

Universidade Estadual de Campinas

Faculdade de Engenharia Elétrica

Departamento de Comunicações

Sistema Automático de Reconhecimento de Impressões Digitas

Por: Eng. Gean Davis Breda

Orientador: Prof. Dr. Lee Luan Ling

Dissertação submetida à Faculdade de Engenharia Elétrica da Universidade Estadual de Campinas, para o preenchimento dos pré-requisitos parciais para a obtenção do Título de Mestre em Engenharia Elétrica.

Este exemplar foi entregue ao autor final da tese defendida por *Gean Davis Breda*

Julgadora em *27/02/1996*

Lee Luan Ling

Fevereiro 1996

16 09208

UNIDADE	BC
N.º CHAMADA	
UNICAMP	
B743s	
V.º	
F.º	27870
D.º	667196
	[X]
VALOR	R\$ 14,00
DATA	03/07/96
N.º CPD	

CM-00089590-1

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

B743s Breda, Gean Davis
Sistema automático de reconhecimento de impressões
digitais / Gean Davis Breda.--Campinas, SP: [s.n.], 1996.

Orientador: Lee Luan Ling.
Dissertação (mestrado) - Universidade Estadual de
Campinas, Faculdade de Engenharia Elétrica.

1. Reconhecimento de padrões. 2. Redes neurais
(Computação). 3. Processamento de imagens.
4. * Impressões digitais. I. Lee, Luan Ling.
II. Universidade Estadual de Campinas. Faculdade de
Engenharia Elétrica. III. Título.

Aos meus pais, que sempre me apoiaram nas decisões que tomei.

Não existo para impressionar o mundo. Existo para viver minha vida de um modo que me faça feliz.

Agradecimentos

Em primeiro lugar, agradeço a minha família. Ao meu pai que continua dando exemplo de luta, honestidade, a minha mãe um referencial de força de vontade, amor. Aos meus irmãos pelo carinho e apoio. Aos meus tios, tias, primos e primas pelo carinho.

À Deus, meu pai supremo, fonte inesgotável de energia.

Agradeço à Karina esta pessoa tão especial e a sua família pelo carinho com que me acolheram, tornando-me parte de suas vidas.

Aos amigos Paulo R. Finger, Paulo e Tânia Zamboni, pessoas que fazem com que a vida tenha mais sentido.

Ao pessoal do NEPAE, principalmente o Nelson, pelas lições valiosas que me ajudaram a crescer como ser humano.

Ao Professor Lee Luan Ling pela orientação e ao Professor Renato Baldini Filho pela co-orientação.

Ao Marcelo a quem considero como um irmão, à Fabrícia pela sua compreensão e carinho, e aos amigos Jackson, Tina, Sandro, Natanael, Marcos, Marta, Ernesto, Luciano, Leonardo, Paulo, Antônio, Gilmar, Geraldo, Cristiana, Alexandre, Edgard, Rodrigo, Mario, Miguel, Jorge, Osmar, Façonny, Rivelino.

Àquelas pessoas que contribuem tanto, que não terão os seus nomes escritos em uma página de livro. Mas sim, nas reminiscências de uma vida.

À Universidade Estadual de Campinas pela oportunidade.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES pelo apoio financeiro.

Somos criaturas que brincam, que se divertem, somos as lontras do universo. Não podemos morrer, não podemos nos ferir mais do que podem ferir as ilusões da tela. Mas, podemos acreditar que estamos feridos, com todos os detalhes agonizantes que quisermos. Podemos acreditar que somos vítimas, mortas e matando, envolvidas pela boa e má sorte.

Sumário

AGRADECIMENTOS	iii
SUMÁRIO	v
LISTA DE FIGURAS	viii
LISTA DE TABELAS	xi
RESUMO	xii
ABSTRACT	xiii
1 Introdução	1
1.1 Estrutura do Trabalho.....	2
2 Aquisição de Imagem	3
2.1 Introdução.....	3
2.2 Elementos de Aquisição de Imagem.....	3
2.2.1 Scanner.....	3
2.2.2 Amostragem e Quantização.....	5
2.3 Métodos de Aquisição.....	9
2.3.1 Método do Papel e Tinta.....	10
2.3.2 Método do Prisma.....	10
3 Processamento de Imagem	17
3.1 Introdução.....	17
3.2 Histograma.....	17
3.2.1 Equalização do histograma.....	18
3.3 Inversão.....	23
3.4 Limiar.....	24

3.4.1	Limiar Global.....	25
3.4.2	Limiar Dinâmico.....	26
3.5	Morfologia Matemática.....	27
3.5.1	Morfologia Matemática em Imagens Binárias.....	28
3.5.2	Morfologia Matemática em Imagens com Níveis de Cinza.....	34
3.6	Algoritmo de Processamento.....	40
3.6.1	Equalização.....	41
3.6.2	Limiar.....	41
3.6.3	Esqueleto.....	46
4	Extração de Características	51
4.1	Introdução.....	51
4.2	Padrões.....	51
4.3	Características.....	51
4.4	Seleção de Características.....	52
4.5	Vetor Distância.....	54
4.6	Vetor Histograma.....	58
4.7	Quadrados Concêntricos.....	58
5	Reconhecimento de Padrões	59
5.1	Introdução.....	59
5.2	Redes Neurais.....	59
5.2.1	Back-Propagation.....	60
5.3	Variações sobre Back-Propagation.....	64
5.3.1	Momento.....	65
5.3.2	Parâmetros Adaptativos.....	66
5.4	Correlação.....	66

5.5 Reconhecimento por Redes Neurais.....	67
5.6 Reconhecimento por Correlação.....	79
5.7 Comentário dos Resultados.....	80
6 Conclusões	82
A Filtro de Mediana	84
B Parâmetros Estatísticos	85
Bibliografia	87

Listas de Figuras

2.1	Princípio de aquisição.....	4
2.2	Layout e funcionamento do scanner.....	5
2.3	Imagens de impressões digitais adquiridas.....	7
2.4	Relevo ideal de uma impressão digital.....	8
2.5	Relevo de uma impressão digital sob pressão.....	9
2.6	Sequência do processo de aquisição.....	10
2.7	Perpectivas do prisma utilizado.....	12
2.8	Refração e reflexão.....	13
2.9	Impressão digital obtida sem o prisma.....	14
2.10	Distância focal.....	15
2.11	Linhas da impressão digital.....	16
3.1	Histogramas de quatro imagens.....	18
3.2	Função de transformação.....	21
3.3	Exemplo de equalização do histograma.....	23
3.4	Histogramas.....	25
3.5	Limiar global.....	26
3.6	Formato da imagem.....	27
3.7	Limiar dinâmico.....	27
3.8	Operações básicas.....	29
3.9	Dilatação.....	29
3.10	Erosão.....	30
3.11	Abertura.....	31

3.12	Fechamento.....	32
3.13	Transformada em tudo ou nada.....	33
3.14	Afinamento.....	33
3.15	Dilatação em níveis de cinza.....	35
3.16	Erosão em níveis de cinza.....	36
3.17	Abertura em níveis de cinza.....	38
3.18	Fechamento em níveis de cinza.....	38
3.19	Top-hat.....	39
3.20	Exemplo de minúcias.....	40
3.21	Esqueleto e equalização.....	42
3.22	Limiar global.....	43
3.23	Limiar dinâmico.....	44
3.24	Relevo de uma pequena porção da impressão digital.....	45
3.25	Top-hat.....	46
3.26	Esqueleto com ramos parasitas.....	48
3.27	Falsos ramos parasitas.....	49
3.28	Resultado sinal.....	50
3.29	Cadeia de processamento.....	50
4.1	Minúcias.....	53
4.2	Novas minúcias.....	54
4.3	Núcleo.....	54
4.4	Exemplo dos valores de colunas da digital.....	55
4.5	Vetores característicos.....	56
4.6	Vetores característicos filtrados.....	57
4.7	Regiões de interesse.....	57

4.8	Quadrados concêntricos.....	58
5.1	Redes de duas camadas com alimentação direta.....	61
5.2	Back-propagation.....	63
5.3	Exemplo de correlação.....	67
5.4	Exemplo de vetores característicos.....	69
5.5	Gráficos do erro quadrático médio e taxa de aprendizado.....	69
5.6	Gráficos do erro quadrático médio e taxa de aprendizado.....	70
5.7	Gráficos do erro quadrático médio e taxa de aprendizado.....	71
5.8	Gráficos do erro quadrático médio e taxa de aprendizado.....	73
5.9	Gráficos do erro quadrático médio e taxa de aprendizado.....	75
5.10	Gráficos do erro quadrático médio e taxa de aprendizado.....	76
5.11	Gráficos do erro quadrático médio e taxa de aprendizado.....	77
5.12	Gráficos do erro quadrático médio e taxa de aprendizado.....	78
5.13	Resultados.....	81

Lista de Tabelas

3.1	Dois pares de elementos estruturantes.....	47
3.2	Elementos estruturantes para ramos parasitas.....	48
3.3	Elementos estruturantes para pontos isolados.....	49
5.1	Erros cometidos.....	70
5.2	Erros cometidos.....	71
5.3	Erros cometidos.....	72
5.4	Ativação das saídas(neurônios).....	72
5.5	Erros cometidos.....	74
5.6	Erros cometidos.....	74
5.7	Erros cometidos.....	75
5.8	Erros cometidos.....	76
5.9	Erros cometidos.....	77
5.10	Erros cometidos.....	78
5.11	Desempenho do classificador por correlação.....	79
5.12	Erros cometidos.....	80

Resumo

Neste trabalho desenvolvemos um sistema automático de reconhecimento de impressões digitais. Este sistema inclui aquisição da imagem, processamento, extração de características e classificação dos padrões. Na aquisição da imagem de impressão digital procuramos estudar os métodos que estão disponíveis e escolher o mais adequado ao nosso propósito. Introduzimos melhoras no método escolhido, o que levou a obtenção de bons resultados. A cadeia de processamento de imagem de impressões digitais é uma contribuição significativa deste trabalho, baseada em morfologia matemática, é desenvolvida especificamente com o objetivo de ser empregada na obtenção do esqueleto de uma imagem de impressão digital. Na extração de características selecionamos atributos para o sistema de classificação. Estes atributos são escolhidos de forma a serem os mais adequados para a tarefa de reconhecimento. Foram extraídos três tipos de características, denominadas: vetor histograma, vetor distância e quadrados concêntricos. Na classificação utilizamos dois tipos de classificadores: redes neurais Back-Propagation e um classificador por correlação. Os melhores resultados na classificação são obtidos utilizando-se o atributo, denominado vetor histograma, conjuntamente com uma rede neural Back-propagation.

Abstract

In this work we develop an automatic system of fingerprint recognition. This system includes image acquisition, processing, extraction of features and pattern classification. In the image acquisition of fingerprint, we study the methods already available and choose the more adequate for our application. We improve the method chosen, that improvement leads to obtain good results. The image processing network, which is based on Mathematical Morphology, is a main contribution of this work. It is developed with the objective of obtaining the skeleton of a fingerprint. In the feature extraction it is chosen the characteristics for the classification system. These characteristics are chosen to fit the task of recognition. The following three kinds of features are utilized: histogram vector, length vector and concentric squares. In the classification we utilize two kinds of classifiers: Back-Propagation neural network and correlation classifier. The best result in the classification is obtained when we utilize the histogram vector with Back-Propagation neural network.

Capítulo 1

Introdução

Identificar uma pessoa parece uma tarefa simples. Pessoas realizam isto todo o tempo no trabalho ou em encontros sociais. Mas a sociedade moderna tem complicado o processo de identificação, por exemplo, podemos ter pessoas tentando se fazer passar por outras das mais diferentes maneiras. Nem todos os métodos de identificação disponíveis são suficientemente seguros. Isto faz com que a pesquisa sobre técnicas de identificação torne-se mais intensa do que nunca.

Uma área onde a tecnologia está melhorando e geralmente simplificando nossa habilidade para reconhecer pessoas é a biométrica. Sistemas biométricos são métodos automatizados de verificação e reconhecimento da identidade de uma pessoa baseando-se em características fisiológicas, como impressões digitais, íris, ou algum aspecto do comportamento, como a dinâmica da assinatura.

Comparação por impressões digitais é um método fundamental na identificação pessoal. A razão para isto é bem conhecida. A impressão digital humana é um dos mais confiáveis meios para se identificar uma pessoa, devido ao fato de que não existe duas impressões digitais iguais de pessoas diferentes.

Através dos anos a polícia foi conhecida como o maior usuário de impressões digitais com o objetivo de identificar criminosos. Os mais complexos e caros sistema já construídos, como enormes bancos de dados de mais de um milhão de imagens armazenadas, são utilizados pela polícia. Devido ao grande volume de impressões e recentes avanços na tecnologia digital, tem havido um crescente interesse em automatizar este sistema de reconhecimento. O Bureau de Investigações Federais dos E.U.A. possui um sistema automatizado de identificações de digitais que armazena cerca de 40 terabytes de dados. Outro exemplo de aplicação que também seria beneficiada com a automatização deste processo é acesso automático.

Muitas técnicas para pré-processamento e reconhecimento de impressões digitais foram propostas; por exemplo, algoritmo de pré-processamento para classificação de impressões digitais, implementado por Ch e Rao [1]; um método sintático para

representar e classificar imagens de impressões digitais, Moayer e Fu[2] e um algoritmo para detecção de sulcos investigado por Verma[3].

Um dos maiores problemas para o reconhecimento automático de impressões digitais é a qualidade da imagem original. Se a qualidade não está em padrões aceitáveis, o reconhecimento automático torna-se extremamente difícil. A razão para isto é que os métodos normais de reconhecimento utilizam-se de pequenas características (conhecidas como minúcias[4]) como padrões para o reconhecimento.

1.1 Estrutura do Trabalho

O objetivo deste trabalho é desenvolver um sistema automático de reconhecimento de impressões digitais. Este sistema executa os seguintes processos: aquisição da imagem, processamento, extração de características e classificação.

No Capítulo 2, inicialmente são introduzidos conceitos físicos de aquisição de imagem, após é feita uma comparação entre dois métodos de aquisição de imagens de impressões digitais. Em seguida, descreve-se o processo de aquisição de imagens proposto neste trabalho.

O Capítulo 3 é dedicado a apresentação das ferramentas usadas no processamento de imagens. É montada uma cadeia de processamento que tem como objetivo melhorar a qualidade da imagem de impressões digitais.

No Capítulo 4 descrevemos o processo de extração de características das impressões digitais.

O capítulo 5 descreve os processos de classificação propostos. Dois tipos de classificadores são utilizados: classificadores do tipo rede neural e classificador por correlação. Em todos os experimentos analisamos a performance do sistema. No capítulo 6 apresentamos as conclusões do nosso trabalho.

Capítulo 2

Aquisição de Imagem

2.1 Introdução

Neste capítulo daremos uma visão geral do princípio de aquisição de imagem. Mostraremos o caminho percorrido por uma imagem desde sua captura até a chegada ao computador, definiremos as variáveis importantes no processo de aquisição e, por último, comparamos dois métodos utilizados na captura de imagens de impressões digitais, analisando suas vantagens e limitações.

2.2 Elementos de Aquisição de Imagem

Existem diversas maneiras de captarmos uma imagem. Podemos, por exemplo, utilizar uma câmera de vídeo, uma câmera CCD ou um scanner de mesa. Na figura 2.1 é mostrado o princípio geral de aquisição. Neste trabalho estamos utilizando um scanner de mesa, razão pela qual daremos maior atenção a este instrumento de aquisição.

2.2.1 Scanner

No processo de aquisição de imagens, existem dois elementos fundamentais. O primeiro é um componente físico que deve ser sensível à banda no espectro eletromagnético escolhida (tal como raio-x, ultravioleta, luz visível), o qual produz sinais elétricos em sua saída proporcionais aos níveis de energia medidos. O segundo, chamado digitalizador, tem como objetivo converter os sinais elétricos provenientes do componente físico numa forma digital, para que estes possam ser processados por computador.

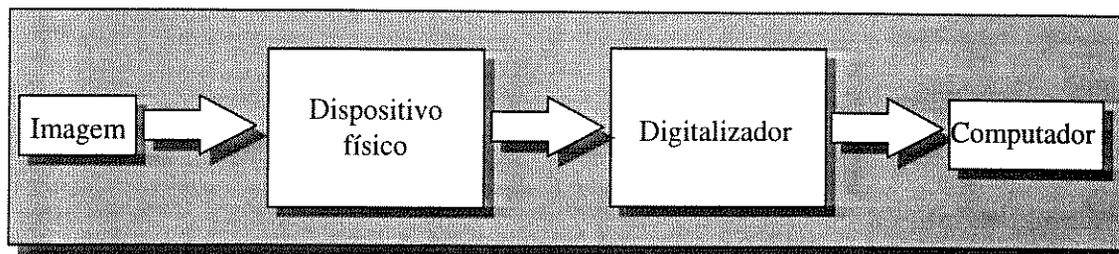


Figura 2.1- Princípio de aquisição

Scanners são equipamentos capazes de adquirir imagens. Seu princípio de funcionamento é emitir luz sobre a imagem a ser adquirida; uma parte desta luz será absorvida e a outra parte refletida sendo posteriormente captada pelo scanner (veja figura 2.2). Esses equipamentos possuem internamente um componente sensível à luz visível, bem como um digitalizador. Um exemplo de componente sensível à luz são células photo-voltáicas que produzem tensão em sua saída proporcional ao nível de intensidade de luz incidente sobre elas. Fisicamente estas células podem ser organizadas em dois modos: uma linha de sensores ou uma área de sensores. Uma linha de sensores consiste em uma linha de células photo-voltáicas, ou melhor, um array de células unidimensionais. Conseguimos produzir uma imagem bi-dimensional pelo movimento relativo entre a imagem a ser captada e a linha de sensores. Uma aplicação deste princípio são os scanners de mesa. Uma área de sensores é composta por uma matriz de células photo-voltáicas, array bi-dimensional, capaz, portanto, de capturar imagens em duas dimensões, sem a necessidade de movimento dos sensores.

Após isto, passamos o sinal resultante por um conversor analógico-digital. Este conversor analógico-digital, nada mais é do que um digitalizador, que transforma sinais analógicos provenientes dos sensores óticos em sinais digitais.

Neste trabalho utilizamos um scanner Microtek ScanMaker IISP, que possui resolução entre 100 d.p.i (pontos por polegada) e 1200 d.p.i com possibilidade de atingir até 256 níveis de cinza.

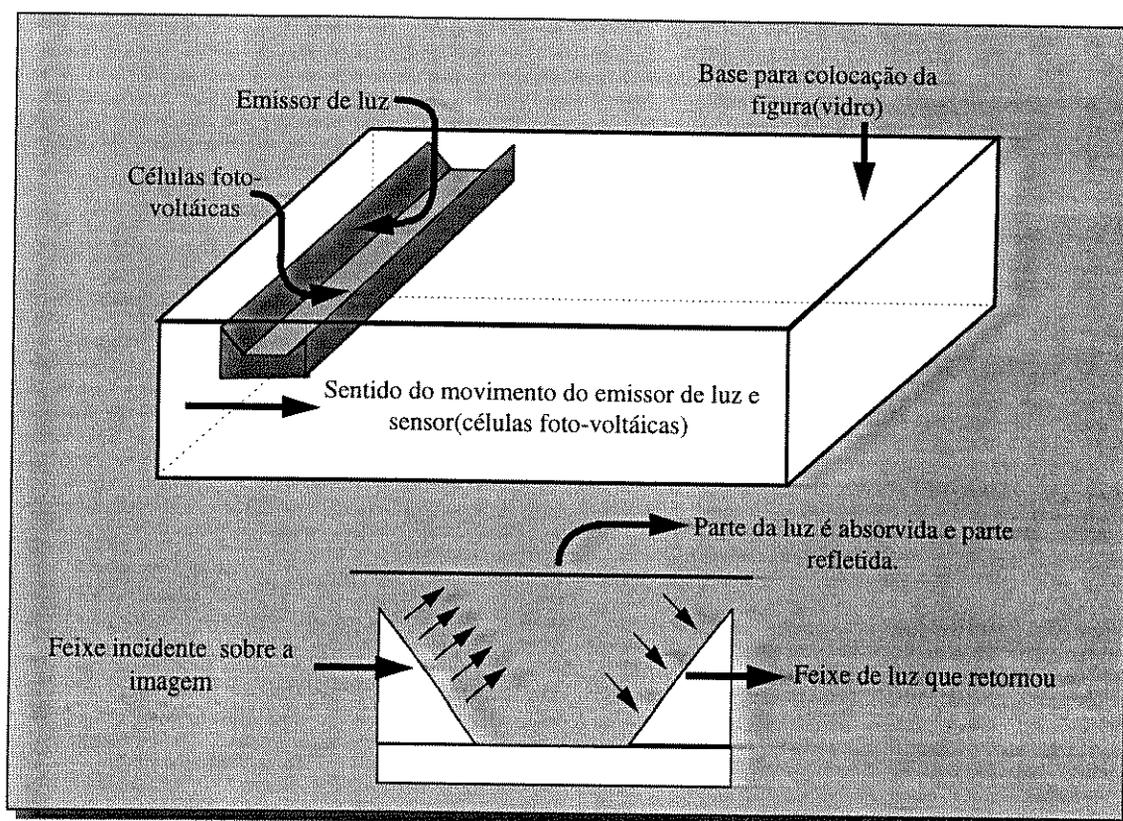


Fig 2.2- Layout e funcionamento do scanner

2.2.2 Amostragem e Quantização

Para uma imagem estar adequada ao uso em um computador, a imagem precisa ser digitalizada[14] tanto no domínio espacial como em amplitude. Digitalização das coordenadas espaciais é chamada de amostragem da imagem, e a digitalização da amplitude é chamada de quantização em níveis de cinza. Suponha que uma imagem contínua $g(x, y)$ é amostrada igualmente no espaço e organizada na forma de uma array de $N \times M$ elementos, onde cada elemento do array é uma quantidade discreta, representando uma pequena área da imagem. Esta representação é o que define uma imagem digital. Cada elemento do array é referido como um elemento da imagem, elemento da figura, pixel ou pel. Os termos imagem e pixels serão muito utilizados nas discussões seguintes para denotar uma imagem digital e seus elementos.

A seguir, definiremos matematicamente o processo de amostragem e quantização. Sejam Z e R o conjunto dos números reais e o conjunto dos números inteiros, respectivamente. O processo de amostragem pode ser visto como uma partição do plano $X \times Y$ por uma grade, com as coordenadas do centro de cada elemento da

grade caracterizando um elemento do produto cartesiano $Z \times Z$, que é o conjunto de todos os pares ordenados de elementos (a, b) , com a e b sendo inteiros de Z . Consequentemente $g(x, y)$ é uma imagem digital se x e y assumem valores inteiros e se g é uma função que designa um valor em nível de cinza, isto é, um número, para cada distinto par de coordenadas (x, y) . Esta função de designação de um valor de cinza é chamada processo de quantização.

Este processo de digitalização requer que certas decisões sejam tomadas sobre os valores das dimensões da imagem e o número de níveis discretos de cinza. É usual adotar valores de potências de dois, isto é:

$$N = 2^n, M = 2^k \text{ e } G = 2^m \quad (2.1)$$

onde G define o número de níveis de cinza, N e M são respectivamente o número de linhas e colunas da matriz(imagem). Assim, a memória necessária para armazenar a imagem, expressa em termos de número de bytes, será:

$$b = N * M * m \quad (2.2)$$

Por exemplo, para uma imagem de tamanho 128×128 com 64 níveis de cinza é requerido 98304 bytes de memória.

Uma pergunta fundamental que deve ser respondida em qualquer processo de digitalização é: Quantas amostras e quantos níveis de cinza são necessários para que a imagem digital seja uma boa aproximação da imagem original.

A resolução, ou melhor, o grau de reprodução dos detalhes depende quase exclusivamente destes três parâmetros. Quanto maiores são estes parâmetros mais fiel a imagem digital será da imagem original. Cabe citar que quanto maiores tais parâmetros, mais dispêndio no armazenamento e consequentemente, as variáveis do processamento crescem rapidamente em função de N , M e G .

É difícil definir o que é uma boa imagem. Ela não se caracteriza apenas por ser altamente objetiva, mas também deve preencher os requisitos da aplicação a que será submetida.

Nas figura 2.3(a)-(d) podemos ver como uma imagem varia com diferentes níveis de resolução (N e M) mantendo-se G constante e igual a 256. As figuras 2.3(a), (b), (c) e (d) mostram as imagens de uma impressão digital com 150, 250, 400 e 600 d.p.i, respectivamente. Na figura 2.3(e) temos uma imagem de impressão digital com apenas dois níveis de cinza ($G = 2$) e 600 d.p.i de resolução.

Estas imagens ilustram como os valores de N , M e G afetam a qualidade da imagem. É obvio, que a imagem de melhor qualidade é aquela com 600 d.p.i. . Em relação aos níveis de cinza também não foi observado qualquer resultado atípico, ou melhor, quanto maior a quantidade de níveis de cinza, melhor a qualidade da imagem. É importante lembrar que o nível de fidelidade é altamente subjetivo e depende da aplicação. Por isso, também é objetivo deste trabalho determinar a resolução da imagem adequada para nosso trabalho.

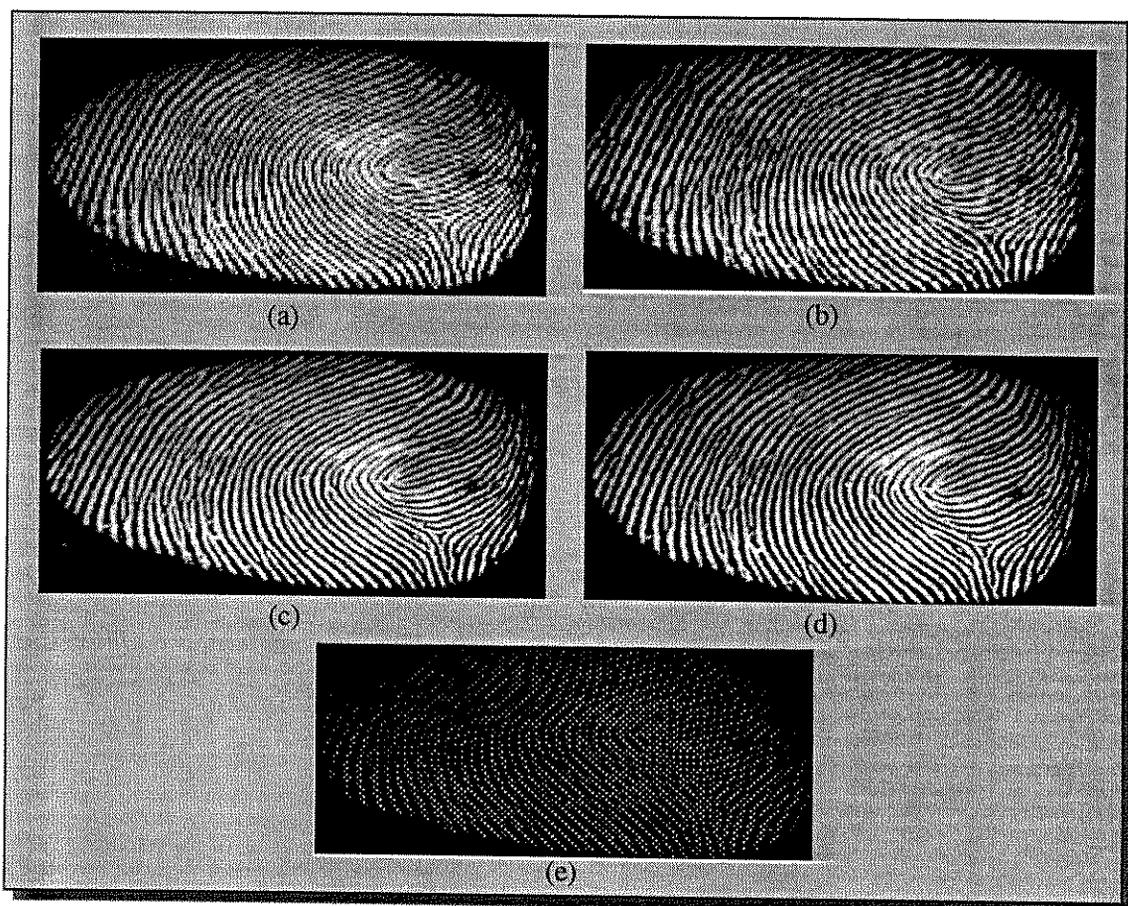


Figura 2.3- Imagens de impressões digitais adquiridas

Na figura 2.4 temos uma forma topológica idealizada do relevo de uma impressão digital, onde introduzimos algumas características importantes da impressão

digital, tais como, montanhas e vales. Nesta representação retratamos apenas o formato de duas montanhas e três vales numa mesma direção, mas é importante lembrar que uma impressão digital é composta por milhares de sulcos em várias direções. A distância entre dois sulcos consecutivos possui um valor médio de aproximadamente 0,5 mm.

Nosso método de aquisição capta uma imagem de uma impressão digital no momento em que o dedo exerce certa pressão sobre uma superfície. A medida que pressionamos a impressão, suas características vão perdendo o formato original, isto é, aumenta-se a superfície dos sulcos e diminui-se a largura dos vales, como ilustrado na figura 2.5 .

