
- [87] RATHA, N. K., AND BOLLE, R. *Automatic Fingerprint Recognition Systems*. SpringerVerlag, 2003.
- [88] ROJAS, D. A., AND ACHING, J. L. An improved method for segmentation of fingerprint images. In *Electronics, Robotics and Automotive Mechanics Conference* (2006), pp. 20–23.
- [89] SENIOR, A. A hidden markov model fingerprint classifier. In *Asilomar Conf. On Signal Systems and Computers (31st)* (1997), pp. 306–310.
- [90] SENIOR, A. A combination fingerprint classifier. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23, 10 (2001).
- [91] SHI, Z., WANG, Y., QI, J., AND XU, K. A new segmentation algorithm for low quality fingerprint image. In *Third International Conference on Image and Graphics* (2004), pp. 314–317.
- [92] SPECHT, D. F. Probabilistic neural networks. *Neural Networks* 3, 10 (1990), 109–118.

- [93] TAHMASEBI, A. M., AND KASAEI, S. A novel adaptive approach to fingerprint enhancement filter design. *Signal Processing:Image Communication* 17, 10 (nov 2002), 849–855.
- [94] TAN, X., BHANU, B., AND LIN, Y. Fingerprint classification based on learned features. *IEEE Transactions on Systems, Man, and Cybernetics* 35, 3 (2005), 287–300.
- [95] TICO, M., AND KUOSMANEN, P. Fingerprint matching using an orientation-based minutia descriptor. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25, 8 (2003), 1009–1014.
- [96] WANG, S., AND WANG, Y. Fingerprint enhancement in the singular point area. *IEEE Signal Processing Letters* 11, 1 (jan 2004), 823–833.
- [97] WANG, Y., HU, J., AND HAN, F. Enhanced gradient-based algorithm for the estimation of fingerprint orientation fields. *Applied Mathematics and Computation* 185, 2 (2007), 823–833.
- [98] WATSON, C. I., CANDELA, G. I., AND GROTH, P. J. Comparison of FFT fingerprint filtering methods for neural network classification. Tech. Rep. 5493, National Institute of Standards and Technology (NIST) Intelligent Systems Division, 1994.
- [99] WATSON, C. I., AND WILSON, C. L. Nist special database 4. Tech. rep., National Institute of Standards and Technology (NIST), 1992.
- [100] WEI, D., YUAN, Q., AND JIE, T. Fingerprint classification system with feedback mechanism based on genetic algorithm. In *International Conference on Pattern Recognition (14th)* (1998), vol. 1, pp. 163–165.
- [101] WILLIS, A. J., AND MYERS, L. A cost-effective fingerprint recognition system for use with low-quality prints and damaged fingertips. *Pattern Recognition* 34, 2 (2001), 255–270.
- [102] WILSON, C. L., CANDELA, G. T., GROTH, P. J., WATSON, C. I., AND WILKINSON, R. A. Massively parallel network fingerprint classification system. Tech. Rep. 4880, National Institute of Standards and Technology (NIST) Intelligent Systems Division, Oct 1992.
- [103] WILSON, C. L., CANDELA, G. T., AND WATSON, C. I. Neural network fingerprint classification. *Journal Artificial Neural Networks* 1, 2 (1994).

- [104] YAGER, N., AND AMIN, A. Evaluation of fingerprint orientation field registration algorithms. In *17th International Conference On Pattern Recognition (2004)*, vol. 4, pp. 641–644.
- [105] YAO, Y., FRASCONI, P., AND PONTIL, M. Fingerprint classification with combinations of support vector machines. In *International Conference On Audio and Video-Based Biometric Person Authentication (3rd)* (2001), vol. 2091, pp. 253–258.
- [106] ZHANG, Q., AND YAN, H. Fingerprint classification based on extraction and analysis of singularities and pseudo ridges. *Pattern Recognition* 37, 11 (2004).
- [107] ZHU, E., YIN, J., HU, C., AND ZHANG, G. A systematic method for fingerprint ridge orientation estimation and image segmentation. *Pattern Recognition* 39, 8 (Ago 2006).