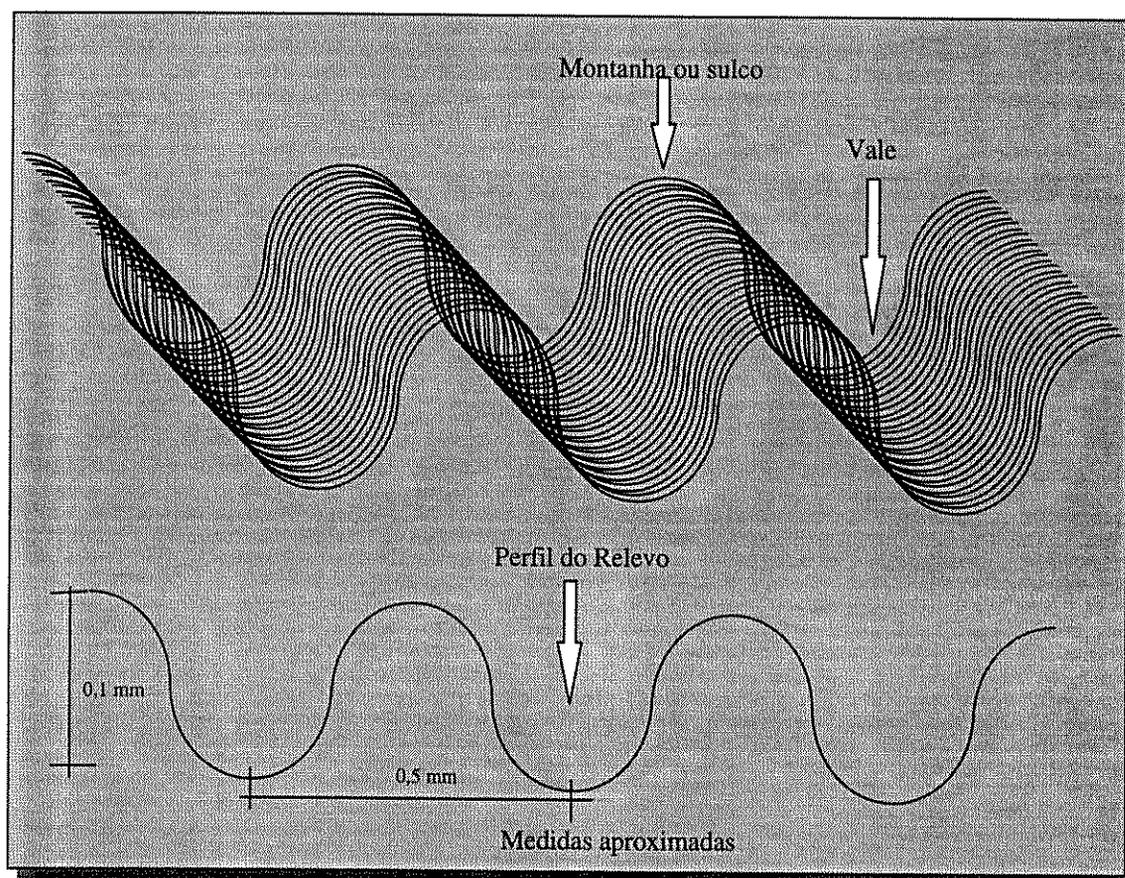


Figura 2.4- Relevo ideal de uma impressão digital

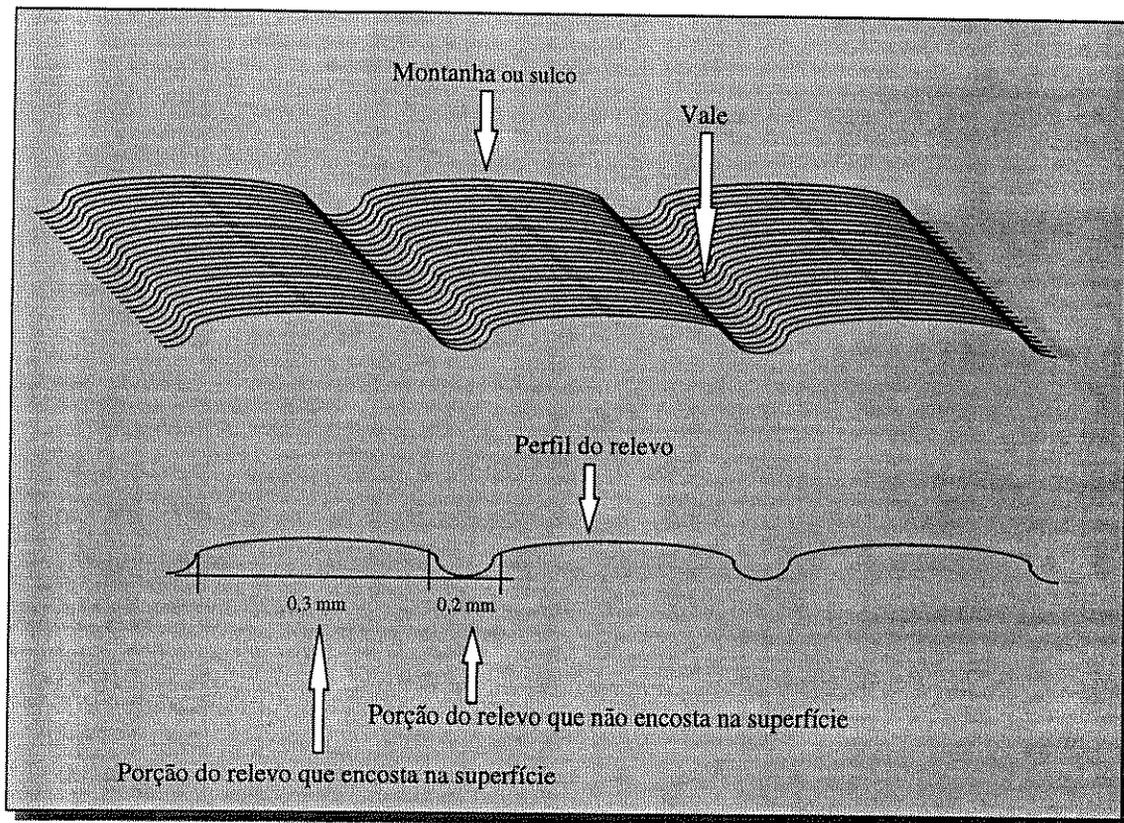


Figura 2.5- Relevo de uma impressão digital sob pressão

Muitas vezes, dependendo da pressão com que a digital está sendo comprimida, ocorrem distorções na imagem obtida. Por exemplo, se a digital é pressionada com muita pressão, pode fazer com que vales desapareçam entre sulcos, formando portanto uma área plana. Por outro lado, pouca pressão pode fazer com que o sensor ótico não detecte alguns sulcos, unindo portanto dois ou mais vales.

Verificamos que a distância média útil entre sulcos é de aproximadamente 0,2 mm. Em nossa cadeia de processamento de imagens utilizamos certas operações que exigem uma quantidade média de 5 pixels entre sulcos, logo:

$$5 \text{ pixels por } 0,2 \text{ mm} = 25 \text{ pixels por } 1 \text{ mm}$$

$$25 \text{ pixels por } 1 \text{ mm} \cong 630 \text{ pixels por polegada(d.p.i)}$$

Adotamos então um valor aproximado de 600 d.p.i, com 256 níveis de cinza.

2.3 Métodos de Aquisição

Uma vez analisado o equipamento de aquisição e suas variáveis, devemos agora estudar os métodos de aquisição. Estes métodos procuram, em suma, melhorar o

contraste da impressão digital no momento da aquisição. Nós investigaremos dois métodos de aquisição[6], métodos de papel e tinta e método do prisma.

2.3.1 Método de Papel e Tinta

Este método é mais comumente utilizado pela polícia. A impressão digital é rolada sobre a tinta e então rolada sobre o papel. Esta técnica transforma uma imagem tri-dimensional da impressão digital em uma imagem bi-dimensional que pode ser fotografada por uma câmera de vídeo ou capturada por um scanner, conforme ilustrada na figura 2.6 .

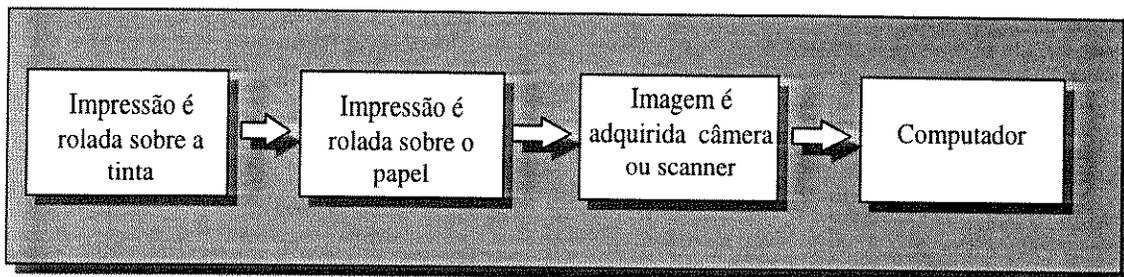


Figura 2.6- Sequência do processo de aquisição

Em geral, este processo não garante uma boa qualidade da imagem captada, visto que a tinta pode borrar partes ou toda a impressão digital, sem mencionar a inconveniência de sujar os dedos com tinta. Além disso, os movimentos e pressões desiguais do dedo sobre o papel causam distorções na imagem . Uma outra desvantagem deste método é o tempo consumido no processo de aquisição, o que impede sua aplicação para casos de reconhecimento automático de impressões digitais em tempo real.

2.3.2 Método do Prisma

O processo de aquisição empregado neste estudo tem como objetivo eliminar ou minorizar algumas das desvantagens do método tradicional. O método proposto resulta em imagens de melhor qualidade através de um processo “dirtyless”. Além disso, fornece a possibilidade de reconhecimento em tempo real.

Fisicamente, nosso sistema de aquisição consiste de um scanner e um prisma de 90° com um ângulo lateral de 45°. A figura 2.7 ilustra várias perspectivas do prisma

utilizado neste trabalho. O princípio de aquisição da impressão digital é o indivíduo encostar sua impressão digital na face inclinada do prisma.

A seguir, faremos uma pequena abordagem sobre fenômenos óticos[21], ocorridos no prisma, para melhor compreensão sobre a função do prisma neste sistema.

Reflexão - Reflexão é um fenômeno que ocorre quando a luz que incide sobre um novo meio de propagação é refletida, ou melhor, esta luz não é absorvida pelo novo meio de propagação. Podemos ver pela figura 2.8(a) um feixe de luz chegando a uma superfície e sendo totalmente refletido. Na figura 2.8(b) temos também um feixe de luz chegando à uma superfície, mas com a diferença de que, sobre a superfície inclinada, existem objetos e estes alteram o índice de refração. Conforme o índice de refração dos meios, este feixe vai ser refletido ou refratado.

Refração - Refração é um fenômeno que ocorre quando a luz que incide sobre um novo meio de propagação é absorvida por este. A luz passa de um meio para outro, porque a velocidade da luz é diferente nos dois meios. Em geral podemos achar que a velocidade da luz, em qualquer material é menor que a velocidade da luz no vácuo. De fato a velocidade da luz possui seu maior valor no vácuo. O índice de refração(η) de um meio é definido como:

$$\eta = \frac{\text{velocidade da luz no vacuo}}{\text{velocidade da luz no meio}} \quad (2.3)$$

O índice de refração é adimensional e sempre maior que 1.

Se um feixe de luz for incidir sobre a superfície do prisma da maneira mostrada na figura 2.8(a), o feixe será totalmente refletido. Quando um objeto é colocado sobre a superfície inclinada, a situação muda devido a alteração no índice de refração, veja figura 2.8(b).

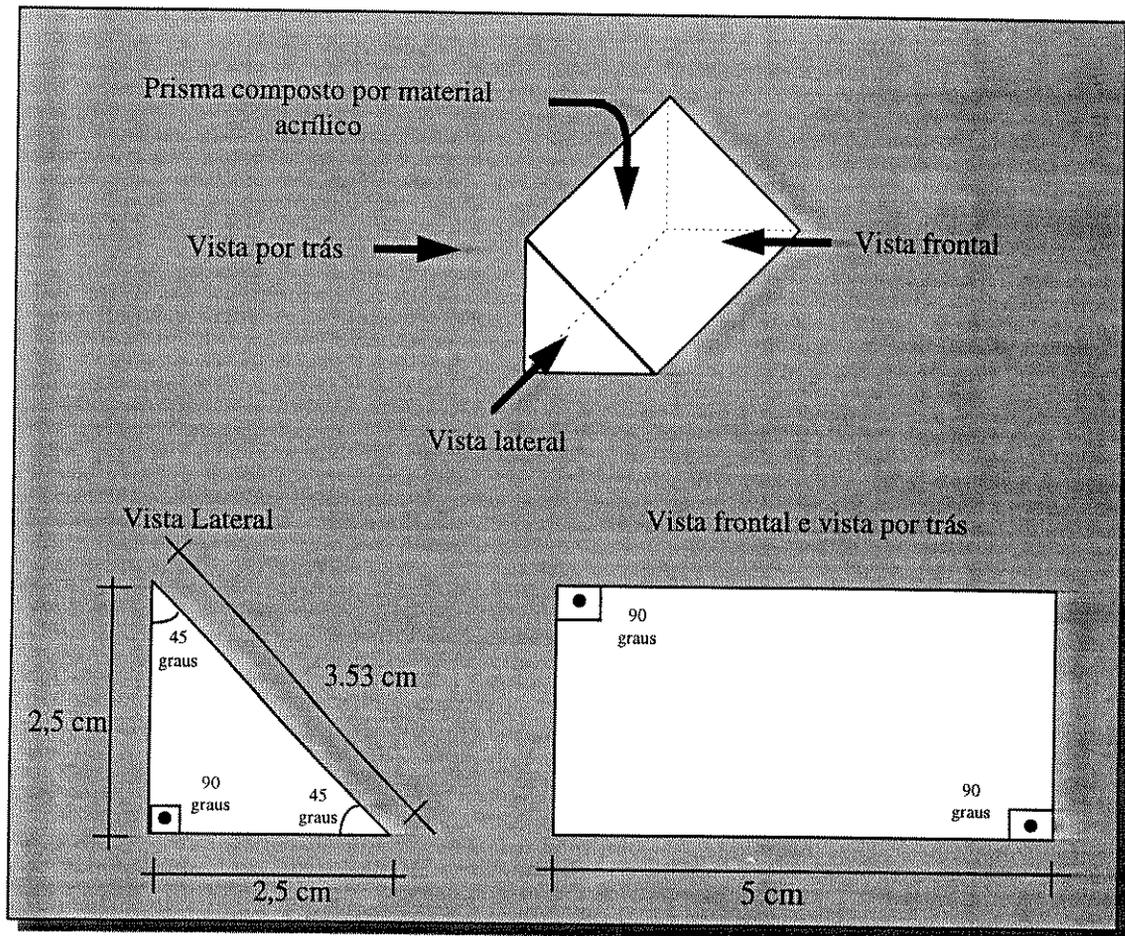


Figura 2.7- Perspectivas do prisma utilizado

A figura 2.8(c) mostra como um raio de luz é refratado para um meio com índice de refração η_3 . Podemos ver também que existe um ângulo de incidência, θ_1 , e um ângulo de refração, θ_3 , v_1 e v_3 são, respectivamente, as velocidades de propagação da luz nos meios 1 e 3. As seguintes relações são válidas:

$$\frac{\text{sen} \theta_3}{\text{sen} \theta_1} = \frac{v_3}{v_1} = \text{constante} \quad (2.4)$$

$$\frac{\text{sen} \theta_3}{\text{sen} \theta_1} = \frac{\eta_1}{\eta_3}$$

Considerando que o índice de refração do ar é $\cong 1$ e do prisma é $\cong 1,7$, podemos dizer que quando um feixe de luz chegar a superfície inclinada com um ângulo θ_1 entre

$|36^\circ \text{ a } 90^\circ|$, esse feixe será totalmente refletido, como na figura 2.8(a). À medida que o feixe de luz chega com inclinação entre $\theta_1 = |0^\circ \text{ a } 35^\circ|$, ele será refratado.

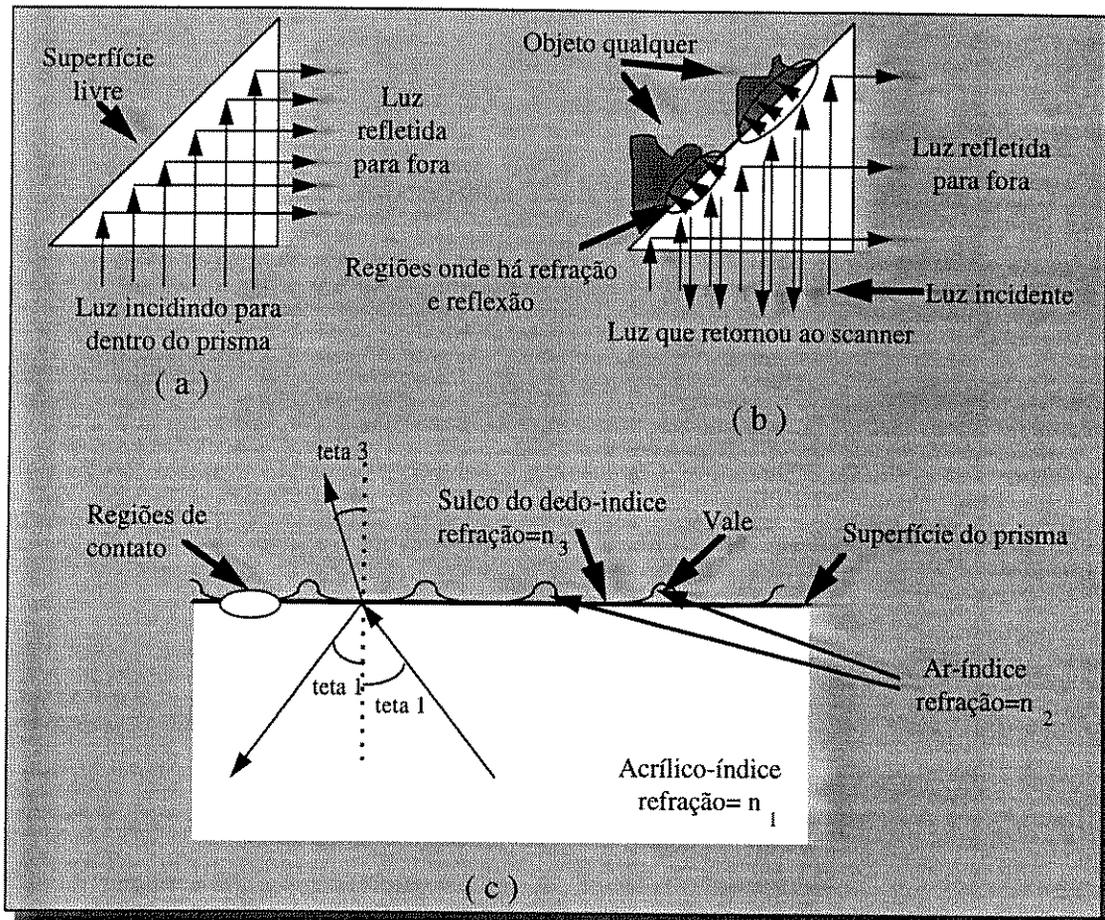


Figura 2.8- Refração e reflexão

Não conhecemos o índice de refração da pele humana, n_3 , mas seu valor deve fazer com que o ângulo com o qual os raios sejam refratados fique compreendido numa faixa bastante ampla. Ou seja, acreditamos que um raio de luz chegando à face com um inclinação entre $\theta_1 = |0^\circ \text{ a } 75^\circ|$ deva ser refratado.

Logo, quando não há objeto posto sobre o prisma, concluímos que um feixe de luz que chega à face inclinada do prisma é totalmente refletido como mostrada na figura 2.8(a). Por outro lado, quando o dedo for pressionado sobre a superfície inclinada do prisma, as linhas da impressão digital, que tocam sua face, alteram o índice de refração. Nestes pontos a luz não é mais refletida para fora do prisma, mas refratada e

parcialmente refletida de volta na direção do scanner. Como consequência, as linhas da impressão digital aparecerão brancas na imagem e os vales ou fundo, pretos.

As imagens capturadas desta maneira possuem uma qualidade muito superior às imagens adquiridas colocando-se o dedo diretamente sobre o scanner. Figura 2.9 mostra uma imagem de impressão digital captada quando o dedo é colocado diretamente sobre o scanner. Note que esta imagem não apresenta uma qualidade para fins de reconhecimento. Veja que é inviável tentar trabalhar com uma imagem com esta qualidade.

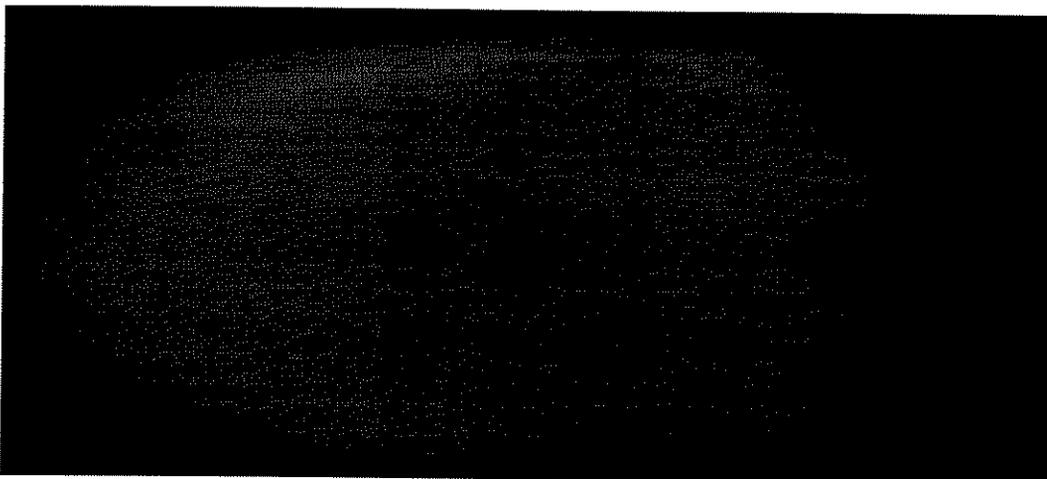


Figura 2.9- Impressão digital obtida sem o prisma

De fato, para atingirmos a qualidade da imagem da figura 2.3(d), tivemos que fazer certas mudanças no prisma. Observamos que a luz proveniente do meio poderia penetrar para dentro do mesmo causando interferência. Percebemos também que a luz refletida na face inclinada do prisma e que deveria sair para fora do mesmo pela outra face, poderia não ter esse comportamento devido a superfícies imperfeitas do prisma e outros motivos. Para contornar estes problemas, a seguinte medida foi tomada: pintar todas as superfícies do prisma em que não há contato com o dedo. Isto garante que a luz exterior não causará interferência e ao mesmo tempo a luz interior refletida pela superfície onde não há contato com o dedo, seja totalmente absorvida pela parede pintada de preto. Notamos uma melhora significativa na qualidade da imagem após este ajuste.

Apesar das modificações realizadas no prisma, este método continua apresentando certas inconveniências. Um dos problemas encontrados é a presença de um gradiente nos valores dos níveis de cinza da imagem. Este gradiente nada mais é do que um aumento nos valores dos níveis que caracterizam os pixels. Quando colocamos o dedo sobre a superfície inclinada do prisma (veja figura 2.10), a distância entre o scanner e a face do dedo, cresce à medida que aproxima-se do ângulo reto. O princípio de funcionamento do scanner não leva em consideração o fato de se variar o foco. Quando adquirimos uma imagem via scanner, o foco estará sempre ajustado sobre a superfície do mesmo. Nosso sistema de aquisição de imagem de impressões digitais contraria a distância focal ideal para o scanner. Isto faz com que, à medida que a imagem se afasta do foco, ocorra um aumento nos valores dos níveis de cinza que caracterizam a imagem. Na figura 2.11 mostramos os níveis de cinza de duas linhas da imagem de uma impressão digital. Veja como os valores que caracterizam os sulcos da impressão crescem a medida que afastamos da superfície plana do scanner. Isto cria inúmeros problemas no processamento da imagem.

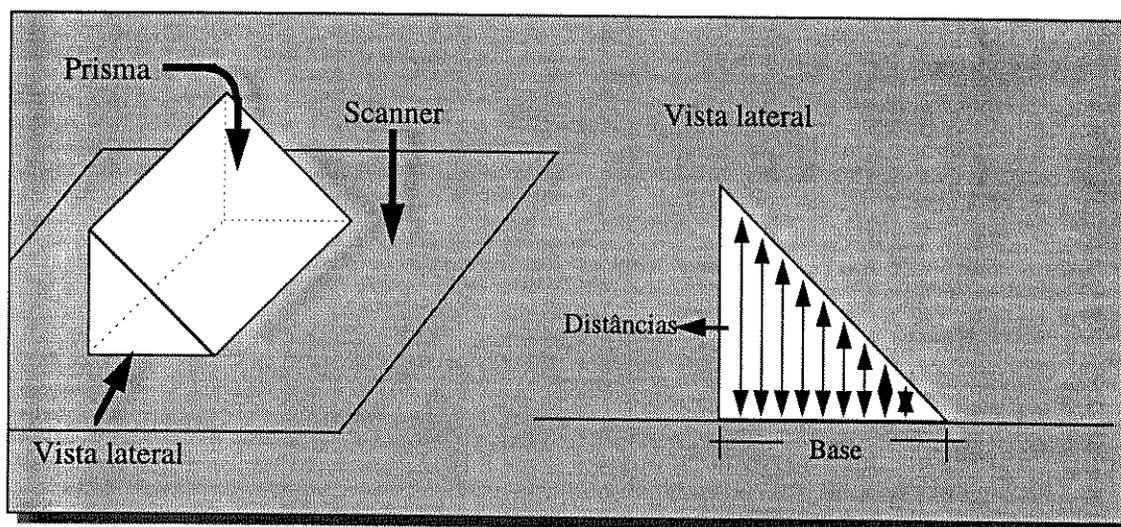


Figura 2.10- Distância focal

É importante ressaltar que a qualidade do prisma, assim como a qualidade física do dedo afetam significativamente a qualidade da imagem captada. As imperfeições do prisma, tais como, bolhas de ar interiores e superfícies imperfeitas causam erros, o suor no dedo também introduz ruído na imagem captada.

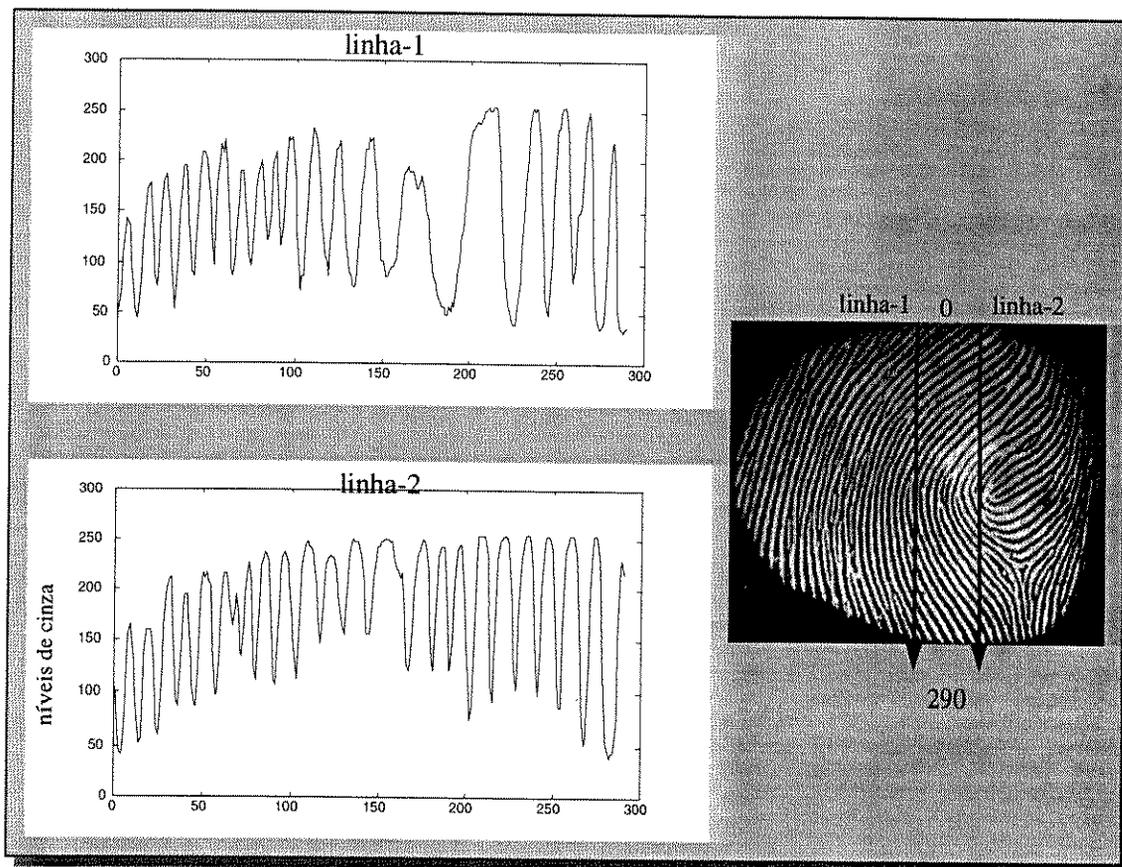


Figura 2.11- Linhas da impressão digital

Por último, mas não menos importante, é o fato de que a imagem que é adquirida possui uma característica de ser escura, ou melhor, com pouca iluminação. Nestas imagens predominam os níveis de cinza que possuem valores baixos. Isto se deve ao fato de que a quantidade de luz que volta ao scanner, luz que caracteriza os sulcos, é muito pequena.

Capítulo 3

Processamento de Imagem

3.1 Introdução

Este capítulo tem por objetivo desenvolver técnicas de processamento de imagens adequadas para manipular impressões digitais. Inicialmente, definiremos certas operações matemáticas que serão utilizadas. Posteriormente, utilizaremos estas ferramentas para compor nosso algoritmo.

3.2 Histograma

O histograma[14,15] de uma imagem digital com níveis de cinza na faixa $[0, L-1]$ é uma função discreta dada por $p(r_k) = \frac{n_k}{n}$, $k = 0, 1, 2, \dots, L-1$, onde r_k é o k -ésimo nível de cinza, n_k é o número de pixels na imagem com este nível de cinza e n é o número total de pixels na imagem.

Podemos dizer que $p(r_k)$ nos dá uma estimativa da probabilidade de ocorrência do nível r_k . A própria função histograma nos dá uma descrição da aparência da imagem. Por exemplo, o histograma de níveis de cinza da figura 3.1(a) mostra que os níveis de cinza da imagem estão concentrados na região abaixo do valor médio da escala. Assim, podemos afirmar que esta imagem possui característica de ser escura. Na figura 3.1(b) temos uma situação contrária do caso anterior. O histograma da figura 3.1(c) possui um espectro estreito, o que indica que a faixa dinâmica é pequena e corresponde a uma imagem com pouco contraste. Como todos os níveis ocorrem ao redor do valor médio da escala de níveis de cinza, a imagem aparecerá com grande concentração de cinza. Finalmente, temos a figura 3.1(d) que mostra um histograma com espalhamento significativo, correspondendo a uma imagem com alto contraste.

Apesar das propriedades discutidas serem descrições globais que não nos dizem nada com respeito a características específicas, a forma do histograma de uma imagem nos dá informações muito úteis com relação ao contraste. A discussão seguinte desenvolve métodos para manipulação de histogramas.

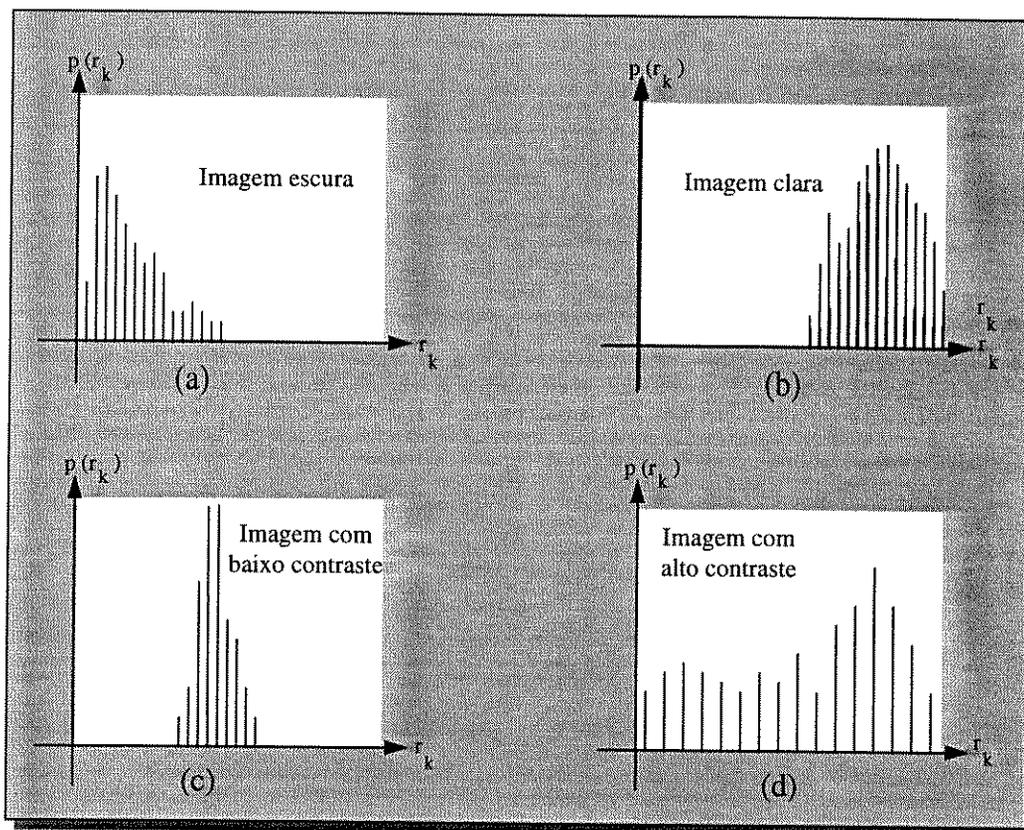


Figura 3.1- Histogramas de quatro imagens

3.2.1 Equalização do Histograma

Vamos fazer com que a variável r que representa os níveis de cinza da imagem seja melhorada. Na primeira parte de nossa discussão, assumiremos que os valores dos pixels são quantidades contínuas que estão sendo normalizadas, tal que eles pertençam ao intervalo $[0,1]$, com $r = 0$ representando preto e $r = 1$ representando branco. Depois, consideraremos uma formulação discreta e permitiremos que os valores dos pixels estejam no intervalo $[0, L - 1]$.

Para qualquer r no intervalo $[0,1]$, nós fixaremos nossa atenção nas transformações da forma

$$s = T(r) \quad (3.1)$$

que produz um outro nível s para cada pixel de valor r da imagem original. É assumido que a transformação dada pela equação 3.1 satisfaz as seguintes condições:

- (a) $T(r)$ é uma correspondência biunívoca e monotonicamente crescente no intervalo $0 \leq r \leq 1$; e
- (b) $0 \leq T(r) \leq 1$ para $0 \leq r \leq 1$.

A condição (a) preserva a ordem do preto para o branco na escala de cinza e a condição (b) garante um mapeamento que é consistente com a largura permitida para os valores dos pixels. A figura 3.2 ilustra uma função de transformação que satisfaz estas condições.

A transformação inversa de s para r é denotado

$$r = T^{-1}(s), \quad 0 \leq s \leq 1 \quad (3.2)$$

onde devemos assumir que $T^{-1}(s)$ também satisfaz as condições (a) e (b) com respeito à variável s .

Os níveis de cinza numa imagem podem ser vistos como quantidades aleatórias no intervalo $[0,1]$. Se eles são variáveis contínuas, os níveis de cinza originais e os transformados podem ser caracterizados por suas respectivas funções de densidade de probabilidade $p_r(r)$ e $p_s(s)$.

Da teoria da probabilidade elementar, se $p_r(r)$ e $T(r)$ são conhecidas e $T^{-1}(s)$ satisfaz a condição (a), a função densidade de probabilidade da transformação dos níveis de cinza é

$$p_s(s) = \left[p_{r(r)} \frac{dr}{ds} \right]_{r=T^{-1}(s)} \quad (3.3)$$

A técnica de equalização baseia-se na modificação da aparência de uma imagem pelo controle de densidade de probabilidade dos seus níveis de cinza via uma função de transformação $T(r)$.

Considere a função transformação

$$s = T(r) = \int_0^r p_r(w) dw, \quad 0 \leq r \leq 1 \quad (3.4)$$

O lado direito da primeira equação 3.4 é também conhecido como função distribuição acumulada de probabilidade de r . As condições (a) e (b) apresentadas anteriormente são satisfeitas por esta função, porque a função distribuição de probabilidade é monotonicamente crescente de 0 até 1 como uma função de r .

Da equação 3.4, a derivada de s em relação à r é

$$\frac{ds}{dr} = p_r(r) \quad (3.5)$$

substituindo $\frac{dr}{ds}$ dentro da equação 3.3, temos

$$\begin{aligned} p_s(s) &= \left[p_{r(r)} \frac{dr}{ds} \right]_{r=T^{-1}(s)} \\ &= [1]_{r=T^{-1}(s)} \\ &= 1, \quad 0 \leq s \leq 1 \end{aligned} \quad (3.6)$$

que $p_s(s)$ é uma densidade uniforme no intervalo de definição da variável transformada s .

O desenvolvimento anterior indica que utilizando-se uma função de transformação igual à distribuição de probabilidade de r produzimos uma imagem cujos níveis de cinza possuem uma densidade uniforme. Em termos de enriquecimento, este resultado implica num crescimento da faixa dinâmica dos pixels, que afeta consideravelmente a aparência da imagem.

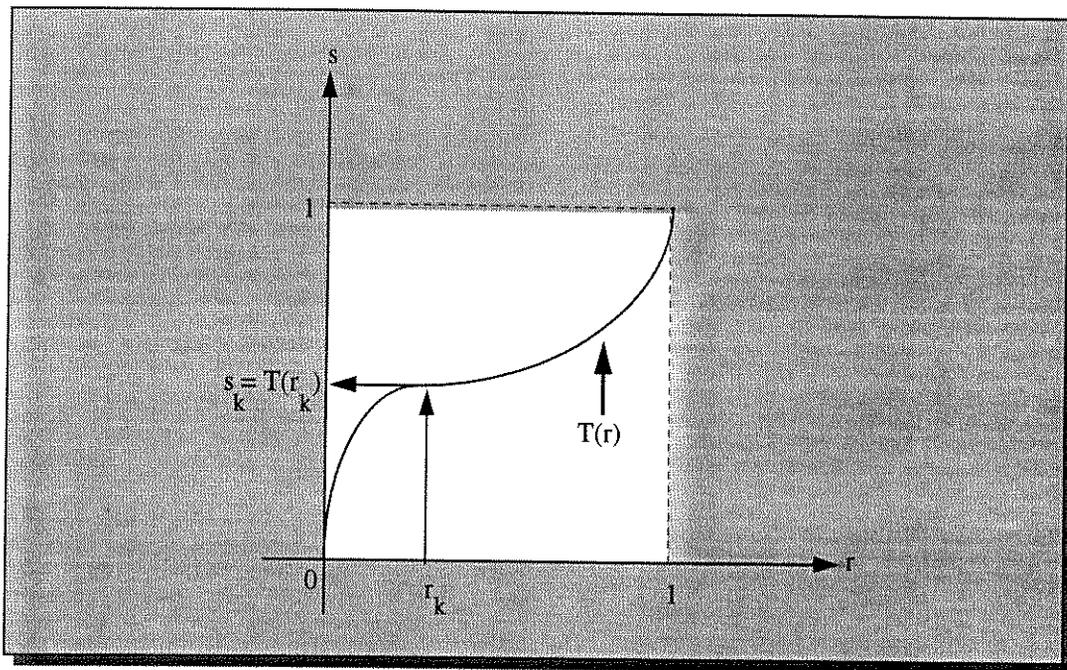


Figura 3.2- Função de transformação

Para que possamos aplicar esta técnica de transformação dos níveis de cinza no processamento digital de imagens, os conceitos formulados anteriormente devem ser adaptados para o caso discreto. Para os níveis de cinza assumindo valores discretos, temos as seguintes probabilidades:

$$p_r(r_k) = \frac{n_k}{n}, \quad 0 \leq r_k \leq 1 \text{ e } k = 0, 1, \dots, L-1 \quad (3.7)$$

onde, como indicado anteriormente, L é o número de níveis, $p_r(r_k)$ é a probabilidade do k -ésimo nível de cinza, n_k é o número de vezes com que este nível aparece na imagem, e n é o número total de pixels na imagem. O gráfico $p(r_k)$ versus r_k , é um

histograma e a técnica utilizada para obtermos um histograma uniforme é conhecida como equalização do histograma.

A forma discreta da equação 3.4 é dada pela seguinte relação

$$\begin{aligned} s_k = T(r_k) &= \sum_{j=0}^k \frac{n_j}{n} \\ &= \sum_{j=0}^k p_r(r_j), \quad 0 \leq r_k \leq 1 \quad \text{e} \quad k = 0, 1, \dots, L-1. \end{aligned} \quad (3.8)$$

A transformação inversa é denotada por

$$r_k = T^{-1}(s_k), \quad 0 \leq s_k \leq 1 \quad (3.9)$$

onde, ambos $T(r_k)$ e $T^{-1}(s_k)$ satisfazem as condições (a) e (b) estabelecidas previamente nesta seção. A transformação $T(r_k)$ pode ser computada diretamente da imagem utilizando-se a equação 3.8.

Para ilustrar a utilidade da equalização, consideremos a imagem da figura 3.3(a) e seu histograma de níveis de cinza na figura 3.3(c). O eixo horizontal representa a quantidade de níveis de cinza. O eixo vertical mostra o número de pixels para cada nível de cinza, ao invés de probabilidades. Esta representação é comum na prática porque ela é mais natural de se interpretar. A conversão de número de pixels para valores de probabilidade é trivial, basta dividirmos os valores pelas dimensões da imagem multiplicadas; isto é,

$$probabilidade = \frac{\text{valor}}{(M * N)}. \quad (3.10)$$

De forma similar, podemos normalizar o eixo horizontal para o intervalo $[0,1]$; basta dividirmos os valores por 255. Como podemos ver pelo histograma da figura 3.3(c), a imagem possui seus níveis de cinza agrupados numa faixa estreita, numa região de valores baixos. Logo a imagem é relativamente escura e sua dinâmica muito pobre. A

figura 3.3(b) é a imagem da figura 3.3(a) equalizada, com seu histograma mostrado na figura 3.3(d). Note que houve uma melhora considerável na dinâmica da imagem, ou seja uma melhora considerável do contraste da imagem.

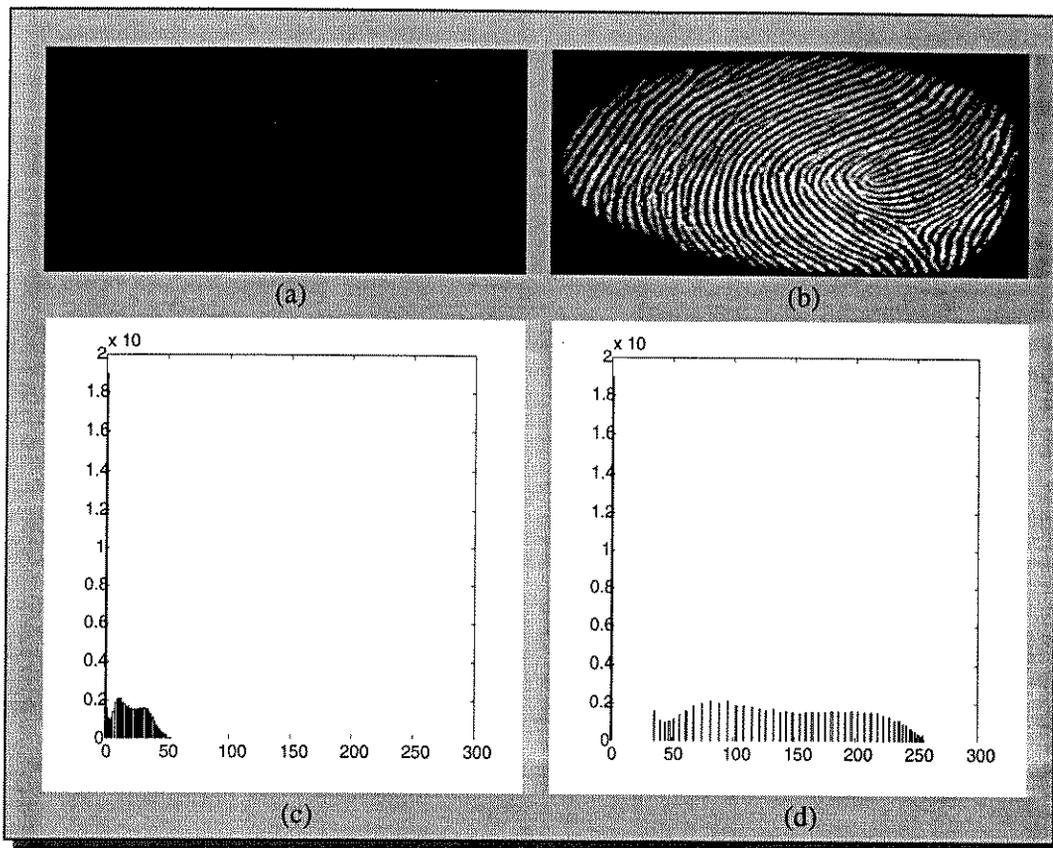


Figura 3.3- Exemplo de equalização de histograma

3.3 Inversão

Inversão é uma operação em que transformamos a imagem em sua forma negativa, ou seja, as partes brancas se tornarão escuras e vice-versa.

Conceitualmente, a operação de inversão de uma imagem é bastante simples. Suponha uma imagem digital com níveis de cinza na faixa $[0, L-1]$, se desejamos inverter esta imagem, basta que realizemos a seguinte operação em cada pixel:

$$n = (L-1) - v \quad (3.11)$$

Onde n é o novo valor de cada pixel e v o antigo valor.

3.4 Limiar

Limiar[14,15], mais conhecido por “threshold”, é um dos mais importantes meios para segmentação de imagens. Aqui vamos ilustrar o conceito de operação de limiar através de um exemplo abaixo.

Suponha que o histograma em níveis de cinza mostrado na figura 3.4(a) corresponde a uma imagem, $f(x,y)$, composta de objetos luminosos sobre um fundo escuro, de tal forma que os pixels que caracterizam o objeto e o fundo estejam agrupados em dois grupos distintos. Uma maneira óbvia de extrairmos o objeto, é selecionar um limiar T que separe os dois grupos. Então, para qualquer ponto (x,y) para o qual $f(x,y) > T$, o ponto é considerado como um ponto do objeto; caso contrário, o ponto é dito pertencer ao fundo. A figura 3.4(b) mostra um caso mais usual, onde temos três grupos dominantes caracterizando a imagem, como por exemplo, dois grupos de objetos luminosos sobre um fundo escuro. Neste caso, a regra de classificação é formulada da seguinte maneira. Se $T_1 < f(x,y) \leq T_2$ for satisfeita, este ponto fará parte do objeto A , e se $f(x,y) > T_2$ for satisfeito este ponto pertencerá ao objeto B , ou o ponto pertencerá ao fundo caso $f(x,y) \leq T_1$ for verdadeiro. Este último tipo de limiar, geralmente é menos confiável que o limiar simples. A razão é a dificuldade de estabelecer múltiplos limiares que isolem regiões de interesse, especialmente quando o número de grupos no histograma é grande. Estes problemas são melhor resolvidos utilizando-se um limiar dinâmico.

Uma imagem que passou por uma operação limiar ou “threshold” é definida da seguinte forma:

$$g(x,y) = \begin{cases} 1 & \text{se } f(x,y) > T \\ 0 & \text{se } f(x,y) \leq T \end{cases} . \quad (3.10)$$

onde $g(x,y)$ é uma imagem binária. Assim os pixels denominados 1 (ou qualquer outro nível de intensidade) correspondem aos objetos e os pixels 0 correspondem ao fundo.

Quando T depende somente da função $f(x,y)$, o limiar é dito global, mas se T depende de $f(x,y)$ e $p(x,y)$ (que denota propriedades locais do ponto) então o limiar é

dito local. Se, em adição, T depende das coordenadas espaciais x e y , o limiar é dito dinâmico.

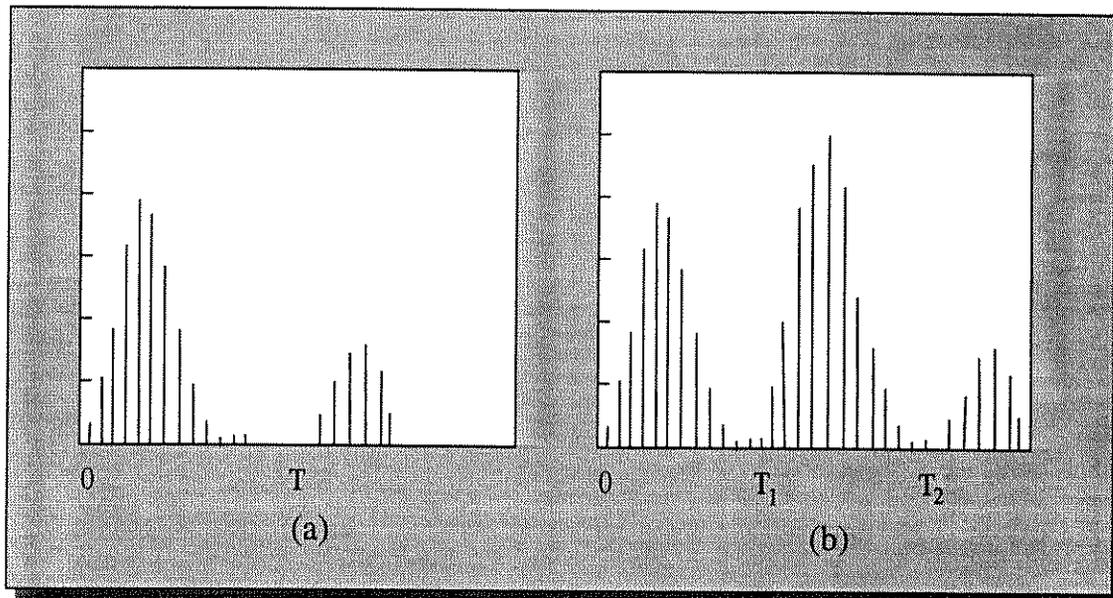


Figura 3.4- Histogramas

3.4.1 Limiar Global

Com referência à discussão que tivemos anteriormente, a operação mais simples de limiar ou “threshold” é quando estabelecemos um limiar simples ou global, T , como ilustrado na figura 3.4(a). A segmentação da imagem é feita varrendo-se a imagem pixel por pixel, classificando cada pixel como pertencente ao objeto ou ao fundo, de acordo com o valor do pixel em relação ao limiar T . Como dito anteriormente o sucesso da segmentação depende exclusivamente da possibilidade de partição do histograma.

A figura 3.5 mostra o resultado de uma operação de limiar global sobre a imagem da figura 3.3(b), na figura 3.5(a) temos o histograma da imagem original. A imagem da figura 3.5(b) é o resultado da segmentação utilizando-se um limiar igual a 125. Os objetos de nosso interesse são os sulcos da impressão digital. Podemos notar que todos os pontos com nível $\leq T$ foram mapeados para o valor 0 (preto), e todos os pixels com nível $> T$ foram mapeados para o valor 255 (branco). De fato, o principal objetivo é gerar uma imagem binária, branco e preto. Note que a operação limiar é irreversível, ou seja, não se pode voltar à imagem original.

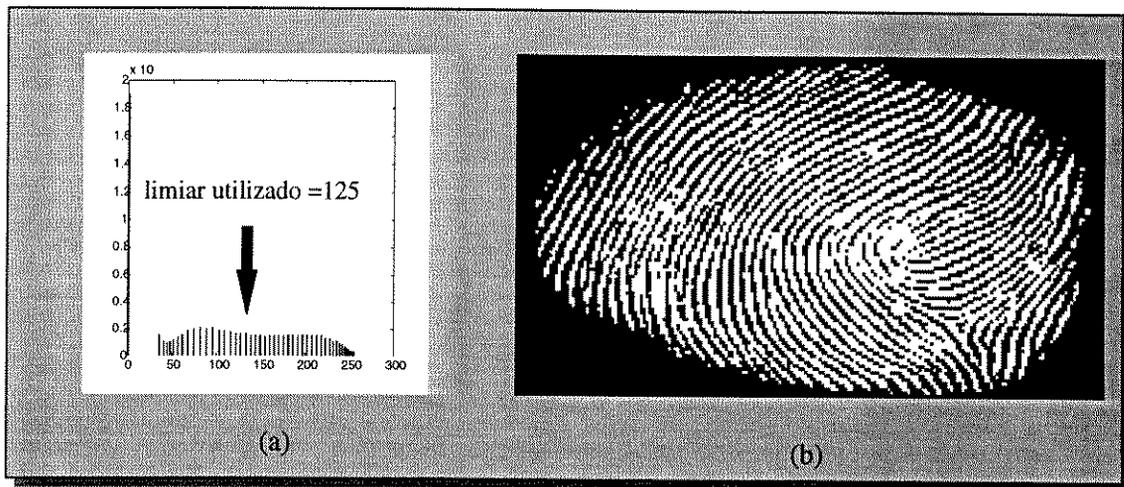


Figura 3.5- Limiar global

3.4.2 Limiar Dinâmico

Quando temos problemas no sentido de isolar objetos em uma imagem, devido à forma de seu histograma, devemos tentar utilizar o limiar dinâmico. No capítulo anterior vimos que podemos representar uma imagem digital como sendo uma matriz. Se dividirmos esta matriz em pequenas matrizes, teremos a imagem particionada. Podemos então dizer que nossa imagem é composta por sub-imagens. A cada sub-imagem extraímos o histograma e a partir dele é adotado o valor médio ou mediana como limiar, ou seja, para cada sub-imagem é estabelecido um limiar. Este limiar é global dentro da sub-matriz, mas quando consideramos a matriz total, temos vários limiares para a imagem, logo ele é dinâmico. Na figura 3.6(a) vemos a forma original de uma imagem, ou melhor, uma matriz $N \times M$. Já na figura 3.6(b), a imagem está dividida em imagens menores.

Na figura 3.7 temos dois resultados obtidos utilizando-se limiar dinâmico. O primeiro, figura 3.7(a), foi obtido subdividindo a imagem original em pequenas imagens de dimensão 31×31 . Na figura 3.7(b), utilizamos matrizes menores, 11×11 . Em ambos casos, o valor médio de cada histograma foi utilizado como limiar.

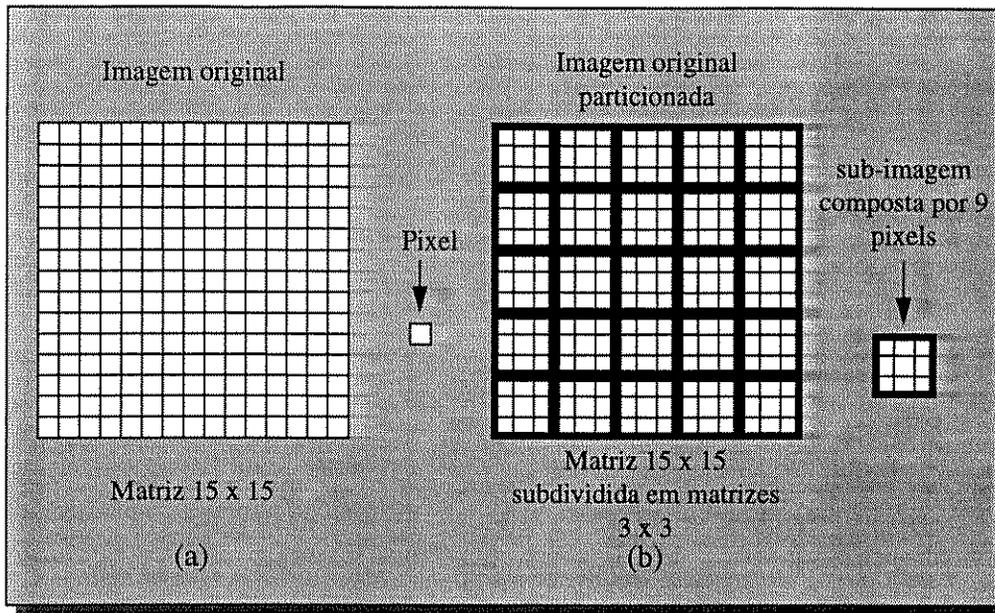


Figura 3.6- Formato de uma imagem

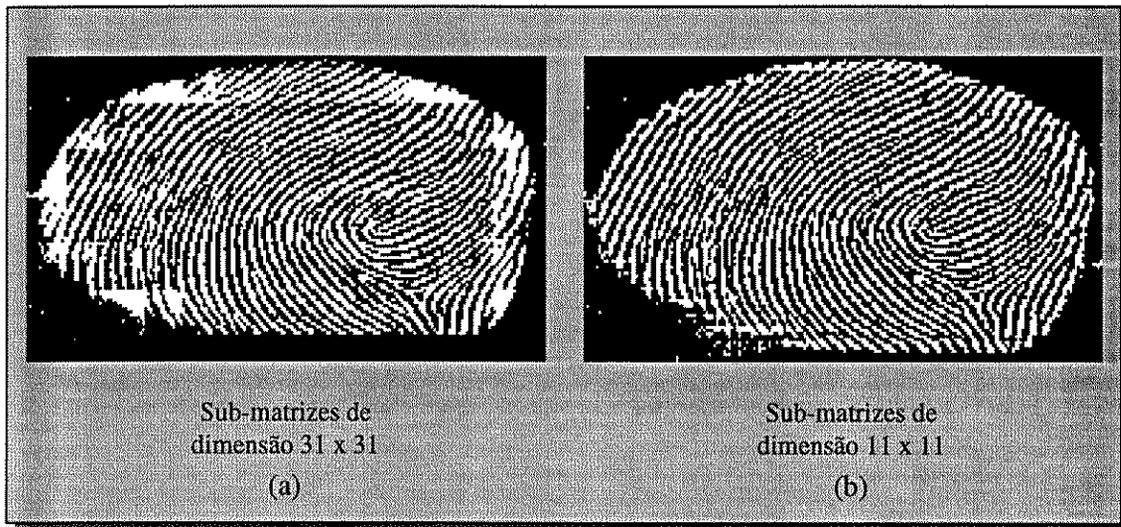


Figura 3.7- Limiar dinâmico

3.5 Morfologia Matemática

A Morfologia Matemática[14] é uma ferramenta nova no processamento de imagens, devido a sua grande aplicabilidade. De fato, ela já é utilizada para solucionar problemas em diversas áreas, tais como: reconhecimento de caracteres, imagens médicas, microscopia, inspeção industrial, metalurgia, visão computacional, etc. A idéia básica desta ferramenta é comparar os objetos de uma imagem com outro objeto de

forma conhecida, denominado elemento estruturante, através de certas operações. O formalismo de tais operações está baseado na teoria de conjuntos, geometria integral e topologia. Deste modo, uma imagem pode ser vista como um conjunto de objetos elementares. Nesta seção, descreveremos algumas operações morfológicas. Começaremos tratando de duas operações básicas: erosão e dilatação, e com o auxílio destas, descreveremos outras.

3.5.1 Morfologia Matemática em Imagens Binárias

Algumas definições básicas: Sejam A e B dois conjuntos em um plano real, R^2 , com elementos $a = (a_1, a_2)$ e $b = (b_1, b_2)$, respectivamente.

- A translação de A por $x = (x_1, x_2)$, denotada $(A)_x$, é definida como $(A)_x = \{c | c = a + x, \text{ para } a \in A\}$.

- A reflexão de B , denotada por \widehat{B} , é definida como $\widehat{B} = \{x | x = -b, \text{ para } b \in B\}$.

- O complemento do conjunto A é $A^c = \{x | x \notin A\}$.

- A diferença de dois conjuntos A e B , denotada por $A - B$, é definida como $A - B = \{x | x \in A, x \notin B\} = A \cap B^c$.

Exemplo: A figura 3.8 ilustra as definições apresentadas anteriormente. A Figura 3.8(a) mostra um conjunto A . A figura 3.8(b) mostra a translação de A por $x = (x_1, x_2)$. Observe que a translação é feita pela adição (x_1, x_2) para cada elemento de A . A figura 3.8(c) mostra um conjunto B e na figura 3.8(d) está a reflexão de B sobre a origem. Finalmente, a figura 3.8(e) mostra o conjunto A e seu complemento. Na figura 3.8(f) temos a diferença entre o conjunto A e o conjunto B .

- Dilatação: A dilatação de A por B , denotada por $A \oplus B$, é definida como $A \oplus B = \{x | (\widehat{B})_x \cap A \neq \emptyset\}$, onde \emptyset é o conjunto vazio. Em outras palavras, o processo de dilatação de A por B consiste em obter a reflexão de B sobre sua origem e então deslocar sua reflexão de x . A dilatação de A por B é um conjunto de todos os deslocamentos por x , tal que \widehat{B} e A sobrepõem pelo menos um elemento diferente de zero. Baseando-se nesta interpretação, a equação acima pode ser escrita, também, como

$$A \oplus B = \{x | ((\hat{B})_x \cap A) \subseteq A\}.$$

Seja B comumente referido como elemento estruturante na dilatação e em outras operações morfológicas.

Exemplo: A figura 3.9(c) mostra o resultado da operação de dilatação da imagem da figura 3.9(a) pelo elemento estruturante da figura 3.9(b). Na figura 3.9(c) a linha pontilhada mostra a imagem original e as linhas sólidas mostram a imagem dilatada.

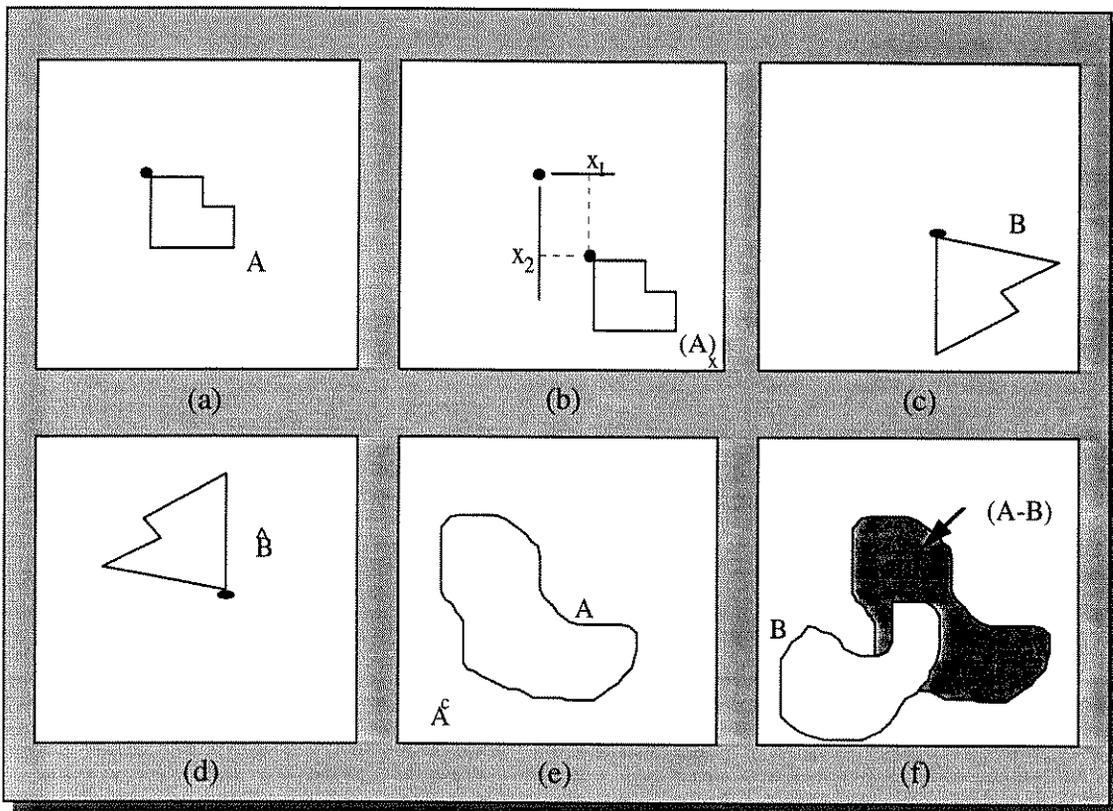


Figura 3.8- Operações básicas

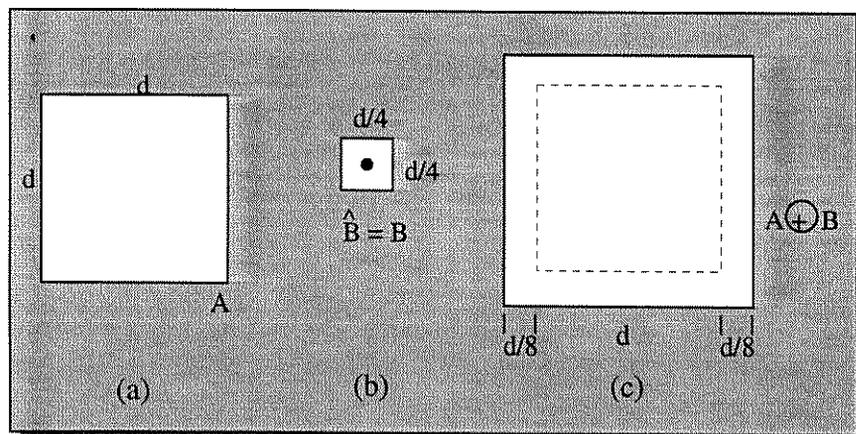


Figura 3.9- Dilatação

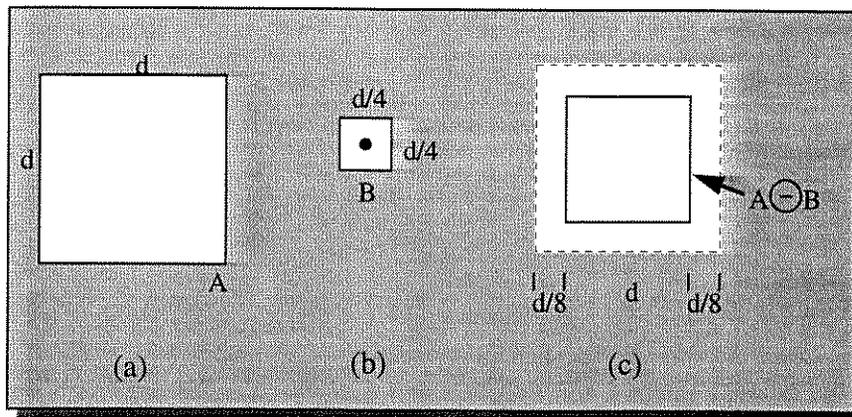


Figura 3.10- Erosão

•Erosão: A erosão de A por B é definida como $A \ominus B = \{x | (B)_x \subseteq A\}$. Em outras palavras, a erosão de A por B é o conjunto de todos os pontos x , tal que B , transladado por x , está contido em A . Note que dilatação e erosão são operações duais com respeito aos conjuntos complementares e refletidos, ou seja, $(A \ominus B)^c = A^c \oplus \hat{B}$. Podemos dizer que a dilatação expande uma imagem, enquanto a erosão a reduz.

Exemplo: A figura 3.10 mostra um processo similar àquele mostrado na figura 3.9. Como anteriormente, o conjunto A é mostrado com linhas pontilhadas no ítem (c) e as linhas sólidas representam o novo conjunto ou objeto resultante da erosão.

•Abertura: Uma operação de abertura visa alisar os contornos de uma imagem, quebrando estreitos elos de ligação e eliminando pequenas ilhas (pontos isolados). A abertura de um conjunto A por um elemento estruturante B , é denotada por $A \circ B = (A \ominus B) \oplus B$, o que significa uma erosão de A por B , seguido por uma dilatação por B .

Exemplo: A figura 3.11 ilustra a operação de abertura de um conjunto A com um disco como elemento estruturante. Na figura 3.11(a) temos o conjunto A , no ítem (b) é mostrada a operação de erosão. Observe o disco percorrendo o conjunto A durante a erosão, que uma vez completada, resulta na imagem do ítem (c). Se nosso elemento estruturante fosse maior teríamos a eliminação do ponto que liga os dois objetos principais. A figura 3.11(d) ilustra o processo de dilatação do conjunto erodido. Finalmente, temos no ítem (e) o conjunto resultante da operação de abertura.

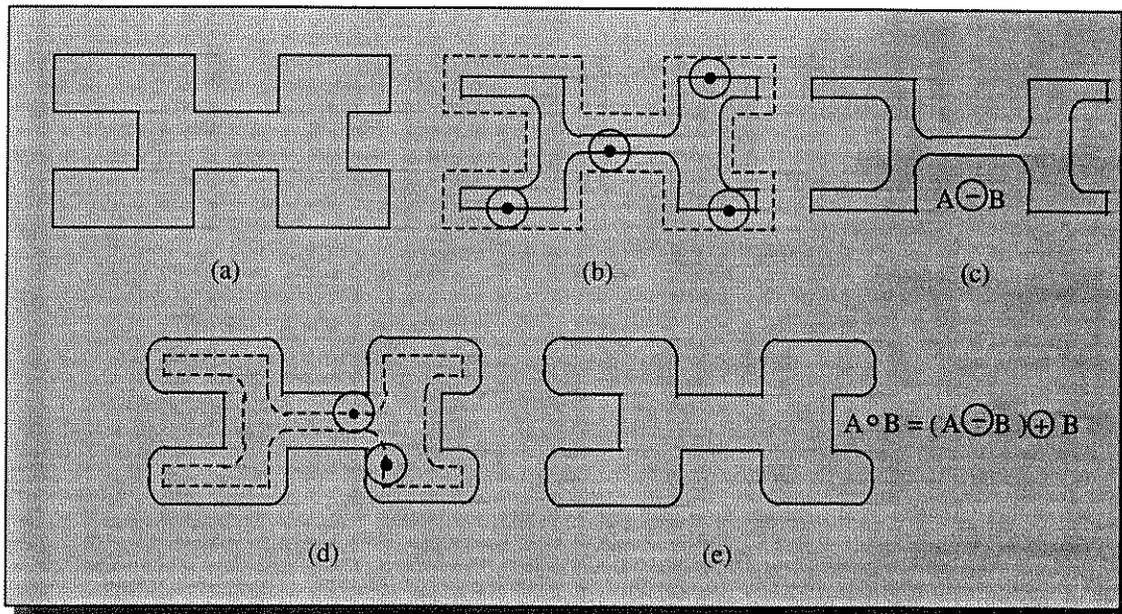


Figura 3.11- Abertura

•Fechamento: O fechamento pode ser considerado como uma operação inversa da abertura, visando fechar pequenos buracos, preencher furos nos contornos, unir grandes golfos, etc. O fechamento do conjunto A por um elemento estruturante B é definido como $A \bullet B = (A \oplus B) \ominus B$, ou seja, dilatação de A por B , seguido por uma erosão por B .

Exemplo: Nas figuras 3.12(a)-(d) ilustramos o processo de fechamento.

•Transformada em Tudo ou Nada: A transformada morfológica em tudo ou nada é uma ferramenta básica na detecção de formas, definida como $A * B = (A \ominus B_1) - (A^c \ominus B_2)$.

Exemplo: Na figura 3.13(a) temos uma imagem composta por três objetos, um retângulo, um quadrado e um quadrado com uma pequena saliência. Se nossa tarefa for marcar a localização do quadrado com saliência, então devemos fazer uma cópia exata dele para servir de elemento estruturante. A erosão da imagem por este elemento estruturante produzirá um ponto único que estará em sua origem quando o elemento estruturante preencher exatamente o quadrado com saliência. Neste caso a erosão consegue localizar o objeto desejado na imagem. O problema é diferente se nós desejarmos achar o quadrado. Erosão por uma cópia exata do quadrado marcará dois

pontos, um para o quadrado perfeito e outro para o quadrado com saliência. Uma solução para este problema é aplicar a transformada em tudo ou nada com o par de elementos estruturantes $B = (B_1, B_2)$, onde B_1 é o quadrado e B_2 é o complemento do quadrado afinado (veja figura 3.13(b)). Aplicando a transformada em tudo ou nada será produzido um único ponto no centro do quadrado desejado.

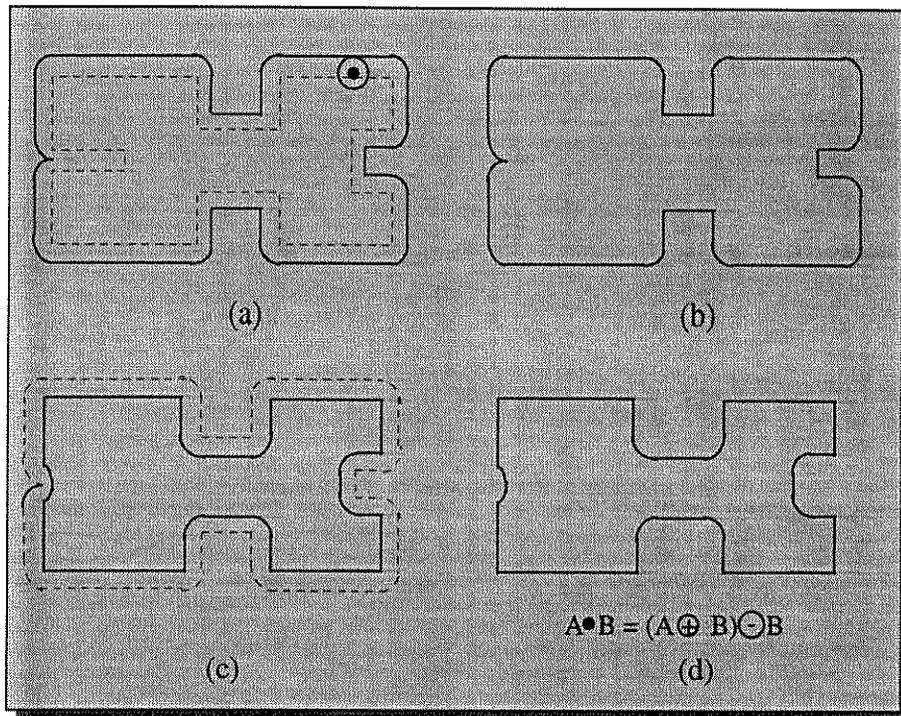


Figura 3.12- Fechamento

•Afinamento: O afinamento de um conjunto A por um elemento estruturante B , denotado $A \otimes B$, é definido como $A \otimes B = A \cap (A * B)^c$. Uma expressão mais útil para o afinamento de A é baseado numa sequência finita de elementos estruturantes: $B = (B^1, B^2, \dots, B^n)$, onde B^i é a versão rotacionada de B^{i-1} . Usando este princípio, podemos definir o afinamento por uma sequência de elementos estruturantes como: $A \otimes (B) = ((\dots((A \otimes B^1) \otimes B^2) \dots) \otimes B^n)$. Em outras palavras, o processo consiste em afinar A pelo elemento B^1 , o resultado afinado por B^2 , e assim por diante. O processo é repetido por um número finito de vezes até não ocorrerem mais mudanças na imagem. Exemplo: Na figura 3.14, temos uma demonstração deste processo.

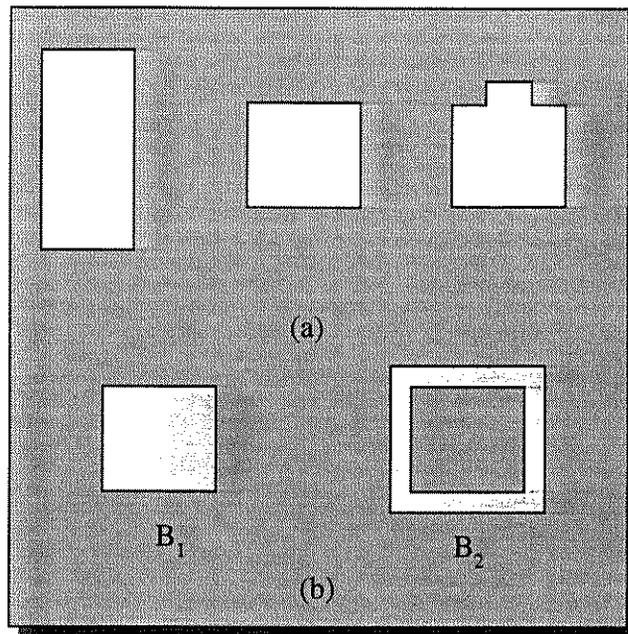


Figura 3.13- Transformada em tudo ou nada

- Espessamento: Espessamento é a operação morfológica dual do afinamento, definida pela expressão: $A \circ B = A \cup (A \otimes B)$, onde B é o elemento estruturante adequado para o espessamento.

- Esqueleto: O esqueleto $S(A)$ de um conjunto A é o conjunto A afinado até a indempotência, ou melhor, até que não ocorram mais mudanças.

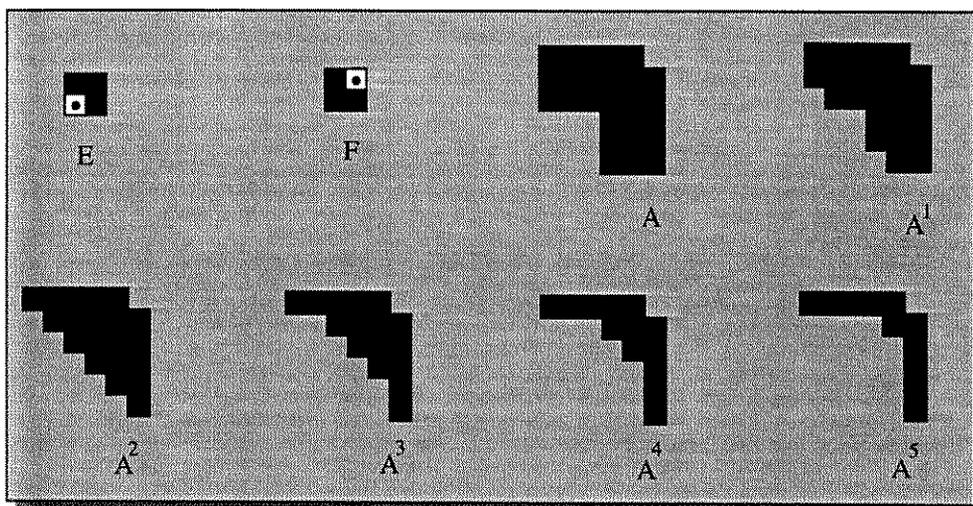


Figura 3.14- Afinamento

3.5.2 Morfologia Matemática em Imagens com Níveis de Cinza

Nesta seção estenderemos os conceitos de morfologia matemática para imagens em nível de cinza para as operações de dilatação, erosão, abertura e fechamento. Como na seção anterior, nossa atenção está voltada para a extração de componentes úteis na representação e descrição de formas. Através das discussões que se seguem trabalharemos com funções imagem da forma $f(x, y)$ e $b(x, y)$, onde $f(x, y)$ denota a imagem de entrada e $b(x, y)$ representa um elemento estruturante. Aqui também devemos assumir que estas funções são discretas, isto é, se Z denota um conjunto de números inteiros, então $(x, y) \in Z \times Z$ e f e b são funções que designam valores em nível de cinza (um número real do conjunto dos números reais, R) para cada distinto par de coordenadas (x, y) . Se os níveis de cinza assumem valores inteiros, Z é trocado por R .

•Dilatação: Sejam $f(x, y)$ e $b(x, y)$, respectivamente a imagem de entrada e o elemento estruturante. Dilatação, em níveis de cinza, de f por b , denotada por $f \oplus b$, é definida como:

$$(f \oplus b)(s, t) = \max\{f(s-x, t-y) + b(x, y) \mid (s-x), (t-y) \in D_f ; (x, y) \in D_b\},$$

onde D_f e D_b são os domínios de f e b , respectivamente. Como anteriormente, b é um elemento estruturante do processo morfológico, mas observe que agora ele é uma função e não mais apenas um conjunto. A expressão acima que define dilatação em níveis de cinza é muito similar à expressão de convolução bidimensional, com a operação de \max substituindo a soma da convolução e a adição substituindo o produto da convolução. O processo de dilatação consiste em escolhermos o valor máximo de $f + b$ numa vizinhança definida pela forma do elemento estruturante. Os efeitos sobre a imagem quando realizamos esta operação são dois: (1) Se todos os valores do elemento estruturante são positivos, a imagem resultante tenderá a ser mais brilhante; e (2) detalhes escuros ou são reduzidos ou eliminados, dependendo de como seus valores e formas relacionam-se com o elemento estruturante utilizado na dilatação. Na figura 3.15 temos uma demonstração deste processo, onde nós demonstramos o mecanismo da dilatação por meio de uma função unidimensional. Neste caso a expressão fica reduzida à $(f \oplus b)(s) = \max\{f(s-x) + b(x) \mid (s-x) \in D_f \text{ e } x \in D_b\}$.

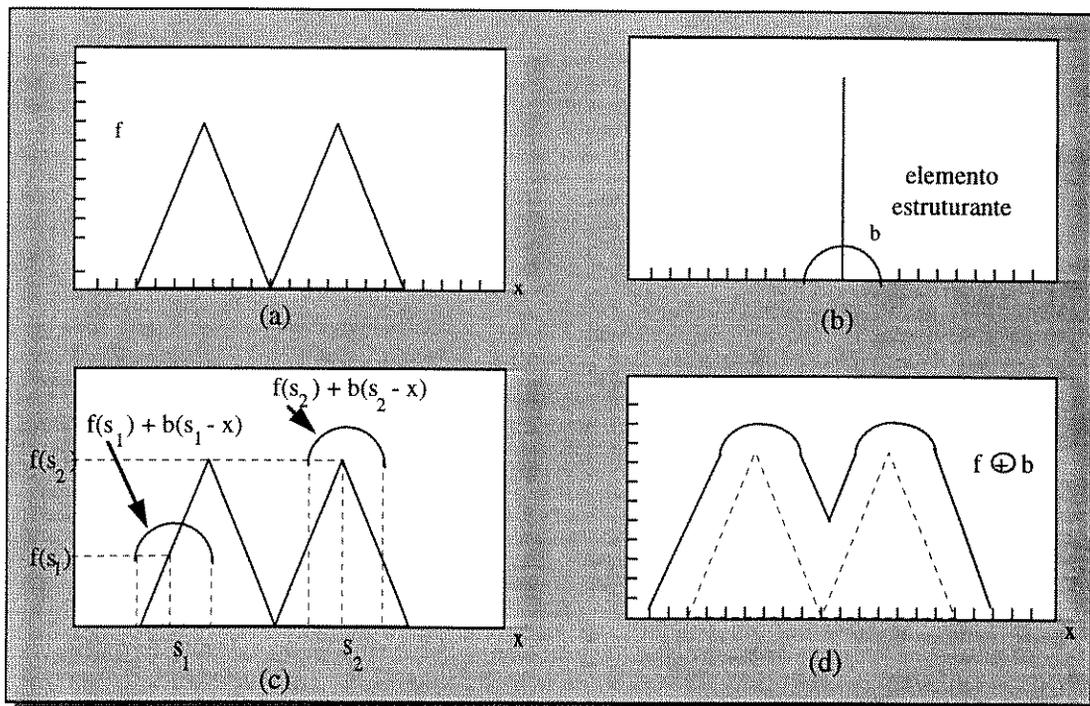


Figura 3.15- Dilatação em níveis de cinza

•Erosão: Erosão de f por b , em níveis de cinza, denotado por $f \ominus b$, é definido como: $(f \ominus b)(s, t) = \min\{f(s+x, t+y) - b(x, y) \mid (s+x), (t+y) \in D_f; (x, y) \in D_b\}$, onde D_f e D_b são os domínios de f e b , respectivamente. Observe novamente a semelhança na forma da equação que rege a erosão com a convolução bidimensional. A operação \min substitui a soma da convolução e a subtração substituindo o produto da convolução. A erosão consiste em escolher o valor mínimo de $f - b$ numa vizinhança definida pela forma do elemento estruturante. Os efeitos gerais da utilização desta operação são dois: (1) se todos os valores do elemento estruturante são positivos, a imagem resultante tende a ser mais escura, e (2) o efeito de detalhes com maior brilho que sejam menores em “área” que o elemento estruturante é reduzido. O grau de redução depende dos valores dos níveis de cinza ao redor do detalhe, da forma e valores de amplitude do elemento estruturante. Na figura 3.16 temos uma ilustração deste processo, onde novamente utilizamos uma função unidimensional para ilustrar o resultado. Em uma dimensão a operação erosão fica reduzida à $(f \ominus b)(s) = \min\{f(s+x) - b(x) \mid (s+x) \in D_f \text{ e } x \in D_b\}$.

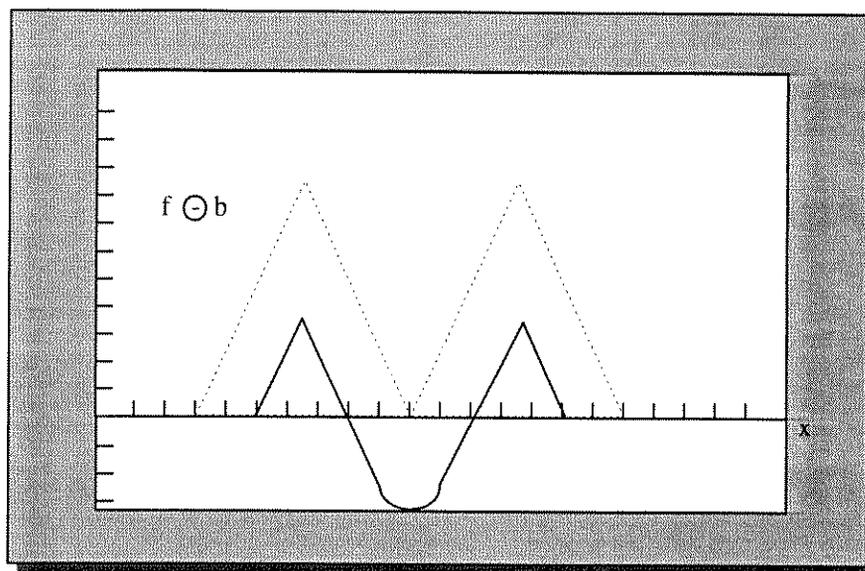


Figura 3.16- Erosão em níveis de cinza

Como anteriormente, a dilatação e erosão são duais em relação à função complementação e reflexão. Isto é $(f \ominus b)^c(x, y) = (f^c \oplus \hat{b})(x, y)$, onde $f^c = -f(x, y)$ e $\hat{b} = b(-x, -y)$.

•Abertura e Fechamento: De forma análoga ao caso da imagem binária, a abertura e o fechamento, ambos em níveis de cinza, de uma imagem f por uma imagem (elemento estruturante) b , são definidos como $f \circ b = (f \ominus b) \oplus b$ e $f \bullet b = (f \oplus b) \ominus b$, respectivamente.

Note que a abertura é simplesmente a erosão de f por b , seguida por uma dilatação do resultado por b . De forma similar, o fechamento é a dilatação de f por b , seguida por uma erosão por b .

A abertura e fechamento para imagens em níveis de cinza são duais com relação ao complemento e reflexão, isto é $(f \bullet b)^c = f^c \circ \hat{b}$.

Estas duas operações possuem uma interpretação geométrica simples. Suponha que nós estejamos visualizando uma função imagem $f(x, y)$ numa perspectiva 3-D, com os eixos x e y sendo as coordenadas espaciais e o terceiro eixo correspondendo ao brilho (isto é, ao valor de f). Nesta representação a imagem aparece como uma superfície discreta, cujos valores da função em qualquer ponto (x, y) são os de f naquelas coordenadas. Vamos supor que nós desejamos realizar a operação de abertura

de f por um elemento estruturante esférico, b . O mecanismo da operação abertura de f por b pode ser interpretado geometricamente como um processo de rolar a esfera sobre toda a superfície inferior de f . A abertura desta imagem é constituída por todos os pontos de maior valor onde a esfera tocou a superfície inferior. Na figura 3.17 temos um exemplo deste mecanismo. No ítem (a) é mostrada uma linha da imagem em níveis de cinza como sendo uma função contínua apenas para efeito de simplificação. No ítem (b) está mostrada a posição da esfera em várias pontos. No ítem(c) temos o resultado da abertura ao longo desta linha. Todos os picos que foram estreitos em relação ao diâmetro da bola são reduzidos em amplitude e estão com sua forma alisada. Em aplicações práticas, a abertura é utilizada para remover pequenos(em relação ao tamanho do elemento estruturante) detalhes brilhantes ou claros, deixando, porém, detalhes maiores relativamente intactos. A erosão inicial remove os pequenos detalhes como também deixa a imagem mais escura. A dilatação subsequente aumenta o brilho da imagem sem reintroduzir os detalhes removidos pela erosão.

Figura 3.18(a) e (b) mostra o resultado do fechamento de f por b . Aqui a bola percorre a superfície, os picos geralmente permanecem em sua forma original. Na prática, o fechamento é utilizado para remover regiões escuras de uma imagem, deixando regiões claras relativamente intactas. A dilatação inicial remove os detalhes escuros da imagem tornando a imagem mais clara e a subsequente erosão escurece a imagem sem reintroduzir os detalhes removidos pela dilatação.

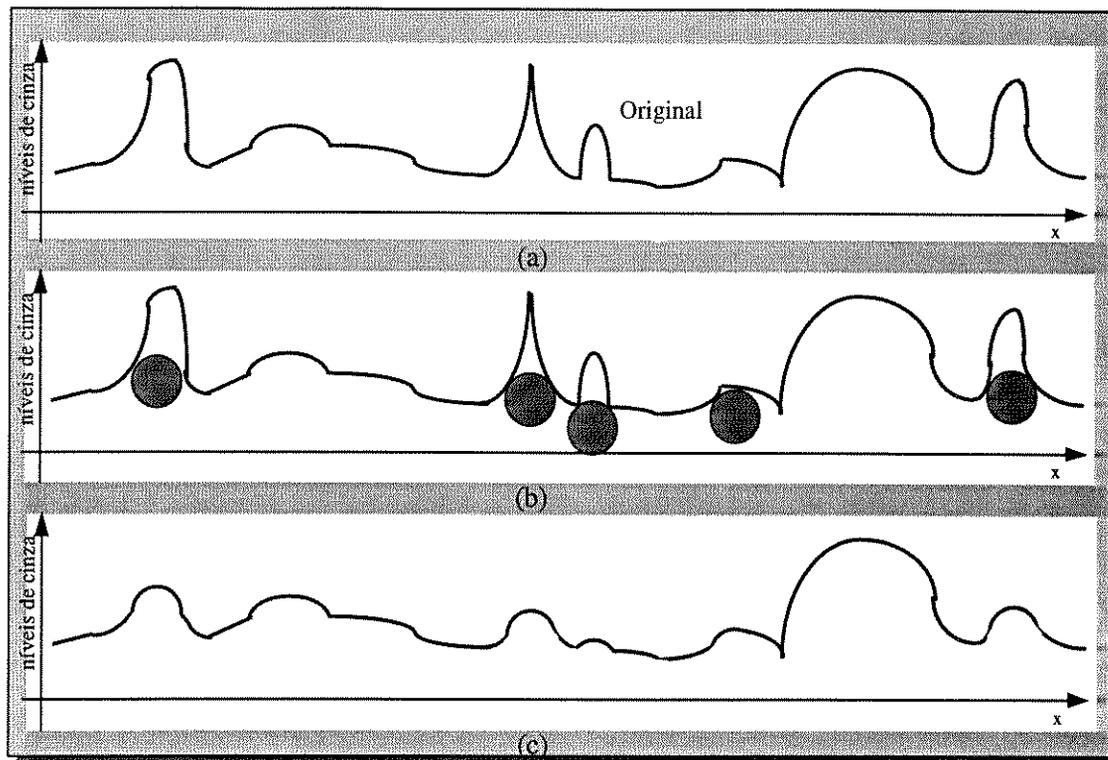


Figura 3.17- Abertura em níveis de cinza

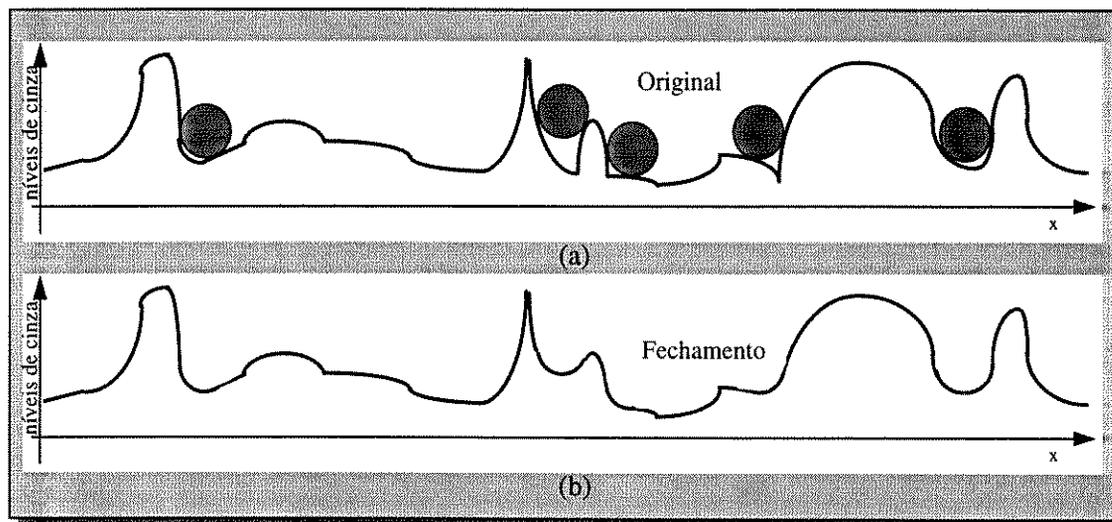


Figura 3.18- Fechamento em níveis de cinza

•Top-hat: Esta operação é definida como $h = f - (f \circ b)$, onde, como anteriormente, f é uma imagem de entrada e b é um elemento estruturante. A transformação top-hat é uma técnica muito utilizada para realçar detalhes de uma imagem, principalmente na presença de formas.

Na figura 3.19(a)-(c) temos uma demonstração da operação top-hat. Podemos ver no item (c), que após a aplicação dessa ferramenta, apenas os detalhes mais claros e estreitos permanecem.

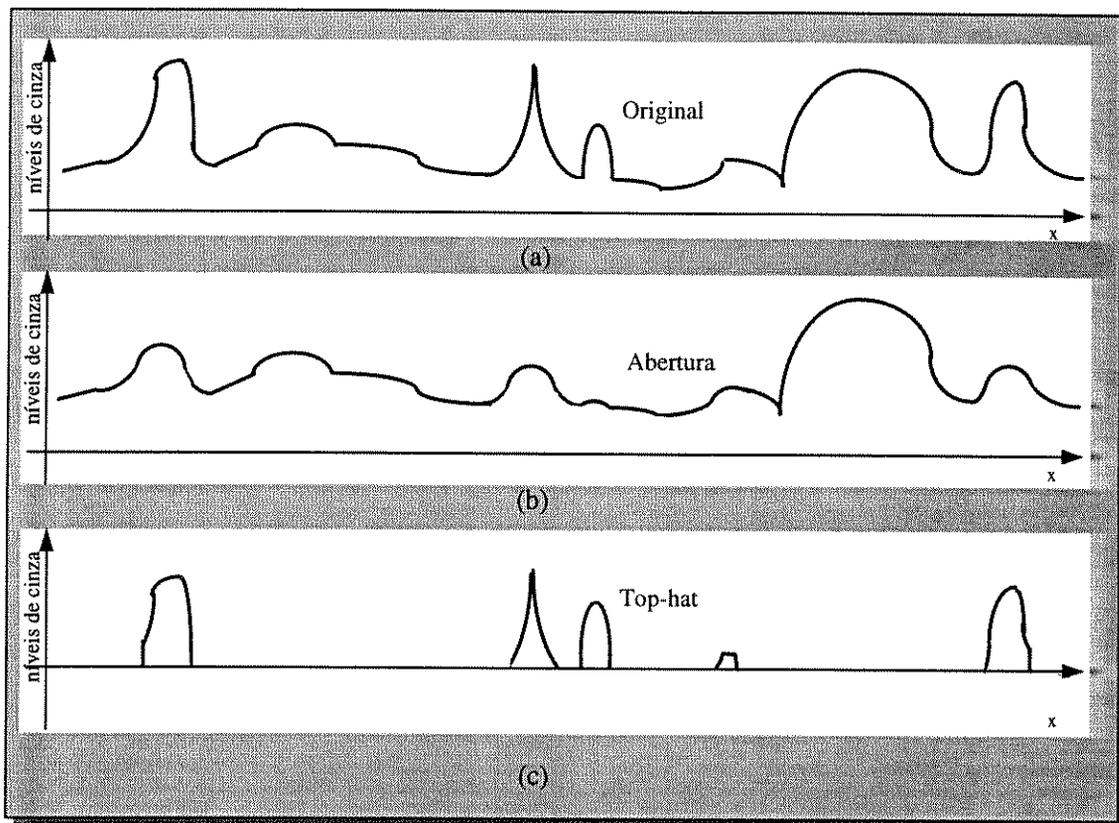


Figura 3.19- Top-hat

3.6 Algoritmo de Processamento

Nesta seção descreveremos a maneira pela qual obtivemos o algoritmo de processamento utilizado no processamento de imagens de impressões digitais.

Inicialmente, nos questionamos quais os caminhos e ferramentas que deveriam ser utilizadas neste processo. Uma maneira natural para solucionar esta questão, seria estudar as técnicas sugeridas na literatura. Surpreendentemente o levantamento bibliográfico realizado mostrou que poucas referências são disponíveis neste assunto particular. Partimos então para o projeto e desenvolvimento de um algoritmo. Decidimos por utilizar a morfologia matemática, baseando-se no fato de que esta ferramenta é um conceito novo no processamento de imagens, possui uma crescente utilização e está sendo utilizada para solucionar os mais diversos problemas, mostrando assim ser uma ferramenta de grande potencial.

Através dos estudos referentes a impressões digitais, podemos observar que existem certas formas que são muito importantes para a caracterização de uma impressão. Estas formas são denominadas minúcias e possuem nomes, tais como, bifurcações, ilhas, pontos de final de linha. Veja um exemplo de minúcias na figura 3.20

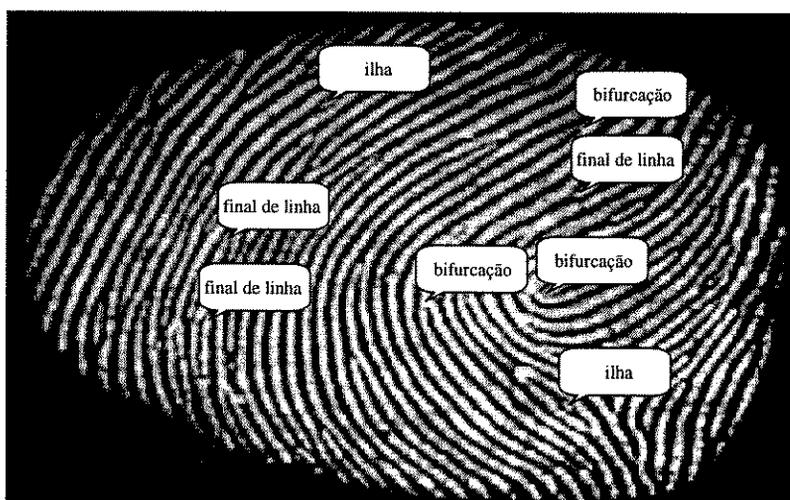


Figura 3.20- Exemplo de minúcias

Trabalhando numa imagem em níveis de cinza, temos uma grande dificuldade em detectar estas minúcias. Porém, a partir de um esqueleto, encontramos muito mais facilidade para localizar. Veja na figura 3.21(a) como estas minúcias são fáceis de serem encontradas com um simples algoritmo de análise de ponto à ponto. Portanto, nosso

objetivo no processamento de imagens é obter o melhor esqueleto possível para impressões digitais.

3.6.1 Equalização

Para a obtenção do esqueleto, nossa primeira tarefa é resolver o problema da imagem em relação a sua característica de ser escura. Isto pode ser solucionado utilizando-se a operação de equalização do histograma. Na figura 3.21(b) temos uma imagem de uma impressão digital, sem nenhum processamento, no ítem (c) temos o resultado após a operação de equalização. Muitas vezes a expressão “imagem original” foi atribuída a uma imagem equalizada, é uma convenção simples que utilizamos para simplificar a explicação. É importante ressaltar que a operação equalização não é uma ferramenta que altere a quantidade de informação da imagem, ela apenas dispõe a informação de uma outra maneira, logo, podemos nos referir a imagem pós-equalização como sendo original.

3.6.2 Limiar

Uma vez equalizada a imagem, nosso novo objetivo é obter uma imagem binária. A ferramenta que realiza esta operação é o limiar ou threshold. Temos que escolher dentre um limiar global ou dinâmico, para isto, devemos analisar as particularidades de uma impressão digital. Uma imagem de uma impressão digital é composta por diversos objetos, estes objetos a que nos referimos são as diferentes linhas na impressão digital. Referimos as linhas como sendo objetos diferentes, visto que cada região de uma impressão digital contém partes de linhas e sulcos e esta região possui características próprias em relação a imagem total. Isto deve-se em certo modo ao problema do gradiente em níveis de cinza na imagem. Podemos, por exemplo dizer que uma região poderia ser segmentada, como sucesso, se fosse utilizado um limiar “ x ”, mas para outra região isto não seria válido e teríamos que adotar um limiar “ y ”. Estas particularidades vão de encontro ao limiar dinâmico explicado na seção 3.3.2 . Para confirmar nossa escolha pelo limiar dinâmico, fizemos testes com o limiar global e esta ferramenta provocou certas distorções, tais como: em algumas regiões diversas linhas foram

consideradas como um mesmo objeto, o contrário também é verdadeiro, quando em algumas regiões tudo foi convertido para o valor correspondente ao fundo, ou melhor, os objetos desapareceram.

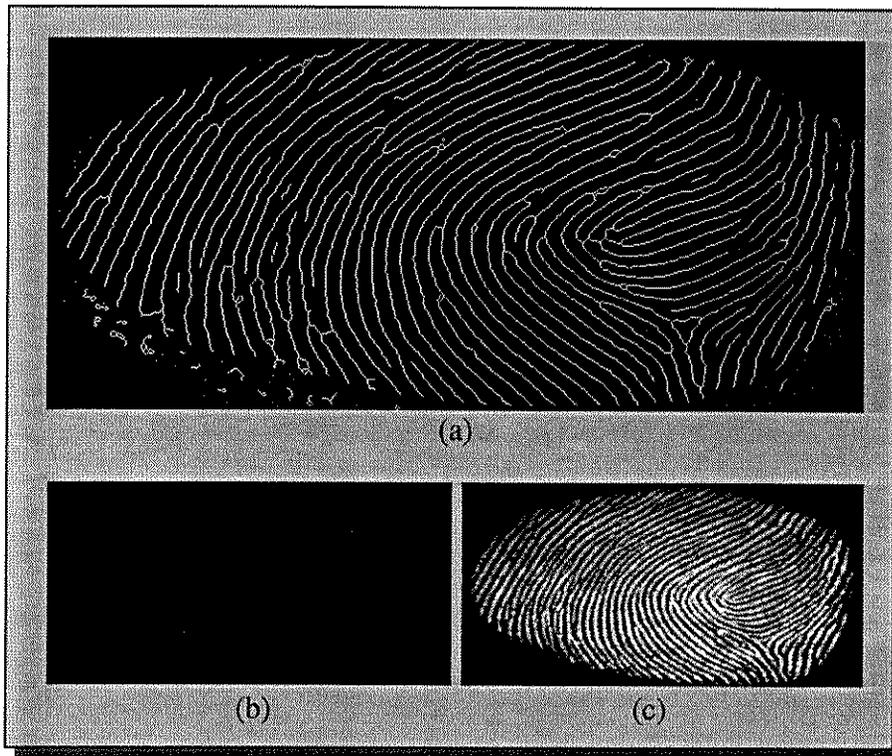


Figura 3.21- Esqueleto e equalização

Veja exemplo na figura 3.22(a)-(c) , onde temos três exemplos de binarização, com diferentes limiares. Observe que em certas regiões destacadas ocorreram erros de segmentação. No ítem (a) temos uma segmentação utilizando-se limiar igual a 100. Note que uma grande região tende a se unir formando um único objeto. A medida que aumentamos o limiar os objetos, linhas da impressão digital, vão se separando, ítem (b), mas certas regiões adquirem à tendência de serem classificadas como fundo, ítem(c).

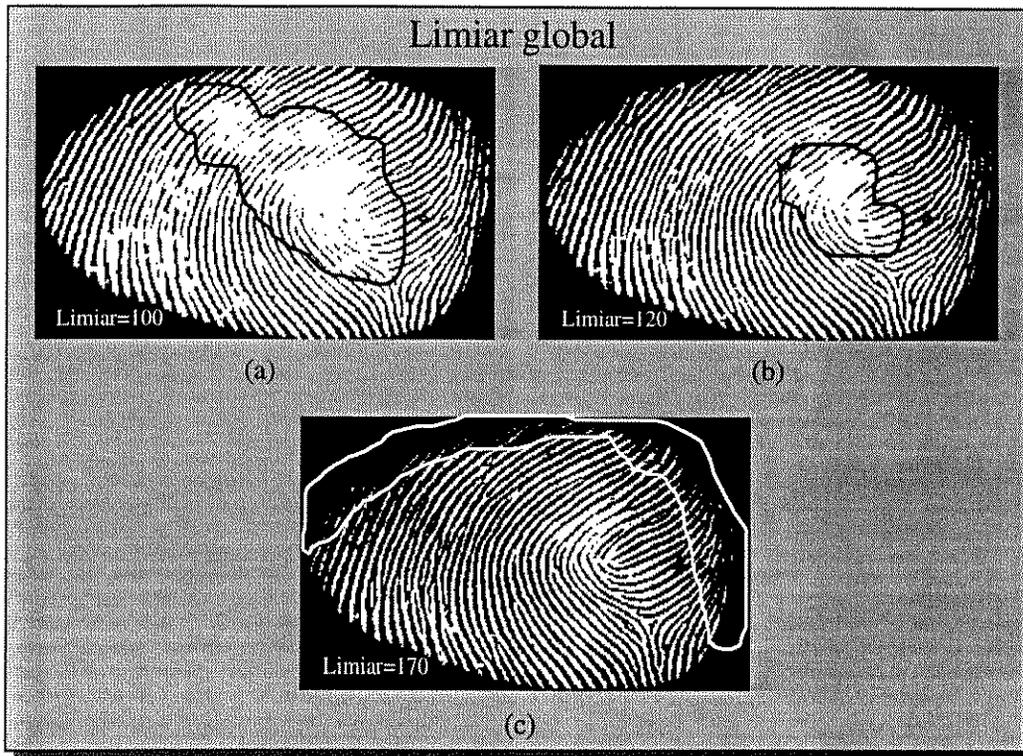


Figura 3.22- Limiar global

Para a utilização do limiar dinâmico devemos primeiramente definir qual é a melhor dimensão para as sub-imagens, ou regiões. A melhor forma para segmentarmos um imagem é quando os pixels que caracterizam os objetos e o fundo estejam agrupados em dois grupos distintos no correspondente histograma(vide seção 3.3). Raciocinando por esta lógica, a melhor dimensão para a sub-imagem é aquela que agrupa apenas um sulco e um correspondente vale, teremos então apenas um objeto luminoso e um vale ou fundo. Através da figura 2.5 vemos que a distância que separa um sulco de outro é de aproximadamente $0,5\text{ mm}$ que corresponde à aproximadamente 13 pontos. Este é o valor da dimensão que deve ser utilizado no limiar dinâmico.

Com este valor de dimensão obtivemos nosso melhor resultado de segmentação, veja na figura 3.23 (a) e (b) dois exemplos.

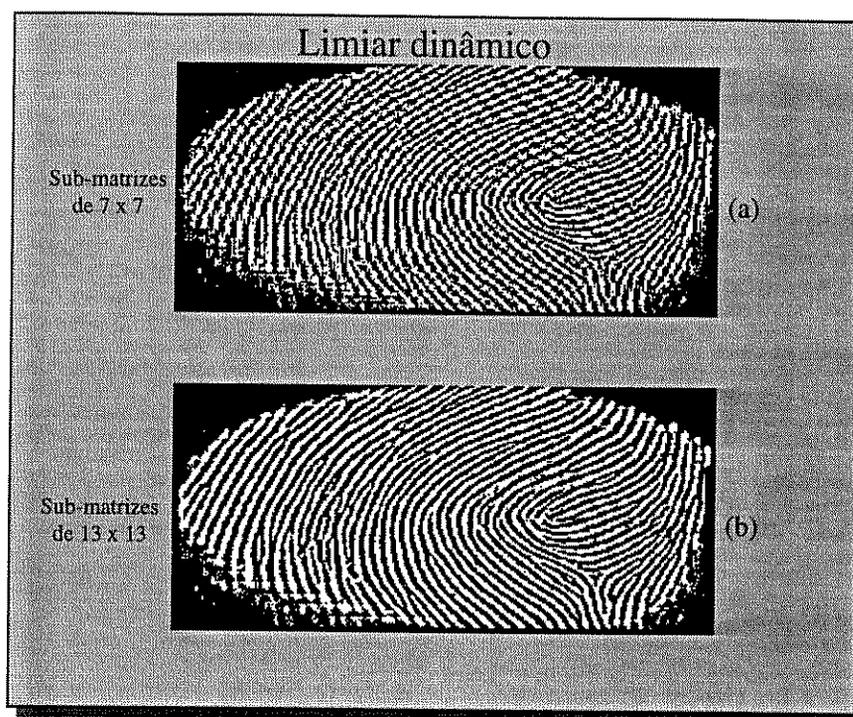


Figura 3.23- Limiar dinâmico

Apesar dos bons resultados obtidos utilizando-se este método, alguns erros ainda são cometidos. Dentre eles podemos citar linhas que são particionadas ou unidas. Muitas vezes isto ocorre devido ao ruído presente na imagem. É importante encontrar um caminho que tente minorizar tais erros, pois eles distorcem as características da impressão digital.

Uma vez que não podemos transformar a imagem original diretamente para uma imagem binária, devemos trabalhar na imagem em níveis de cinza, tentando melhorar sua qualidade, e então sim realizarmos a operação threshold. O segmento dentro da morfologia matemática que trabalha com imagens em tons de cinza é a morfologia matemática em níveis de cinza. Dentro deste ramo da morfologia matemática existe uma ferramenta muito útil para nosso caso, ela é uma composição de dilatações e erosões, chamada Top-hat (vide seção 3.4.2). Esta operação é especial, pois ela literalmente extrai os sulcos da impressão, eliminando o problema do gradiente em níveis de cinza. Para a aplicação desta ferramenta é necessário a utilização de um elemento estruturante, suas dimensões são de vital importância, conforme explicado no desenvolvimento teórico. Utilizamos um elemento estruturante circular com raio de 9 pixels. Este valor foi obtido analisando-se a figura 2.5, onde vemos que um sulco ocupa aproximadamente $0,3\text{mm}$ correspondendo a 8 pixels. O círculo deve possuir um valor ímpar, então

optamos pelo valor 9 pixels. Quando utilizamos um valor superior a este, corremos o risco de não eliminarmos o problema do gradiente e à medida que diminuimos este valor extraímos os sulcos, mas suas dimensões serão cada vez menores. Veja na figura 3.24(a)-(b) como ficam os sulcos após a utilização do top-hat com elementos estruturantes de valor 9 pixels e 7 pixels, respectivamente. No ítem (c) vemos a operação threshold global, com limiar = 0, sendo aplicada com sucesso após a utilização da ferramenta top-hat. Se compararmos esta imagem com o resultado obtido com o mesmo threshold global(veja figura 3.22) veremos que a ferramenta top-hat é muito eficiente. Observe porém, que ainda existe algum ruído presente nos vales, evidenciado através da figura 3.24(b). Este ruído acarreta muitos erros na extração do esqueleto.

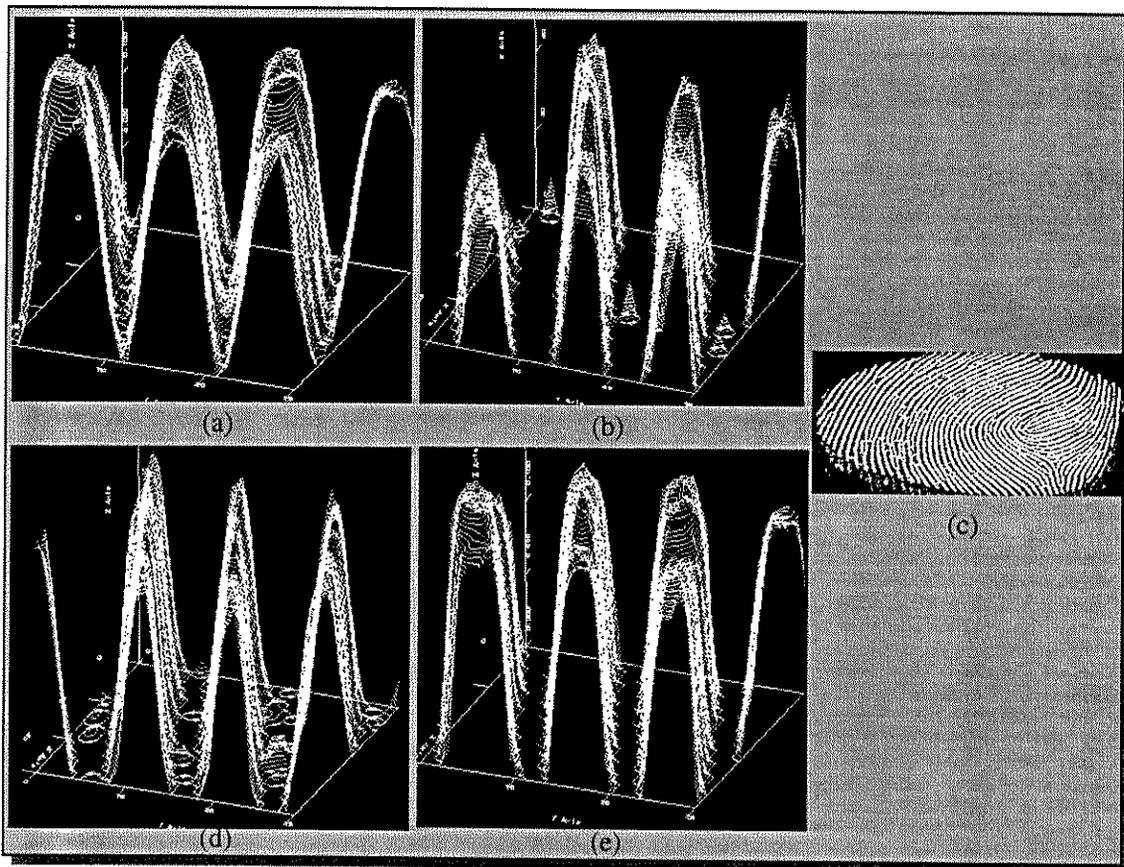


Figura 3.24- Relevo de uma pequena porção da impressão digital

Para eliminarmos este ruído utilizamos a seguinte técnica: invertemos a imagem original da impressão digital, de forma a transformar sulcos em vales e vales em sulcos e a partir daí, utilizamos novamente a ferramenta top-hat para extrairmos o relevo da

impressão digital inversa, veja figura 3.24(d). É importante lembrarmos que este relevo da figura 3.24(d) corresponde aos vales da imagem original. Logo após, subtraímos o resultado da operação top-hat da imagem inversa como o da imagem original. Esta técnica elimina o ruído entre os sulcos. A figura 3.24(e) mostra o resultado final.

Na figura 3.25 temos as imagens de impressões digitais correspondentes a este processo. No ítem (a) está a imagem original, no ítem (b) temos o resultado da inversão da imagem. A impressão correspondente ao ítem (c) é o resultado da operação top-hat sobre a imagem original e ao seu lado, ítem (d), está o resultado da operação top-hat sobre a imagem invertida. Finalmente, no ítem (e), temos o resultado da subtração das duas imagens((c) e (d)) já binarizado. O processo utilizado na binarização foi um threshold global com limiar = 0.

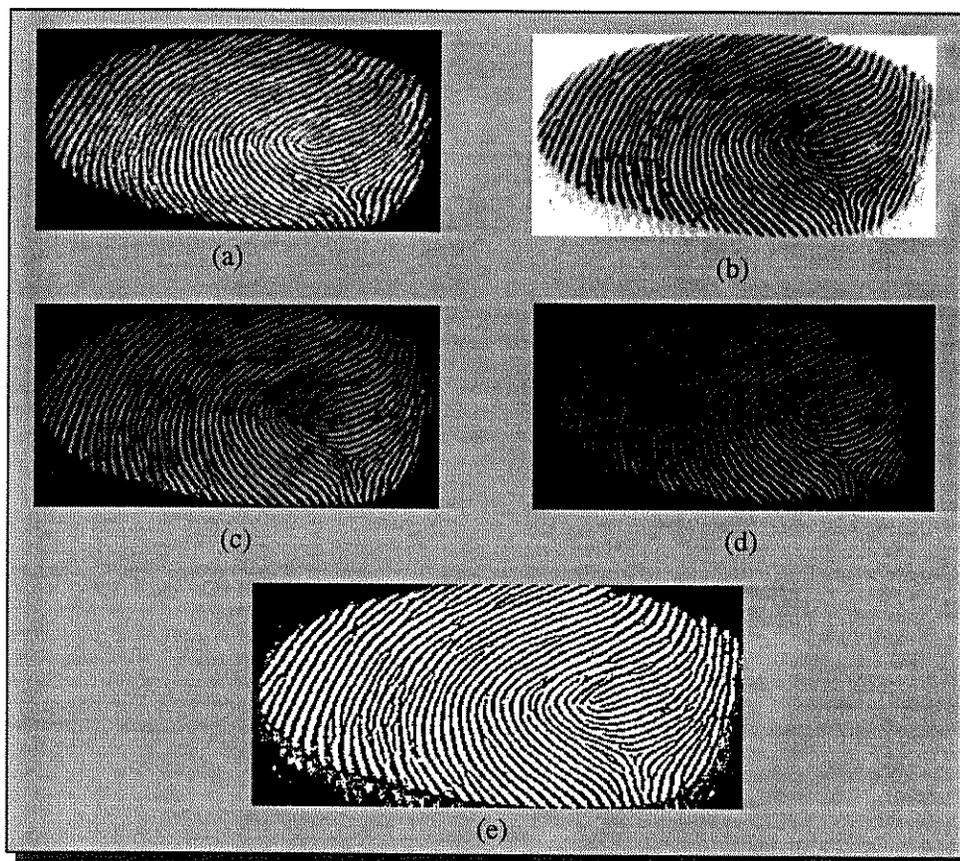


Figura 3.25- Top-hat

3.6.3 Esqueleto

Uma vez obtida uma imagem binária com um grau de qualidade aceitável, podemos prosseguir no processo de extração do esqueleto morfológico. O algoritmo utilizado foi proposto por Roland T. Chin e Ben-Kwei Jnag. Este algoritmo constrói o esqueleto a partir de uma composição de dois esqueletos, cada um construído por um par básico de elementos estruturantes (veja tabela 3.1).

Tabela 3.1- Dois pares de elementos estruturantes do esqueleto

0	0	0	1	1	0	0	0	0	0	1	0
0	1	1	1	0	0	1	1	1	0	0	0
0	1	0	0	0	0	0	1	0	0	0	0

Como resultado da operação esqueleto, temos a imagem mostrada na figura 3.26(a). Nesta figura vemos que o esqueleto contém diversos ramos parasitas, que são ramificações que partem das linhas principais(veja figura 3.26(b)). Além disso, existem diversos pontos isolados, caracterizados por possuírem dimensão de 1(um pixel). Esses ramos parasitas e pontos isolados podem ser facilmente confundidos com minúcias, causando erro na interpretação.

A técnica adotada para eliminar ramos parasitas é utilizar a operação afinamento com a função de eliminar os pontos final de linha. Se realizarmos esta operação inúmeras vezes de forma que a cada vez um único pixel seja eliminado de cada final de linha, chegaremos a uma situação em que a linha que caracteriza o ramo foi totalmente destruída. Os elementos estruturantes, mais uma vez, possuem papel decisivo para a realização desta tarefa, sua configuração está na tabela 3.2 .

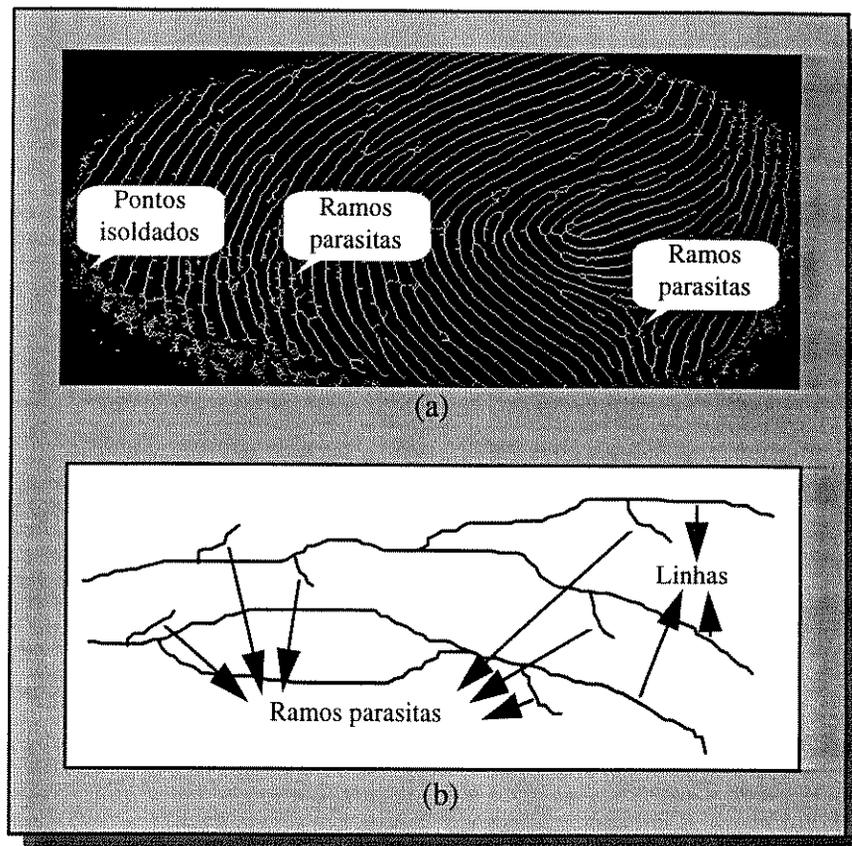


Figura 3.26- Esqueleto com ramos parasitas

Tabela 3.2- Elementos estruturantes para ramos parasitas

0	1	0	0	0	0
0	1	0	1	0	1
0	0	0	1	1	1

Apesar dos elementos estruturantes estarem adequados a nossa tarefa, não há como impedir que todas as vezes quando eliminamos pontos extremos dos ramos parasitas, estejamos também eliminando pontos extremos de linhas da impressão digital que estejam interrompidas por algum motivo, ou também pontos extremos de linhas que estejam caracterizando uma ilha. Observe pela figura 3.27(a) e (b), como não há maneiras de diferenciarmos estas linhas dos ramos parasitas.

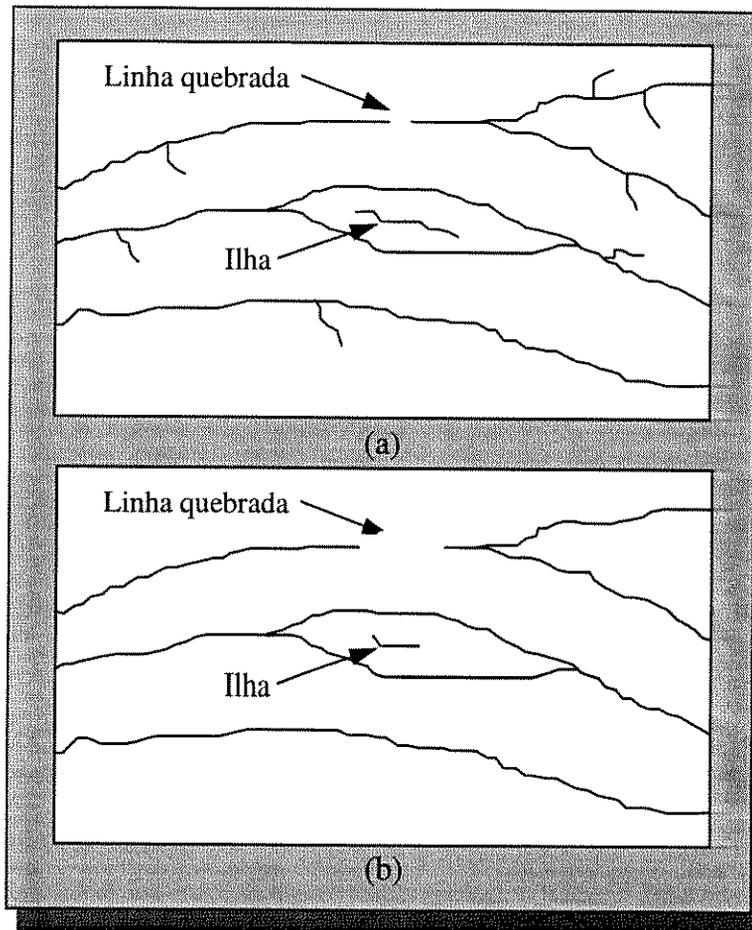


Figura 3.27- Falsos ramos parasitas

Para eliminarmos pontos isolados, utilizamos também a operação afinamento, com um novo conjunto de elementos estruturantes, mostrado na tabela 3.3 .

Tabela 3.3- Elementos estruturantes para pontos isolados

0	0	0	1	1	1
0	1	0	1	0	1
0	0	0	1	1	1

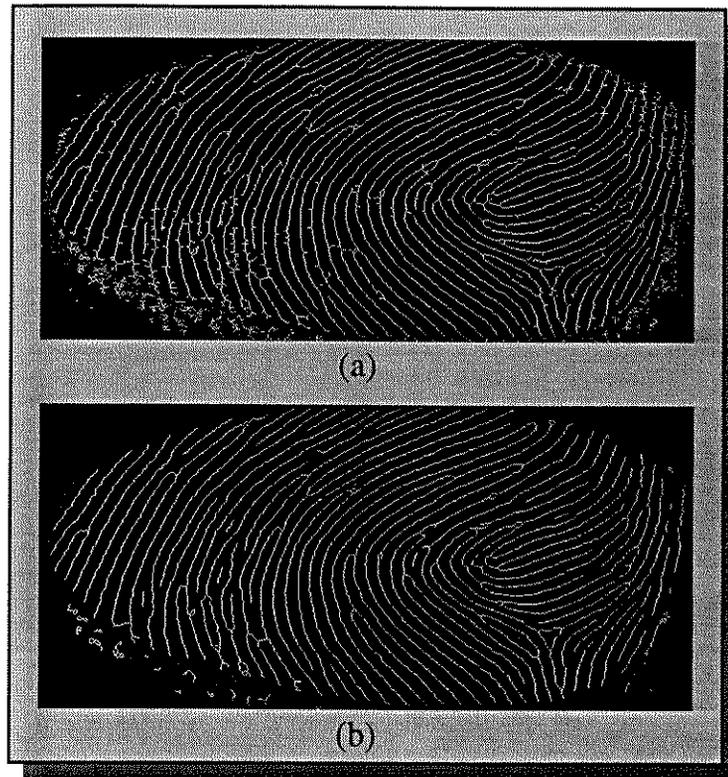


Figura 3.28- Resultado final

Se compararmos o resultado da figura 3.28(a), imagem com ramos parasitas e pontos isolados, com a figura 3.28(b), imagem final, veremos a grande diferença em termos de qualidade. Observe que ainda existe algum ruído ao redor da impressão digital. Porém, para eliminarmos este ruído teríamos que realizar uma quantidade muito grande de afinamentos que poderiam corroer muito as linhas da impressão digital.

Veja na figura 3.29 nossa cadeia final de processamento de impressões digitais.

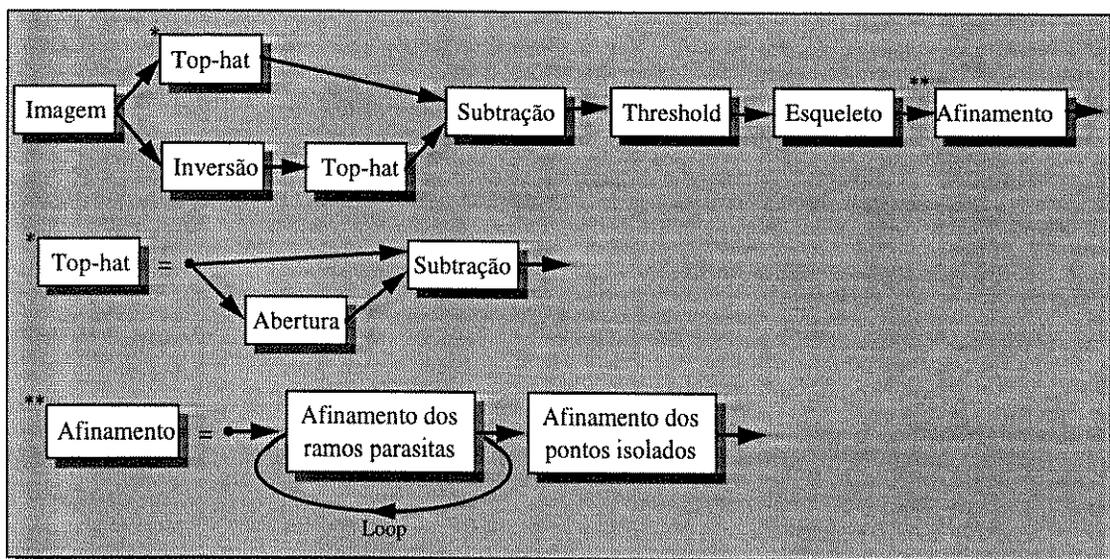


Figura 3.29- Cadeia de processamento

Capítulo 4

Extração de Características

4.1 Introdução

Este capítulo tem por objetivo desenvolver técnicas de extração de características que sejam adequadas para a representação de uma impressão digital. Inicialmente são vistos certos conceitos muito utilizados neste capítulo e nos posteriores, em seguida são propostas várias características para a representação de uma impressão digital.

4.2 Padrões

Muita da informação que nos rodeia manifesta-se em forma de padrões[14,17]. A facilidade com que humanos classificam e descrevem padrões geralmente nos conduz a assumir incorretamente que esta capacidade possa ser facilmente automatizada. Reconhecimento de padrões, naturalmente, é baseado em padrões. Um padrão pode ser tão básico quanto um conjunto de medidas ou observações (pressão sanguínea, idade, peso, altura), talvez representado numa notação de vetor ou matriz.

4.3 Características

Características[14,17] são qualquer medidas extraídas de um padrão. Um exemplo de características de um nível mais baixo são as intensidades dos sinais. Características podem ser simbólicas ou numéricas, por exemplo, cores e dimensões, respectivamente. Características podem também ser resultado da aplicação de um algoritmo de extração ou operador. Adicionalmente, características podem ser de nível mais alto: por exemplo, descrições geométricas de uma região ou um objeto em 3-D aparecendo numa imagem. Observe que: (1) esforços significantes podem ser requeridos

na extração de características; e (2) as características extraídas podem conter erros ou ruído. Características podem ser representadas por variáveis contínuas, discretas, ou discretas-binárias. Os problemas relatados de seleção e extração de características precisam ser resolvidos conforme as necessidades do sistema de reconhecimento de padrões. A chave é escolher e extrair características que (1) são computacionalmente flexíveis; (2) trabalhar com estes dados num bom sistema de reconhecimento; (3) reduzir os dados a uma quantidade de informação administrável sem descartar informações vitais.

4.4 Seleção de Características

Seleção de características é um processo de escolher dados de entrada para o sistema de reconhecimento de padrões e isto envolve julgamento. É importante extrair características que sejam relevantes na tarefa do reconhecimento. As características não devem ser de nível mais baixo. Porém, o nível das características determina a quantidade de pré-processamento da imagem necessária e pode também influenciar a quantidade de erro que é introduzido nestas características extraídas. A cadeia processamento de imagem desenvolvida para impressões digitais introduz erros que prejudicam algumas minúcias. Para entendermos melhor os erros introduzidos, vamos lembrar quais são as minúcias. Veja na figura 4.1(a) um exemplo de ilha; observe que definimos ilha como um ponto isolado. No ítem (b) temos a minúcia chamada ponto de bifurcação, e outro padrão chamado ponto final de linha que pode ser vista no ítem (c).

Muitas vezes nos capítulos anteriores citamos problemas em relação à aquisição de impressões digitais. Dentre todas os problemas citados, o que certamente produz mais erros é a pressão com que a digital é pressionada sobre a superfície do prisma. Se captarmos duas impressões digitais de uma mesma pessoa, teremos duas digitais diferentes em relação às minúcias. Dificilmente conseguiremos colocar a mesma pressão, ou melhor, dificilmente conseguiremos dipor a pressão igualmente sobre a superfície do dedo. É conveniente dizer que quando colocamos uma maior pressão, provocamos distorções do tipo: unir linhas. Isto se deve ao fato de que os vales cada vez mais se aproximam da superfície do prisma, chegando em certos pontos a tocá-la. Nestes pontos forma-se um elo de ligação entre duas linhas ou sulcos da impressão

digital. Esta ligação cria duas novas minúcias (pontos de bifurcação), a mais na impressão (veja figura 4.2(a)).

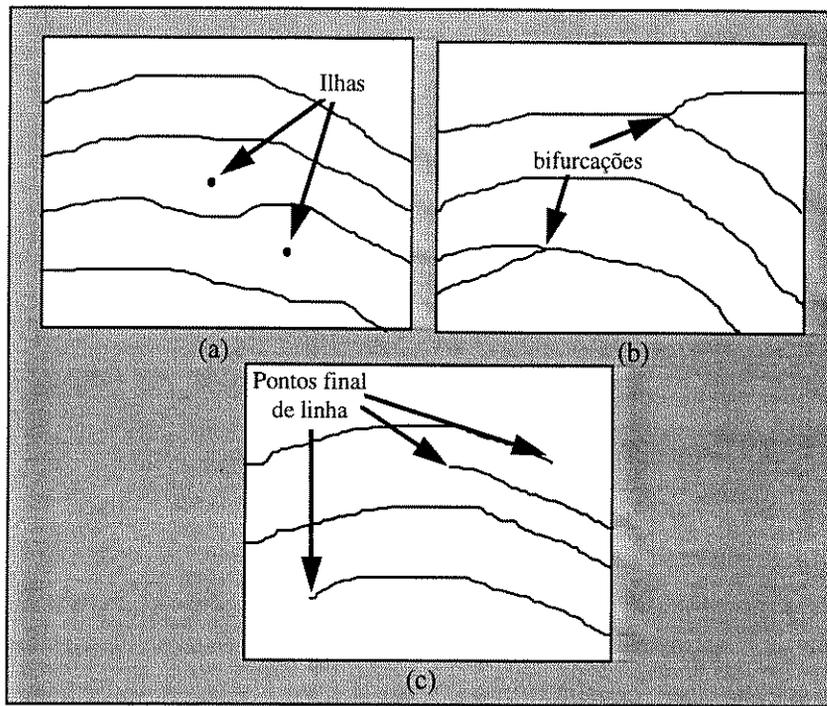


Figura 4.1- Minúcias

Em relação ao fato de colocarmos pouca pressão, podemos dizer que isso acarreta erros no sentido de que algumas linhas serão particionadas. Como resultado de uma linha quebrada teremos dois novos pontos de final de linha sendo criados (veja figura 4.2(b)).

Quando analisamos o padrão chamado ilha, o problema aumenta. Em geral, este padrão está presente em uma digital numa quantidade muito pequena, normalmente só 4 ou 5 ilhas. Ocorre que devido ao processamento surgem falsos pontos isolados, isto faz com que sua utilização como característica de uma impressão fique comprometida.

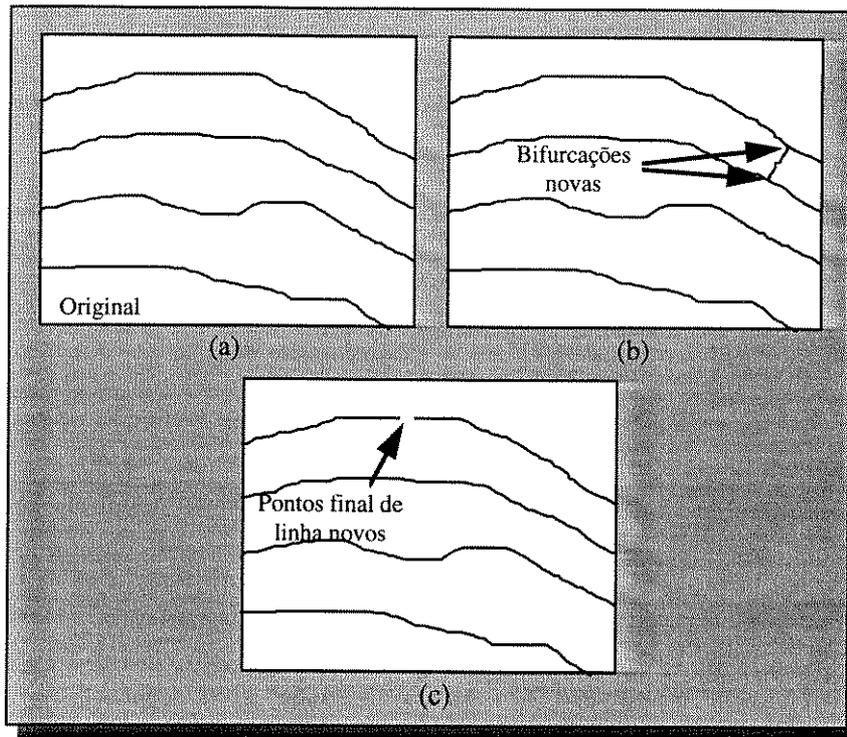


Figura 4.2- Novas minúcias

4.5 Vetor Distância

Precisamos, agora, traduzir padrões em medidas. Por isso trabalhamos para adotar um ponto que servisse como referência para criar um vetor distância entre este ponto e as minúcias. Resolvemos adotar como referência um ponto chamado centro e que está localizado no núcleo de uma impressão digital (veja figura 4.3). Existem cinco tipos de núcleos possíveis em impressões digitais. Em todos estes tipos, observamos uma característica similar; à medida que as linhas da impressão digital vão se aproximando deste núcleo, elas adquirem uma curvatura maior.

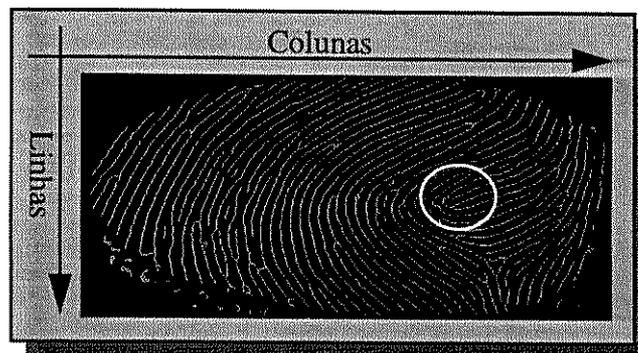


Figura 4.3 - Núcleo

Baseado neste fato, desenvolvemos diversos algoritmos com o objetivo de encontrarmos o núcleo, porém apenas um obteve resultados significativos. Este algoritmo funciona baseado no seguinte princípio: se para cada coluna da imagem traçarmos uma reta e contarmos quantas vezes esta reta corta as linhas da impressão digital (sulcos), teremos um vetor de quantidades unidimensional, que cresce e atinge seu máximo na região do núcleo, decrescendo à medida que se afasta do núcleo. A característica das linhas da impressão digital possuem curvatura maior na região do núcleo é responsável por isso. Se traçarmos uma reta, agora no sentido das linhas da imagem, e contarmos quantas vezes esta reta corta as linhas da impressão digital (sulcos), obteremos um outro vetor de quantidades unidimensional com comportamento similar. Veja na figura 4.4(a) um exemplo das linhas da impressão digital cortadas numa coluna da imagem. Na figura 4.4(b) temos um coluna da região do núcleo sendo amostrada. Observe como existem mais linhas da impressão sendo cortadas.

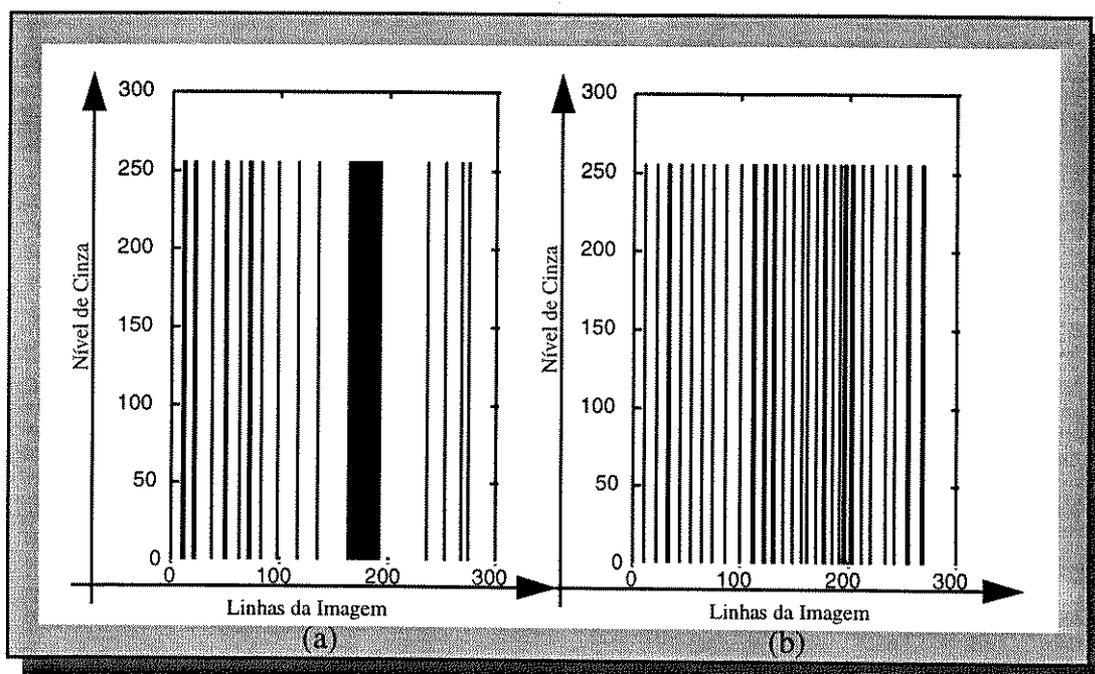


Figura 4.4- Exemplo dos valores de colunas da digital

Veja na figura 4.5(a)-(e) gráficos dos vetores obtidos de cinco digitais tiradas do polegar direito de uma mesma pessoa. A cada digital corresponde um gráfico do vetor obtido através das colunas e outro através das linhas da imagem. Observe a semelhança

na forma dos cinco vetores correspondentes as colunas e linhas da imagem. Em alguns casos a forma do gráfico aparece um tanto quanto deslocada, mas devemos lembrar que dificilmente a digital é colocada sobre a mesma posição na superfície do prisma. Isto acarreta pequenos deslocamentos.

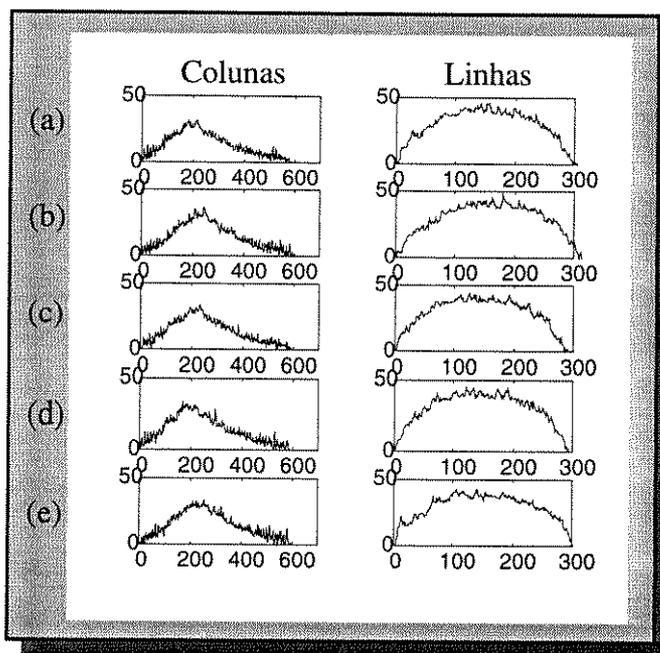


Figura 4.5- Vetores característicos

O ruído remanescente do processamento de imagens ao redor da impressão digital provoca transitórios nos vetores. Com o objetivo de eliminarmos distorções passamos os vetores por um filtro de mediana(veja apêndice A). O resultado está na figura 4.6, observe agora as digitais e seus correspondentes gráficos. Estas digitais estão em níveis de cinza, mas os vetores foram obtidos das correspondentes imagens binárias. Estamos utilizando imagens em nível de cinza apenas para facilitar a visualização do ponto que foi escolhido como centro, ponto preto na região do núcleo. O ponto escolhido como centro corresponde à coordenada de valor máximo dos vetores.

Podemos agora criar os vetores distâncias deste ponto central às posições das minúcias(bifurcações e pontos final de linha). Devemos tomar cuidados com o ruído que está presente ao redor da digital(comentado no final da seção 3.4), para isto delimitamos uma região quadrada centrada na coordenada central da digital. Somente as minúcias presentes que estão presentes dentro deste quadrado serão consideradas(veja figura 4.7). Em relação aos vetores, podemos utilizar a distância euclidiana, que leva nosso vetor a

ser variante à magnitude da digital, ou expressarmos esta distância em função das linhas da impressão digital cortadas entre a coordenada do núcleo e a coordenada das minúcias, que torna este vetor invariante a magnitude. Nos dois casos temos invariância à rotação.

Estes dois vetores distâncias, um em relação às bifurcações e outro em relação aos pontos final de linha, devem agora nos fornecer características que possam ser utilizadas no reconhecimento. Neste caminho, adotamos os seguintes parâmetros estatísticos que julgamos ser importantes: média, desvio padrão, mediana, valor máximo, soma dos elementos.

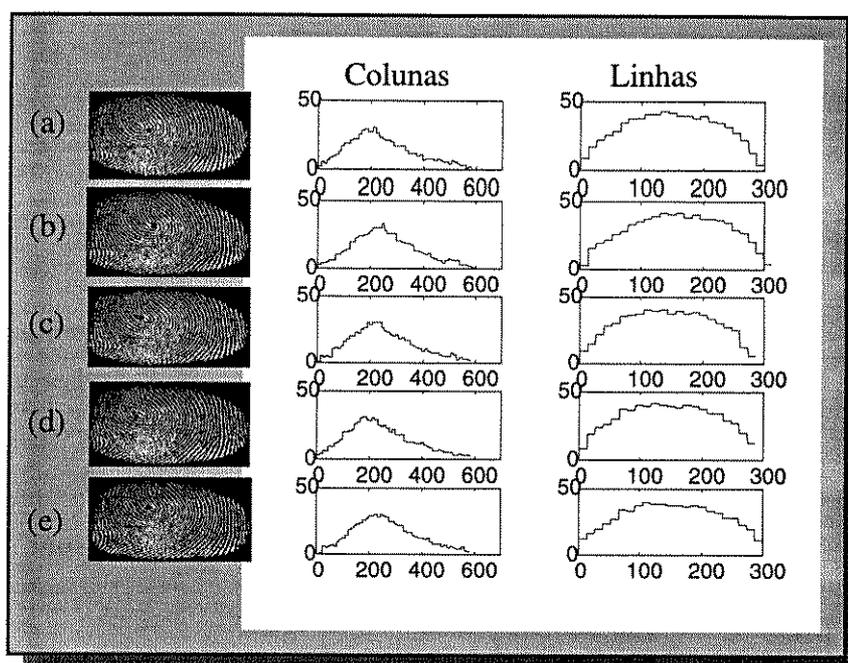


Figura 4.6- Vetores característicos filtrados

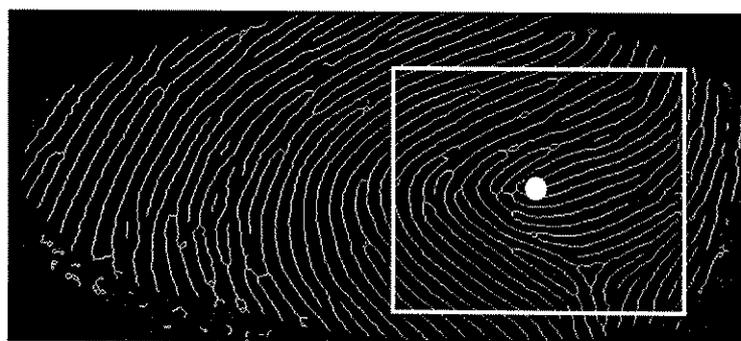


Figura 4.7- Região de interesse

4.6 Vetor Histograma

Outro padrão que adotamos para caracterizar as impressões digitais foram os vetores utilizados anteriormente para se localizar o centro. Observamos que eles possuem formas peculiares para cada pessoa. Destes vetores também retiramos os parâmetros estatísticos: média, desvio padrão, mediana, valor máximo, somados elementos. Chamaremos estes vetores de vetores histograma.

4.7 Quadrados Concêntricos

No último padrão utilizado, fizemos uma contagem da quantidade de pixels interiores a quadrados concêntricos. Imagine o centro da digital, agora tracemos um primeiro quadrado, interiormente a este quadrado teremos vários pixels que formam as linhas, ou pontos isolados da impressão digital. A contagem desses pixels nos dá uma quantidade. Se formos aumentando as dimensões deste quadrado, e a cada novo quadrado, contarmos a quantidade de pixels interior, teremos um vetor quantidades. Este vetor pode ser utilizado como característica no reconhecimento. Adotamos a forma quadrada, pois é de fácil implementação. Se utilizássemos um círculo teríamos que este vetor é totalmente invariante a rotação da impressão digital. Com a forma quadrada podemos dizer que o vetor é invariante a rotação com um certo erro associado. Veja na figura 4.8 como são dispostos este quadrados.

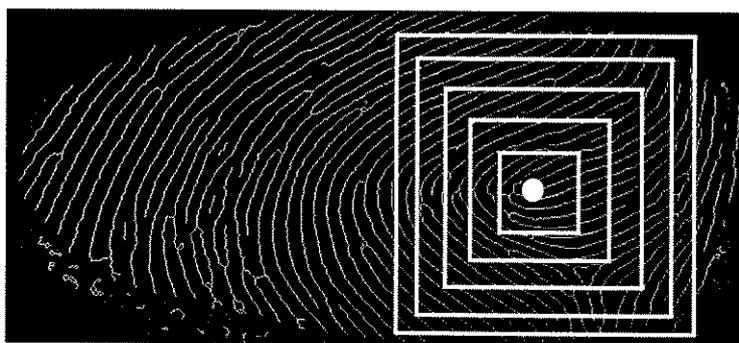


Figura 4.8- Quadrados concêntricos

Capítulo 5

Reconhecimento de Padrões

5.1 Introdução

Neste capítulo, começaremos explorando redes neurais com o objetivo de aplicarmos este classificador no reconhecimento. Após daremos uma introdução sobre método de correlação que também foi utilizado no processo de reconhecimento. Nos dois métodos mostraremos os resultados obtidos.

5.2 Redes Neurais

Sistemas biológicos realizam o reconhecimento de padrões via interconexão de células físicas chamadas neurônios. Esta propriedade forneceu motivação para que fosse tentado simular tal sistema computacionalmente. Um completo campo de estudos, centrado ao redor da criação e estudo de sistemas inteligentes pela formulação de sistemas computacionais que simulem a estrutura do funcionamento do cérebro humano, está emergindo com grande força nesta última década. Este movimento é conhecido por muitos nomes, incluindo Redes Neurais, Modelos Conexionistas, Modelamento Neuromorfo, e Processamento Paralelo Distribuído. A idéia de que podemos simular o comportamento humano, por meio de interações de um grande número de simples unidades de processamento é bastante nova. A adaptabilidade, tolerância ao erro, grande capacidade da memória, e a capacidade de processamento de informações em tempo real sugere uma arquitetura alternativa competitiva. Podemos dizer que o fato de que uma célula nervosa ser extremamente lenta (Faixa de milisegundos), mas o processamento no cérebro da informações ser relativamente rápido, nos leva a crer que a base computacional de nosso cérebro possui poucos passos em série, é maciçamente paralela.

5.2.1 Back-Propagation

O algoritmo de Back-Propagation[17,18,19] é hoje em dia o que está mais sendo utilizado. Este algoritmo foi inventado independentemente muitas vezes, por Bryson e Ho(1969), Werbos(1974), Parker(1985) e Rumelhart, Hinton, e Willinas(1986). Uma versão muito próxima a este algoritmo foi sugerida por Le Cun(1985). A base do funcionamento de uma rede Back-Propagation é um simples gradiente descendente. Considere primeiro, uma rede com duas camadas , tal como a ilustrada na figura 5.1 . Nossa convenção está mostrada nesta figura, as unidades de saída estão denotadas por O_i , as unidades escondidas por V_j , e os terminais de entrada por Y_k . Há conexões w_{jk} da entrada a camada intermediária, e W_{ij} da camada intermediária até à saída. Observe que o índice i sempre se refere a uma unidade de saída, j a uma unidade intermediária, e k a um terminal de entrada.

As entrada estão sempre fixadas a valores previamente definidos. Denotaremos os diferentes elementos por um subscrito μ , assim a entrada k é o conjunto y_k^μ . Os valores de entrada podem ser binários(0/1,ou ± 1) ou valores contínuos. Nós utilizamos N para denotar o número de unidades de entrada e p para o número de características de entrada($\mu = 1,2,\dots,p$).

Dado um elemento μ , a unidade intermediária j recebe da rede de entrada

$$h_j^\mu = \sum_k w_{jk} y_k^\mu \quad (5.1)$$

e produz a saída

$$V_j^\mu = g(h_j^\mu) = g\left(\sum_k w_{jk} y_k^\mu\right) \quad (5.2)$$

A unidade de saída i recebe assim,

$$h_i^\mu = \sum_j W_{ij} V_j^\mu = \sum_j W_{ij} g\left(\sum_k w_{jk} y_k^\mu\right) \quad (5.3)$$

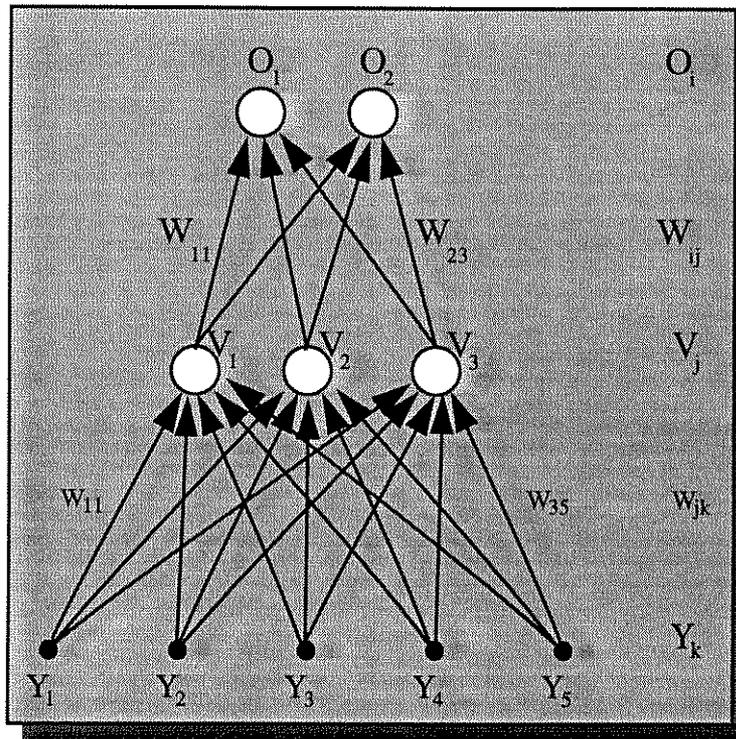


Figura 5.1- Rede de duas camadas com alimentação direta

e produz para como saída final

$$O_i^\mu = g(h_i^\mu) = g\left(\sum_j W_{ij} V_j^\mu\right) = g\left(\sum_j W_{ij} g\left(\sum_k w_{jk} y_k^\mu\right)\right). \quad (5.4)$$

Nossa função erro medida é a seguinte

$$E[W] = \frac{1}{2} \sum_\mu [\zeta_i^\mu - O_i^\mu]^2 \quad (5.5)$$

ou

$$E[W] = \frac{1}{2} \sum_\mu \left[\zeta_i^\mu - g\left(\sum_j W_{ij} g\left(\sum_k w_{jk} y_k^\mu\right)\right) \right]^2. \quad (5.6)$$

Esta é uma função contínua diferenciável em relação à W_{ij} e w_{jk} . Logo podemos utilizar um algoritmo de gradiente descendente para ensinar a atualizar os pesos. Podemos dizer

que isto é a essência do algoritmo de Back-propagation, mas há grande importância na forma com que as atualizações vão ser feitas.

Para as conexões da camada intermediária à camada de saída, as regras para o gradiente descendente nos dão

$$\begin{aligned}\Delta W_{ij} &= -\eta \frac{\partial E}{\partial W_{ik}} = \eta \sum_{\mu} [\zeta_i^{\mu} - O_i^{\mu}] g'(h_i^{\mu}) V_j^{\mu} \\ &= \eta \sum_{\mu} \delta_i^{\mu} V_j^{\mu}\end{aligned}\quad (5.7)$$

se definirmos os erros(ou deltas) δ_i^{μ} por

$$\delta_i^{\mu} = g'(h_i^{\mu}) [\zeta_i^{\mu} - O_i^{\mu}]. \quad (5.8)$$

Para as conexões da camada de entrada a camada intermediária Δw_{jk} devemos diferenciar, pois eles estão mais “profundos” na rede.

$$\begin{aligned}\Delta w_{jk} &= -\eta \frac{\partial E}{\partial w_{ik}} = -\eta \sum_{\mu} \frac{\partial E}{\partial V_j^{\mu}} \frac{\partial V_j^{\mu}}{\partial w_{jk}} \\ &= \eta \sum_{\mu} [\zeta_i^{\mu} - O_i^{\mu}] g'(h_i^{\mu}) W_{ij} g'(h_j^{\mu}) y_k^{\mu} \\ &= \eta \sum_{\mu} \delta_i^{\mu} W_{ij} g'(h_j^{\mu}) y_k^{\mu} \\ &= \eta \sum_{\mu} \delta_j^{\mu} y_k^{\mu}\end{aligned}\quad (5.9)$$

com

$$\delta_j^{\mu} = g'(h_j^{\mu}) \sum_i W_{ij} \delta_i^{\mu} \quad (5.10)$$

A constante η é chamado taxa de aprendizado. Observe que a equação 5.9 tem a mesma forma da equação 5.7, mas com a definição diferente de δ 's. Em geral, com um número arbitrário de camadas, a regra de atualização sempre tem a forma

$$\Delta w_{pq} = \eta \sum_{\text{elementos}} \delta_{\text{saida}} * V_{\text{entrada}} \quad (5.11)$$

onde a saída e entrada referem-se aos dois finais p e q da conexão feita. Na figura 5.2 temos o algoritmo Back-Propagation sendo utilizado para uma rede com três camadas. As linhas sólidas mostram a propagação para frente, enquanto as linhas pontilhadas mostram a propagação para trás do erro.

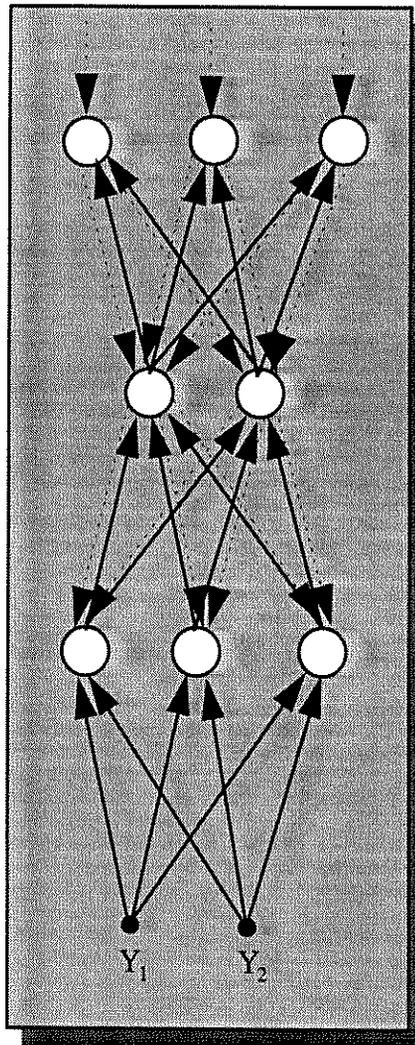


Figura 5.2- Back-Propagation

Apesar de havermos escrito as regras para a atualização, equações 5.7 e 5.9, como somas sobre todos os elementos μ , utilizaremos uma filosofia um pouco diferente: Um elemento μ é apresentado a entrada da rede e então todos os pesos são atualizados, antes do próximo elemento ser considerado. Isto diminui o gasto em tempo de cada passo. Podemos implementar este processo sugerindo que a ordem de apresentação dos elementos à entrada da rede seja aleatória. Se atualizarmos os pesos somente após todos os elementos passarem pela rede, teremos que ter um espaço adicional para armazenarmos variáveis. A eficiência relativa dos dois métodos depende do problema, mas o primeiro parece superior na maioria dos casos, especialmente para conjuntos com elementos regulares ou redundantes.

Podemos utilizar a função sigmóide (equação 5.12) ou também a função tangente hiperbólica (equação 5.13) como função de transferência. A função é claro deve ser diferenciável, e normalmente é desejável que ela sature nos dois extremos. Para as funções abaixo temos os seguintes extremos: 0/1 ou ± 1 , respectivamente.

$$g(h) = f_{\beta}(h) = \frac{1}{1 + \exp(-2\beta h)} \quad (5.12)$$

$$g(h) = \tanh \beta h \quad (5.13)$$

O parâmetro β é geralmente setado para 1, ou 1/2. As derivadas destas funções são expressas da seguinte forma $g'(h) = 2\beta g(1-g)$ para a equação 5.12 e $g'(h) = \beta(1-g^2)$ para a equação 5.13.

5.3 Variações sobre Back-Propagation

O algoritmo Back-Propagation tem sido muito estudado, e muitas extensões e modificações foram produzidas. O método acima dado é relativamente lento para convergir com múltiplas camadas, e muitas variações têm proposto torná-lo mais rápido. Outras melhoras incluem cuidados com pontos mínimos e na eficiência geral do método.

5.3.1 Momento

O gradiente descendente pode ser muito lento se η é pequeno, e pode oscilar largamente se η é muito grande. Há um grande número de maneiras de tratarmos com este problema, incluindo até a substituição do gradiente descendente por algoritmos mais sofisticados de minimização. Porém, uma maneira simples, é a adição de um termo chamado momento [Plaut et al., 1986][19], é geralmente eficiente e muito fácil de ser utilizado. A idéia é dar alguma inércia ou momento para cada conexão. Conforme o algoritmo converge para um ponto de mínimo esta inércia se ajusta de forma que não haja oscilação para se encontrar este ponto. A taxa de aprendizado efetiva pode ser aumentada sem provocar oscilações quando o algoritmo estiver longe de pontos de mínimo. Este esquema está implementado abaixo:

$$\Delta w_{pq}(t+1) = -\eta \frac{\partial E}{\partial w_{pq}} + \alpha \Delta w_{pq}(t). \quad (5.14)$$

O parâmetro do momento α precisa estar entre 0 e 1; o valor 0.9 é aconselhável.

Se o algoritmo estiver trabalhando numa região plana, então a derivada $\partial E / \partial w_{pq}$ será praticamente sempre a mesma a cada passo e a equação 5.14 convergirá para

$$\Delta w_{pq} \approx -\frac{\eta}{1-\alpha} * \frac{\partial E}{\partial w_{pq}} \quad (5.15)$$

com taxa efetiva de aprendizado de $\eta / (1-\alpha)$. Se por exemplo chegarmos a uma região de oscilação, w_{pq} responde somente com o coeficiente η . Podemos concluir que produzimos uma aceleração quando possível, mas quando chegamos numa região onde exista uma tendência a um mínimo esta aceleração diminui gradativamente até zero.

5.3.2 Parâmetros Adaptativos

Não é fácil escolher valores adequados para os parâmetros η e α para um problema particular. Além disso os melhores valores no começo do treinamento podem

não ser tão bons no final. Assim, muitos autores têm sugerido que estes parâmetros sejam ajustados automaticamente com um processo de aprendizado [exemplo, Cater, 1987; Franzini, 1987; Vogl et al., 1988; Jacobs, 1988][19].

O procedimento usual é checar se a atualização dos pesos realizada, reduziu a função erro. Se isto não aconteceu, então o processo não está convergindo, e η precisa ser diminuído. Se por outro lado, em alguns passos o erro diminuiu, então talvez nós estejamos sendo muito moderados, e poderíamos tentar aumentar η . Parece melhor incrementar η por uma constante, mas decrementar ele geometricamente para permitir um decaimento rápido quando necessário. Teremos então os seguintes casos:

$$\Delta\eta = \begin{cases} +a & \text{se } \Delta E < 0; \\ -b\eta & \text{se } \Delta E > 0; \\ 0 & \text{caso contrario} \end{cases} \quad (5.16)$$

onde ΔE é a média da função erro, e a e b são constantes apropriadas.

5.4 Correlação

Correlação[17] é uma simples e extremamente popular técnica de detecção de similaridades que é aplicável para sinais, vetores, e conjuntos. Um conjunto de padrões de referência é utilizado conjuntamente com um padrão desconhecido, frequentemente o padrão desconhecido pode estar contido dentro do conjunto de padrões de referência. Correlação entre o conjunto de padrões e o padrão desconhecido é achado pelo deslocamento deste sobre todas as possíveis localizações, utilizando uma métrica adequada, e calculando a função correlação a cada instante. Um exemplo visual de correlação está na figura 5.3 .

Definindo-se:

g : O conjunto de padrões que servem como referência.

f : O padrão desconhecido.

R : A extensão sobre a qual o casamento ocorre. Esta extensão, em algumas aplicações de reconhecimento de padrões , é todo g . Porém, isto pode ser apenas sobre uma pequena região .

As equações 5.17 ou 5.18 servem para calcular o casamento entre o padrão desconhecido com o conjunto de referência.

$$m_1 = \sum_R |f - g| \quad (5.17)$$

ou

$$m_2 = \sum_R (f - g)^2. \quad (5.18)$$

Intuitivamente, m_1 e m_2 serão pequenos(idealmente zero) quando f e g forem idênticos, e grandes quando eles forem significativamente diferentes.

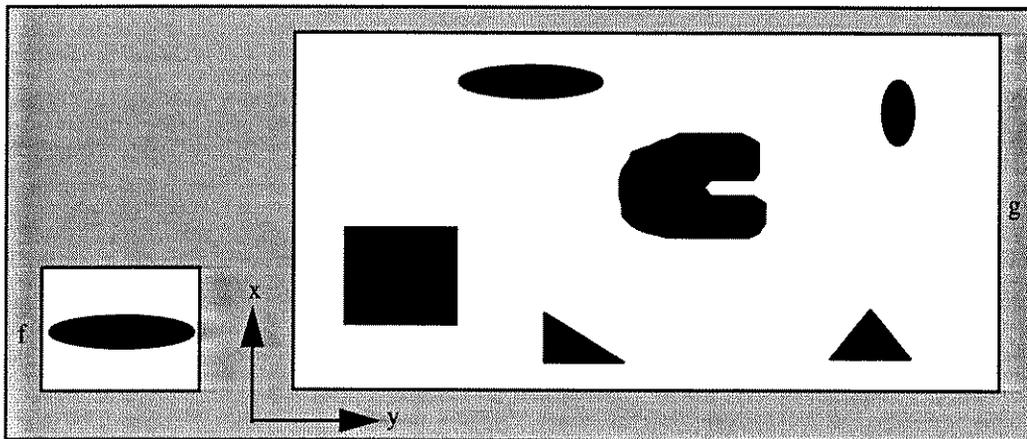


Figura 5.3- Exemplo de correlação

5.5 Reconhecimento por Redes Neurais

Inicialmente utilizamos redes neurais para o reconhecimento. O algoritmo utilizado foi Back-Propagation com momento e taxa de aprendizado variável. Aqui estão os parâmetros que foram utilizados na rede.

Número máximo de interações = 10000 ;

Erro a ser alcançado(desejado) = $E = 0.001$;

Taxa de aprendizado = $\eta = 0.045$;

Incremento na taxa de aprendizado = $a = 1.05$;

Decremento na taxa de aprendizado = $b = 0.7$;

Momento = $\alpha = 0.95$;

Número de entradas = $k = 10$;

Número de neurônios na camada intermediária = $j = 50$;

Função de transferência na camada intermediária = tangente hiperbólica;

Número de saídas = $i = 10$;

Função de transferência da camada de saída = sigmóide;

Em nosso experimento estamos utilizando as impressões digitais de dez pessoas. Para manter o sigilo em relação ao nome dos indivíduos que cederam suas impressões, designaremos as pessoas por letras do alfabeto de “A” até “J”, respectivamente. De cada pessoa captamos cinco vezes a impressão digital do seu polegar direito, logo temos um total de cinquenta impressões. O número de saídas da rede é correspondente ao número de pessoas, ou melhor, para cada pessoas temos um correspondente neurônio. O Neurônio mais ativado corresponde a pessoa que foi reconhecida. O número de neurônios da camada intermediária não possui regra para ser dimensionado, temos que observar a convergência da rede e conforme esta aumentamos ou diminuimos sua dimensão.

Para uma primeira série de treinamentos, adotamos os vetores utilizados na seção 4.4 para extração do centro da digital, ou vetores histograma. Cada pessoa tem dois vetores histogramas unidimensionais(veja figura 5.4), um em relação as colunas e outro em relação as linhas da imagem, e de cada um destes vetores retiramos cinco parâmetros; valor máximo, média, desvio padrão, soma dos elementos, e mediana(veja apêndice B). Isto totaliza dez valores que caracterizam cada pessoa.

No primeiro treinamento da rede neural utilizamos apenas uma digital por pessoa e outras duas digitais para teste de reconhecimento. Na figura 5.5 estão os gráficos do erro quadrático médio e taxa de aprendizado pelo número de interações. Veja através do gráfico do erro quadrático como a rede converge(erro quadrático tende ao valor desejado, 0,001). Após a rede ter convergido, devemos testar a eficiência da rede no reconhecimento. Para isto são utilizadas duas impressões digitais de cada pessoa, estas

digitais serão sempre utilizadas para testes em todos os posteriores treinamentos da rede neural. As digitais são injetadas na entrada da rede e são verificados os valores nas correspondentes saídas.

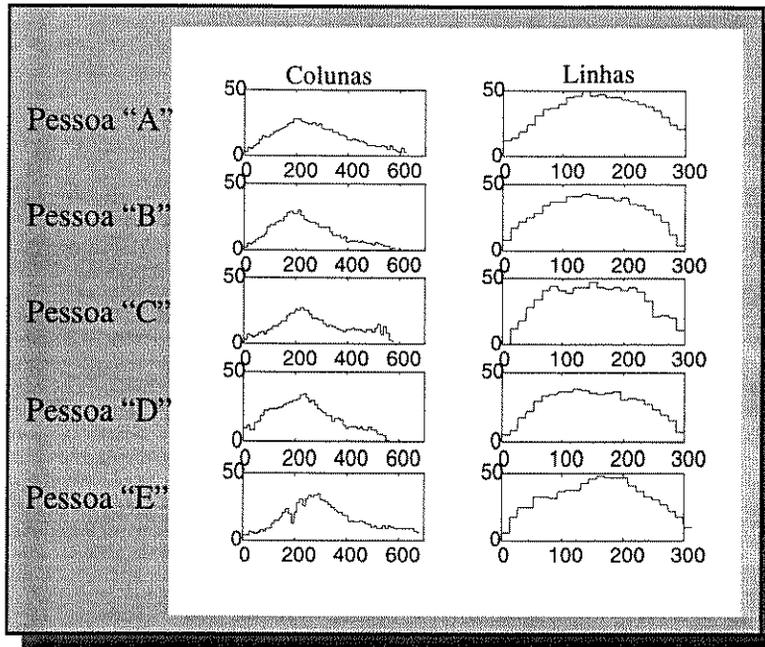


Figura 5.4- Exemplo de vetores característicos

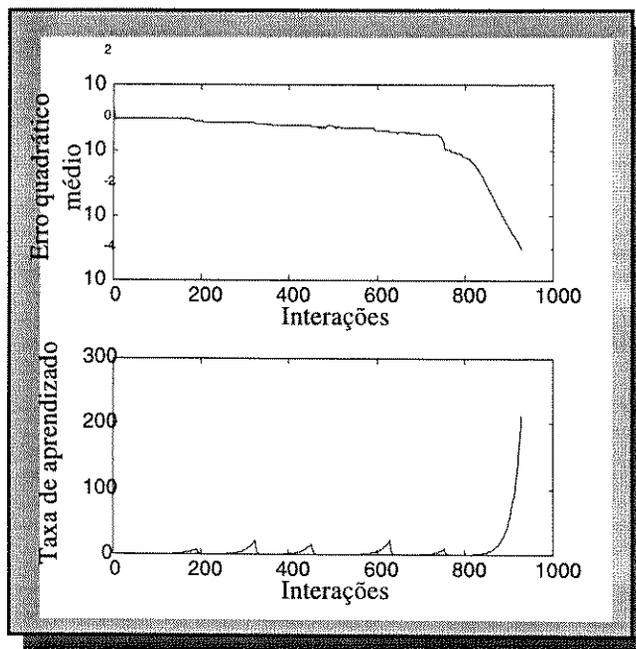


Figura 5.5- Gráficos de erro quadrático médio e taxa de aprendizado

Foram cometidos seis enganos no reconhecimento(30% de erro). Veja na tabela 5.1 os erros cometidos.

Tabela 5.1- Erros cometidos

Primeira série de digitais injetadas na rede		Segunda série de digitais injetadas na rede	
Correta	Reconhecida	Correta	Reconhecida
indivíduo "D"	indivíduo "C"	indivíduo "B"	indivíduo "A"
		indivíduo "C"	indivíduo "B"
		indivíduo "D"	indivíduo "C"
		indivíduo "H"	indivíduo "J"
		indivíduo "J"	indivíduo "H"

No próximo teste utilizamos duas impressões digitais para cada pessoa no treinamento da rede neural. Veja na figura 5.6 os gráficos correspondentes ao erro quadrático e taxa de aprendizado. Observe como a convergência foi muito mais rápida do que a anterior.

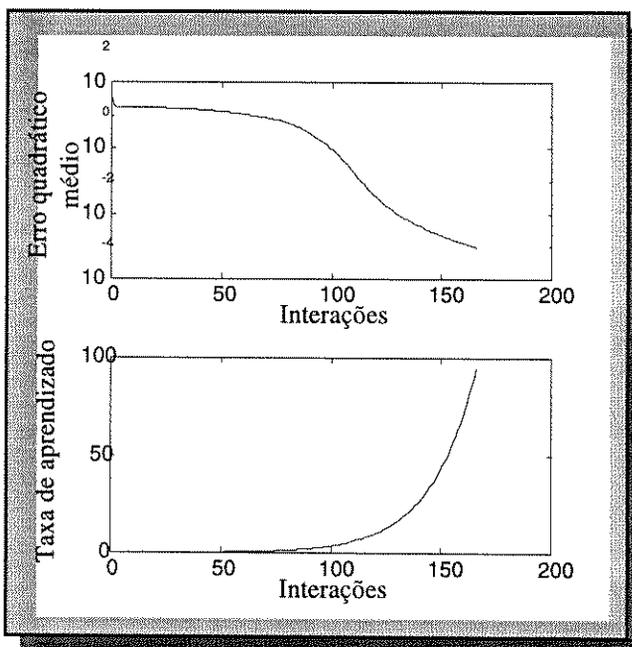


Figura 5.6- Gráficos do erro quadrático médio e taxa de aprendizado

Testamos a eficiência da rede da mesma forma que fizemos anteriormente, injetando as digitais de testes. Observamos que o número de equívocos diminuiu. Foram

cometidos quatro enganos no reconhecimento(20% de erro). Veja na tabela abaixo os erros cometidos.

Tabela 5.2- Erros cometidos

Primeira série de digitais injetadas na rede		Segunda série de digitais injetadas na rede	
Correta	Reconhecida	Correta	Reconhecida
indivíduo "D"	indivíduo "J"	indivíduo "C"	indivíduo "B"
		indivíduo "D"	indivíduo "J"
		indivíduo "H"	indivíduo "J"

Com o objetivo de diminuir ainda mais os erros cometidos no reconhecimento, treinamos a rede com três impressões por pessoa. Os resultados da convergência estão na figura 5.7 . Observe agora, que em relação ao treinamento anterior não houve melhora na convergência, porém foram cometidos três equívocos no reconhecimento(15% de erro). Veja tabela 5.3 quais foram os enganos.

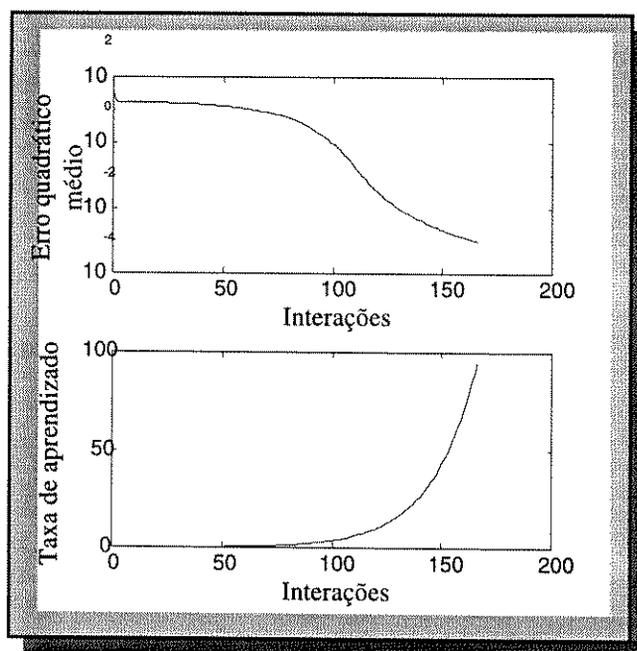


Figura 5.7- Gráficos do erro quadrático médio e taxa de aprendizado

Tabela 5.3- Erros cometidos

Primeira série de digitais injetadas na rede		Segunda série de digitais injetadas na rede	
Correta	Reconhecida	Correta	Reconhecida
indivíduo "C"	indivíduo "D"	indivíduo "C"	indivíduo "D"
		indivíduo "D"	indivíduo "G"

Na tabela 5.4 estamos mostrando a ativação dos neurônios de saída para a primeira série de digitais de testes injetadas na rede. O valor sombreado foi onde ocorreu o engano, observe que este valor corresponde ao neurônio do indivíduo "D", mas a digital é da pessoa "C".

Tabela 5.4- Ativação das saídas(neurônios)

Primeira série de digitais injetadas na rede										
	Neurônio 1="A"	Neurônio 2="B"	Neurônio 3="C"	Neurônio 4="D"	Neurônio 5="E"	Neurônio 6="F"	Neurônio 7="G"	Neurônio 8="H"	Neurônio 9="I"	Neurônio 10="J"
"A"	0,99	0,1	0	0	0	0	0	0	0	0
"B"	0	0,99	0	0	0	0	0	0	0	0
"C"	0	0	0,61	0,88	0	0	0	0	0	0
"D"	0	0	0	0,2	0	0	0	0	0	0
"E"	0	0	0	0	1,0	0	0	0,27	0	0
"F"	0	0	0	0	0	0,98	0	0,06	0,01	0
"G"	0	0	0	0	0	0,04	0,93	0	0,01	0
"H"	0	0	0	0	0	0	0	0,99	0	0
"I"	0	0	0	0	0	0,04	0	0	1,0	0
"J"	0	0,2	0	0,3	0	0	0,4	0	0	1,0

Em nosso próximo passo, extraímos novos parâmetros dos vetores histogramas. Estes parâmetros foram escolhidos de tal forma que caracterizam melhor a forma dos vetores histograma, trabalhamos com quartiz(veja apêndice A). Foram retirados cinco novos parâmetros que adicionados aos anteriores totalizam dez valores para cada vetor histograma. Isto totaliza vinte valores que caracterizam cada pessoa.

Para o treinamento da rede neural, redimensionamos certos valores da rede neural:

Número de entradas = $k = 20$;

Número de neurônios na camada intermediária = $j = 70$;

Fizemos três treinamentos da rede neural, utilizando uma, duas e três digitais por pessoa. Para o treinamento com uma digital por pessoa a rede não conseguiu convergir. Na figura 5.8 estão os gráficos do erro quadrático médio e taxa de aprendizado dos dois outros processos. Podemos observar que praticamente não existe diferenças no processo de convergência. Porém, quando analisamos os enganos cometidos no reconhecimento, vemos que com três digitais por pessoa o processo comete menos equívocos. Para o treinamento com duas impressões por pessoa ocorreram quatro enganos(veja tabela 5.5) no reconhecimento(20% de erro), para três digitais por pessoa ocorreu um equívoco(veja tabela 5.6) no reconhecimento(5% de erro). Podemos observar que os enganos cometidos são constantes, ou melhor, sempre os mesmos equívocos são cometidos. Isto nos leva a crer que os novos parâmetros extraídos dos vetores são redundantes em relação a certas digitais.

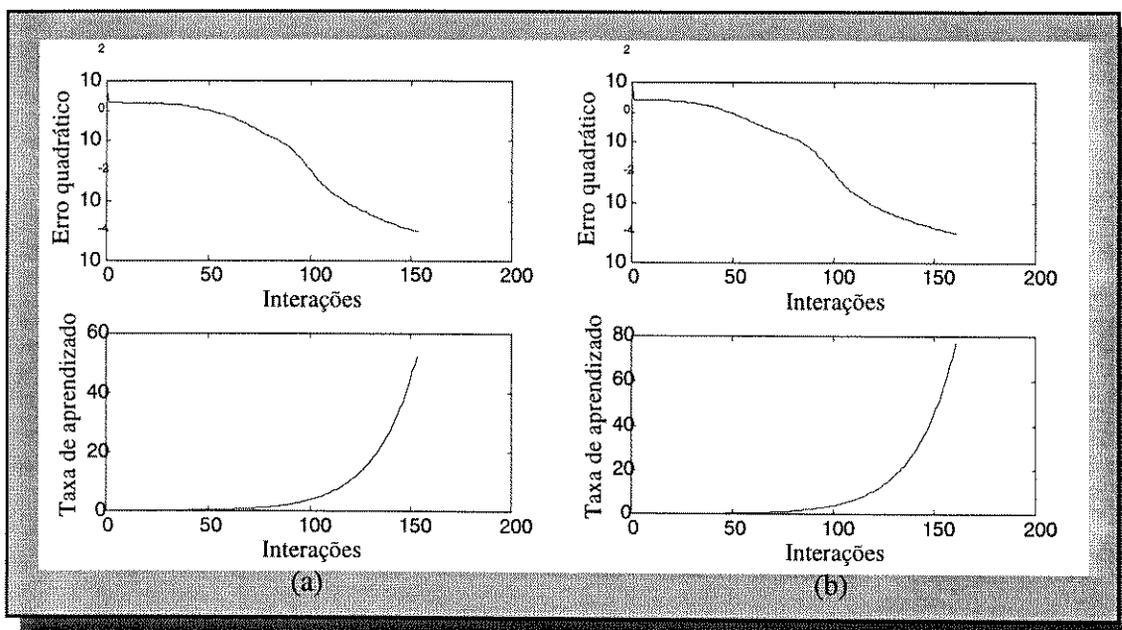


Figura 5.8- Gráficos do erro quadrático médio e taxa de aprendizado

Tabela 5.5- Erros cometidos

Primeira série de digitais injetadas na rede		Segunda série de digitais injetadas na rede	
Correta	Reconhecida	Correta	Reconhecida
indivíduo "D"	indivíduo "G"	indivíduo "D"	indivíduo "G"
indivíduo "H"	indivíduo "E"	indivíduo "H"	indivíduo "E"

Tabela 5.6- Erros cometidos

Primeira série de digitais injetadas na rede		Segunda série de digitais injetadas na rede	
Correta	Reconhecida	Correta	Reconhecida
		indivíduo "H"	indivíduo "E"

Podemos concluir que à medida que aumentamos a quantidade de digitais utilizadas no treinamento e utilizamos um número maior de parâmetros para caracterizar cada indivíduo obtemos maior confiabilidade no reconhecimento.

Em nosso próximo passo, utilizamos outro padrão para caracterizar as digitais. Este padrão foi comentado na seção 4.4, trata-se do vetor distância euclidiana entre o centro da digital e os pontos final de linha, pontos de bifurcação. Destes vetores distância extraímos basicamente quatro componentes, quantidade, média, desvio padrão e soma dos elementos. A rede possui a mesma configuração da anterior, apenas com a mudança na quantidade de entradas, $k = 8$. Utilizamos três digitais para o treinamento. Veja na figura 5.9 os gráficos do erro quadrático médio e taxa de aprendizado e na tabela 5.7 os equívocos cometidos no reconhecimento.

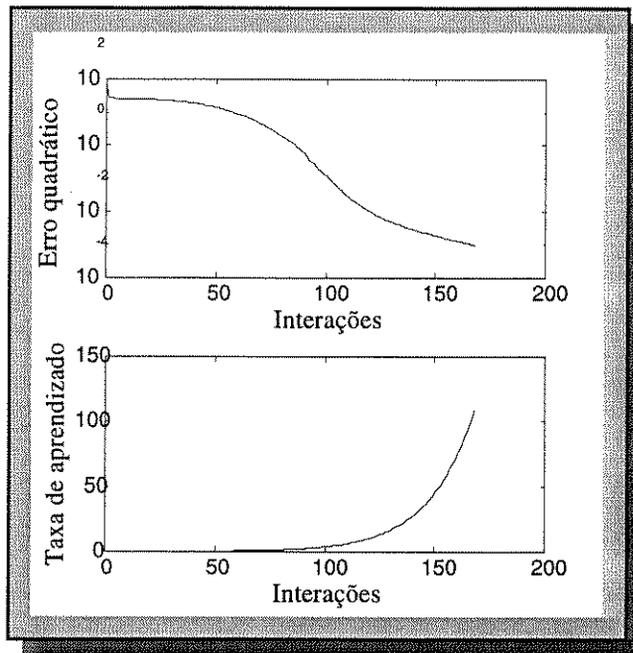


Figura 5.9- Gráficos do erro quadrático médio e taxa de aprendizado

Tabela 5.7- Erros cometidos

Primeira série de digitais injetadas na rede		Segunda série de digitais injetadas na rede	
Correta	Reconhecida	Correta	Reconhecida
indivíduo "B"	indivíduo "I"	indivíduo "D"	indivíduo "E"
indivíduo "C"	indivíduo "J"		
indivíduo "F"	indivíduo "I"		
indivíduo "G"	indivíduo "D"		

Observamos que ocorreram cinco enganos(25% de erro).

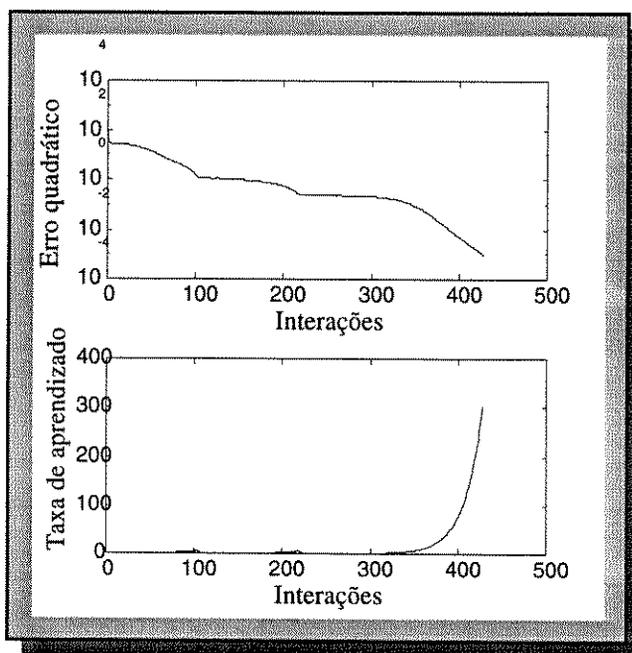
Haja visto que o vetor distâncias depende exclusivamente do centro da impressão digital, resolvemos sugerir centros para as digitais e apartir destes criamos novos vetores distância. Isto é necessário, pois os centro extraídos por nosso algoritmo dependem muito da qualidade da imagem e algumas digitais possuem uma qualidade ruim. Os parâmetros extraídos dos vetores distância foram os mesmos que os anteriores. O treinamento utilizou três impressões por pessoa e a mesma configuração da rede neural anterior. Na figura 5.10 temos o comportamento do erro quadrático médio e taxa de aprendizado, na tabela 5.8 estão os enganos cometidos no reconhecimento.

Tabela 5.8- Erros cometidos

Primeira série de digitais injetadas na rede		Segunda série de digitais injetadas na rede	
Correta	Reconhecida	Correta	Reconhecida
indivíduo "D"	indivíduo "E"	indivíduo "D"	indivíduo "E"
indivíduo "F"	indivíduo "H"	indivíduo "F"	indivíduo "I"
indivíduo "J"	indivíduo "C"	indivíduo "J"	indivíduo "C"

Ocorreram seis equívocos(30% de erro), porém temos agora padrões de erro que se repetem.

Realizamos teste de treinamento com apenas pontos de bifurcação, a rede neural conseguiu convergir com 100 neurônios na camada intermediária. Veja na figura 5.10 como a função erro demorou para convergir para o erro desejado. Ocorreram nove enganos no reconhecimento(45% de erro).

**Figura 5.10-** Gráficos do erro quadrático médio e taxa de aprendizado

A distância euclidiana, como foi dito na seção 4.4, é invariante a rotação. Para testarmos esta invariância, rotacionamos algumas impressões em 90°, extraídos os vetores distância e deste retirados os parâmetros que foram posteriormente jogados numa rede treinada com digitais não rotacionadas. Nenhum novo erro foi produzido.

No último reconhecimento de impressões digitais utilizando redes neurais, adotamos como padrão o vetor quantidades produzido pelos método dos quadrados

concêntricos(veja seção 4.4). Como primeiro experimento, fixamos as dimensões do quadrado inicial no valor de 20x20 pixels, e a cada novo quadrado produzido acrescentamos 20 pixels à sua dimensão. Foram produzidos 10 quadrados, logo, nosso vetor quantidades unidimensional possui magnitude igual a dez. Veja na figura 5.11 os gráficos do erro quadrático médio e taxa de aprendizado, observe que se comparado com experimentos anteriores houve uma certa demora na convergência. Podemos estimar que isto se deva em parte aos vetores não caracterizarem muito bem as digitais. Isto é, os valores que representam os diferentes indivíduos são semelhantes. Na tabela 5.9 estão os erros cometidos no reconhecimento(20% de erro).

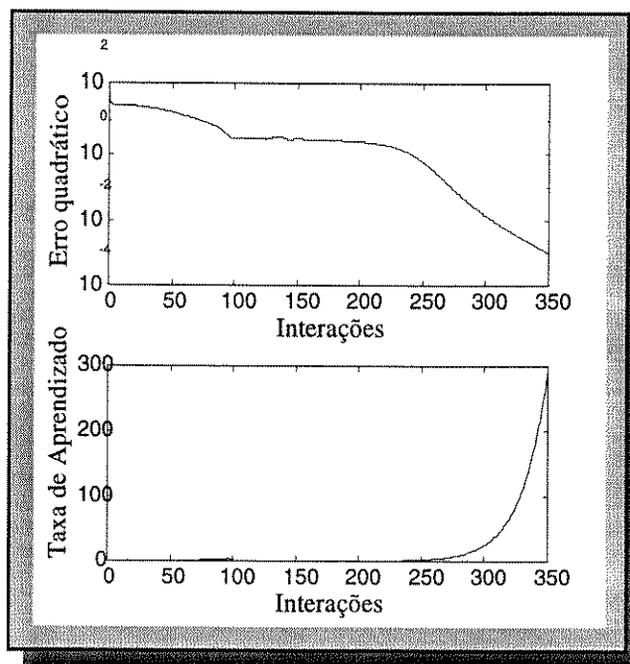


Figura 5.11- Gráficos do erro quadrático médio e taxa de aprendizado

Tabela 5.9- Erros cometidos

Primeira série de digitais injetadas na rede		Segunda série de digitais injetadas na rede	
Correta	Reconhecida	Correta	Reconhecida
indivíduo "I"	indivíduo "H"	indivíduo "D"	indivíduo "C"
		indivíduo "I"	indivíduo "J"
		indivíduo "J"	indivíduo "C"

Com o objetivo de melhorar o vetor característico, adotamos uma maior quantidade de quadrados concêntricos, isto aumentou a dimensão do vetor. Estamos utilizando, agora, 20 quadrados concêntricos, com incremento de cinco pixels nas dimensões de quadrado para quadrado. Veja na figura 5.12 o gráfico do erro quadrático médio e taxa de aprendizado, observe que a convergência melhorou um pouco em relação ao treinamento anterior. Na tabela 5.10 estão os enganos cometidos no reconhecimento(20% de erro). Nos dois experimentos realizados com quadrados concêntricos a configuração básica da rede foi a mesma das anteriores, apenas foram alteradas a quantidade de entrada e camada intermediária para 70 neurônios, $j = 70$.

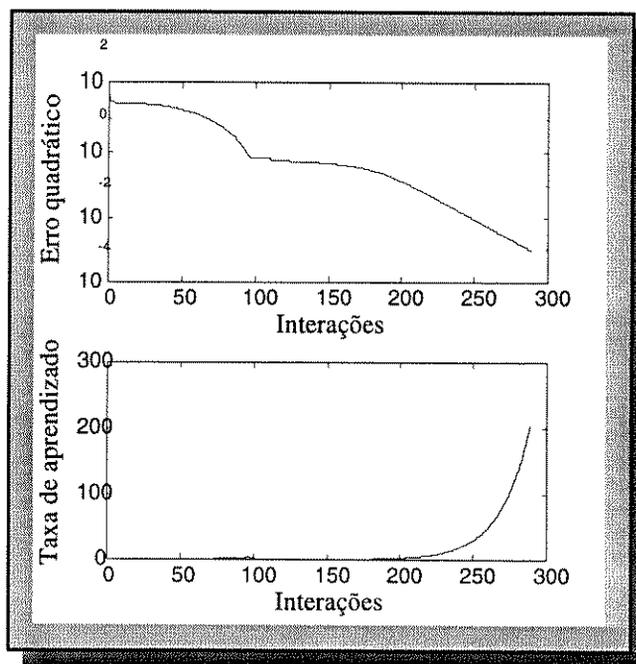


Figura 5.12- Gráficos do erro quadrático médio e taxa de aprendizado

Tabela 5.10- Erros cometidos

Primeira série de digitais injetadas na rede		Segunda série de digitais injetadas na rede	
Correta	Reconhecida	Correta	Reconhecida
indivíduo "D"	indivíduo "B"	indivíduo "D"	indivíduo "B"
indivíduo "I"	indivíduo "H"	indivíduo "J"	indivíduo "C"

5.6 Reconhecimento por Correlação

Classificação utilizando correlação não requer qualquer tipo especial de vetores característicos. Com este classificador as impressões são comparadas diretamente. Uma impressão digital desconhecida é classificada da seguinte maneira. A correlação entre uma digital sob teste e uma digital do banco de dados é computada. Este processo é repetido para todas as impressões do banco de dados. A impressão sob teste é dita pertencer a uma classe quando o valor da correlação com esta classe for o maior de todos. Este mecanismo não provê meios de invariância a rotação e magnitude. A não ser que você crie um banco de dados com rotações e magnitudes variadas. Utilizamos uma impressão digital de cada pessoa no banco de dados e as outras 40 para o reconhecimento. Veja na tabela 5.11 como foram os valores percentuais de correlação entre as digitais do banco de dados e as digitais sob teste. Na tabela 5.12 estão os três enganos cometidos(7,5% de erro) no reconhecimento das 40 digitais desconhecidas.

Tabela 5.11- Desempenho do classificador por correlação

Indivíduos do banco de dados	Correlação em relação a indivíduos de mesma classe	Correlação em relação a indivíduos de classes diferentes
"A"	30%	13%
"B"	30%	13%
"C"	18%	14,5%
"D"	45%	17%
"E"	41%	16,4%
"F"	30%	18,2%
"G"	67%	14,5%
"H"	25,4%	17%
"I"	38,5%	15,2%
"J"	50%	16%

Tabela 5.12- Erros cometidos

Primeira série de digitais		Segunda série de digitais		Terceira série de digitais		Quarta série de digitais	
Correta	Reconhecida	Correta	Reconhecida	Correta	Reconhecida	Correta	Reconhecida
		indivíduo "C"	indivíduo "B"	indivíduo "C"	indivíduo "B"	indivíduo "H"	indivíduo "I"

5.7 Comentário dos Resultados

Analisando o gráfico da figura 5.13, podemos observar que:

À medida que aumentamos a quantidade de elementos de uma mesma classe no treinamento da rede neural reduzimos a probabilidade de erro. Isto ocorre também quando aumentamos a quantidade de parâmetros que caracterizam as digitais. Os parâmetros devem ser escolhidos de forma que não sejam redundantes e caracterizem da melhor forma possível nosso padrão.

Nosso melhor resultado foi obtido utilizando-se o vetor histograma com vinte parâmetros de entrada. Este vetor histograma possui invariância a pequenas rotações e mostrou ser de grande utilidade não somente no reconhecimento, como também na localização dos centros das digitais. É um algoritmo de fácil implementação e rápido.

Em relação ao vetor distâncias, ele se baseia nas minúcias das impressões. Quando a qualidade da imagem é boa nosso vetor distâncias será bom, porém quando a imagem diminui em qualidade este vetor diminui proporcionalmente. Em nosso caso, possuímos digitais com qualidade variada. Certamente, se estivessemos trabalhando com imagens de alta qualidade este vetor seria o mais adequado para o reconhecimento. Devemos lembrar que ele possui invariância a rotação e magnitude.

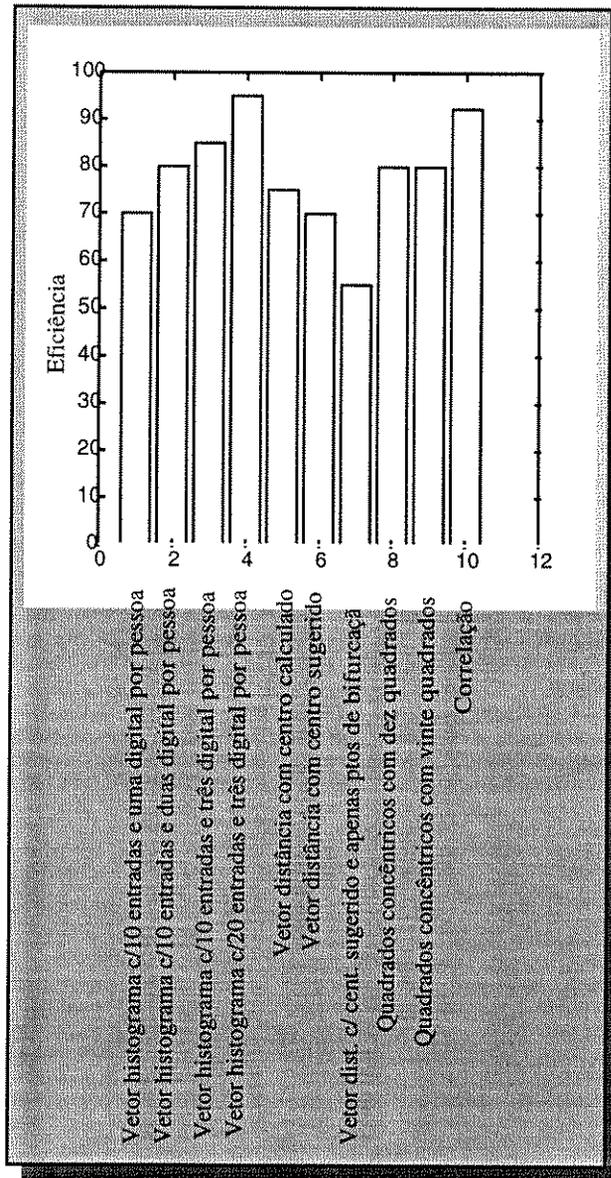


Figura 5.13- Gráfico de desempenho

O vetor produzido pelo método dos quadrados concêntricos mostrou ser outro caminho para o reconhecimento. Podemos destacar como pontos positivos sua invariância a rotação, fácil implementação e ser rápido.

Em relação ao reconhecimento por correlação, podemos dizer que apresentou resultados muito bons. Porém possui pontos extremamente negativos; consome muito tempo no processamento, não possui invariância em relação a rotação e magnitude.

Capítulo 6

Conclusões

Neste trabalho desenvolvemos um sistema de automático de reconhecimento de impressões digitais. Este sistema executa os seguintes processos: aquisição da imagem, processamento, extração de características e classificação.

Em relação ao sistema de aquisição de imagens, procuramos mostrar os caminhos que estavam disponíveis a nós. O método do prisma mostrou ser o mais adequado. Procuramos, então aprimorar o método realizando certas alterações. O resultado foi uma melhora significativa no seu desempenho.

A cadeia de processamento de imagem foi totalmente desenvolvida para o uso em impressões digitais. Tivemos grandes dificuldades para eliminarmos ou pelo menos minorarmos os problemas associados as imagens das impressões. Boa parte de nosso tempo foi empregada nesta etapa. Os resultados obtidos foram muito bons se comparados com algoritmos sugeridos em artigos[6].

Quatro esquemas de classificação foram investigados. Em três desses utilizamos redes neurais do tipo Back-Propagation. Os parâmetros utilizados no reconhecimento foram extraídos de padrões, criamos estes padrões de maneira a caracterizar o melhor possível as digitais. A estes padrões foram dados os seguintes nomes; vetor histogramas, vetor distâncias, quadrados concêntricos. Último, porém não menos importante, foi a utilização do classificador por correlação. Todos eles trabalharam sobre imagens esqueletonizadas.

A melhor performance obtida foi de 95% com a utilização do vetor histogramas associados a redes neurais do tipo Back-propagation com momento e taxa de aprendizado variável. Este valor de acerto pode ser considerado muito bom levando-se que nosso método de aquisição de imagens é considerado na literatura como um método de baixa qualidade. Obtivemos muito bons resultados também com o classificador por correlação, mas como dito anteriormente no capítulo 5, ele possui muitos inconvenientes.

Podemos concluir que este trabalho contribuiu significativamente na sua área de interesse, pois através dele novos algoritmos de processamento de imagens, localização do centro da impressão digital, extração de características foram criados, bem como uma melhora no método de aquisição.

Nossas sugestões para trabalhos futuros:

- Melhorarmos o sistema de aquisição, buscando adquirirmos imagens de impressões digitais em tempo real com uma qualidade superior ao que estamos trabalhando hoje. Acreditamos que a qualidade dos dados de entrada possuem uma importância crucial no desempenho das redes de classificação. Estamos hoje tentando adquirir dados referentes a um sensor óptico-eletrônico especialmente construído para captar imagens de impressões digitais.
- Devemos melhorar o algoritmo de obtenção do centro das digitais, torná-lo totalmente invariante a rotação e magnitude.
- Extrairmos padrões que representem melhor as impressões digitais, padrões mais robustos, consistentes.
- Desenvolvermos um sistema para compressão de imagens de impressões digitais.
- Criação de um banco de dados para impressões digitais.

Apêndice A

Filtro de Mediana

O objetivo do filtro de mediana[14] é eliminar o ruído presente na imagem. Imagine um vetor v composto por l elementos. O valor de cada elemento de uma vizinhança pré-estipulada, é substituído pelo valor da mediana desta vizinhança.

A mediana[20], m , de um conjunto de valores é tal que a metade dos valores é inferior a ela e metade dos valores superior. Para fazermos o filtro de mediana numa vizinhança de um elemento, nós primeiros devemos ordenar crescentemente os valores desta vizinhança, determinamos a mediana, e designamos a todos os outros elementos da vizinhança o valor da mediana.

Apêndice B

Parâmetros Estatísticos^[20]

Soma dos elementos- Como o próprio nome diz, é uma operação onde somamos os valores do vetor.

$$\sum_{j=1}^n x_j \quad (\text{b.1})$$

Média- A mais comum medida para localizar o centro de dados, nada mais é do que uma média aritmética.

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (\text{b.2})$$

Mediana- Outra medida da tendência central é a mediana, ou o ponto no qual as amostras estão divididas em duas medidas iguais. Seja $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ as amostras de um vetor arranjadas em ordem crescente de magnitude. Então, a mediana \tilde{x} é definida como a metade ou $([n+1]/2)$ th observação se n é ímpar, e a média aritmética entre os elementos $(n/2)$ e $([n]/2+1)$ se n for par. Expressando matematicamente, temos,

$$\tilde{x} = \begin{cases} x_{([n+1]/2)}, & n \text{ ímpar} \\ \frac{x_{(n/2)} + x_{([n/2]+1)}}{2} & n \text{ par} \end{cases} \quad (\text{b.3})$$

Valor máximo- É o máximo valor que acontece no vetor.

$$\max(x_j) \quad (\text{b.4})$$

Desvio Padrão- É a mais importante medida de variabilidade. Se $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ é uma amostra de n observações, então o desvio padrão é

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad (\text{b.5})$$

Quartil- A média divide os dados em duas partes iguais. Nós podemos dividir os dados em mais do que duas partes. Quando um conjunto de dados é dividido em quatro partes iguais, esta divisão é chamada de quartil. Podemos, agora, tratar os dados como quatro conjuntos de dados e deles extrair vários parâmetros estatísticos.

Bibliografia

- [1] **T.Ch; Malleswara Rao**, Feature extraction for fringerprint classification, *Pattern Recognition* 8, p. 181-192, 1976.
- [2] **B.Moayer; K.S.Fu**, A tree system approach for fringerprint pattern recognition, *IEEE Trans. Pattern Analysis Mach. Intell. PAMI-8(3)*. p. 376-387, may 1986.
- [3] **M.R.Verma; A.K.Majumdar, B.Chatterjee**, Edge detection in fingerprint, *Pattern Recognition* 20, p. 513-523, 1987.
- [4] **A.K.Hrechak; J.A.McHugh**, Automated fingerprint recognition using structural matching, *Pattern Recognition* 23, p. 893-904, 1990.
- [5] **A.D.Brink**, Grey level thresholding of images using a correlation criterion, *Pattern Recognition Lett.* 9, p. 335-341, 1989.
- [6] **L.Coetzel; E.C.Botha**, Fingerprint recognition in low quality images, *Pattern Recognition* 26, No 10, p. 441-1460, 1993.
- [7] **V.S.Srinivasan; N.N.Murthy**, Detection of singular points in fingerprint images, *Patter Recognition* 25, No.2, p.139-153, 1992.
- [8] **D.C.Douglas Hung**, Enhancement and feature purification of fingerprint images, *Pattern Recognition* 26, No.11, p. 1661-1671, 1993.
- [9] **A.K.Hrechak; J.A.Mchugh**, Automated fingerprint using structural matching, *Pattern Recognition* 23, No.8, p. 893-904, 1990.
- [10] **B.M.Mehtre**, Segmentation of fingerprint images-A composite method, *Pattern Recognition* 22, No.4, p. 381-385, 1989.
- [11] **B.M.Mehtre; N.N.Murthy; S.Kapoor**, Segmentation of fingerprint using the directional image, *Pattern Recognition* 20, No.20, p. 429-435, 1987.
- [12] **Q.Xiao; H.Raafat**, Fingerprint image postprocessing: A combined statical and structural aproach, *Pattern Recognition* 24, No.10, p. 985-992, 1991.
- [13] **M.R.Verma; A.K.Majumdar; B.Chatterjee**, Edge detection in fingerprints, *Pattern Recognition* 20, No.5, p. 513-523, 1987.
- [14] **R.C. Gonzalez; R.E.Woods**, *Digital Image Processing*, Addison-Wesley Publishing Company, Massachusetts, 1993.
- [15] **A.Low**, *Introductory Computer Vision and Image Processing*, McGRAW-HILL Book Company, London, 1991.

- [16] **J.S.Lim**, Two-dimensional Signal and Image Processing, Prentice Hall, New Jersey, 1990.
- [17] **R.Schalkoff**, Pattern Recognition: Statical, Structural and Neural Approaches, John Wiley & Sons, Inc., Singapore, 1992.
- [18] **J.A.Freeman; D.M.Skapura**, Neural Networks-Algoritms, Applications, and Programming Techniques, Addison-Wesley Publishing Company, Massachusetts, 1991.
- [19] **J.Hertz; A.Krogh; R.G.Palmer**, Introduction to the Theory of Neural Computation, Addison-Wesley Publishing Company, Massachusetts, 1991.
- [20] **D.C.Montgomery; G.C.Runger**, Applied Statistics and Probability for Engineers, John Wiley & sons, Inc., New York, 1994.
- [21] **R.A.Serway**, Physics for Scientists & Engineers with Modern Physics, Saunders College Publishing, Philadelphia, 1990.
- [22] **G.D.Breda, L.L.Ling, R.B.Filho**, Processamento de Imagens de Impressões Digitais utilizando Morfologia Matemática, ELECTRO '95:XI congreso Chileno de Ingenieria Electrica, Punta Arenas - Chile, 11/95.