

Processamento e Análise de Imagens

Jacques Facon
Pontifícia Universidade Católica do Paraná
Curso em Mestrado em Informática Aplicada

February 21, 2002

Jacques Facon é professor e pesquisador no Departamento de Ciência da Computação e no Curso em Mestrado em Informática Aplicada da Pontifícia Universidade Católica do Paraná (PUC-PR) em Curitiba.

Formou-se em Engenharia Elétrica pelo ESIGElec, (Ecole Supérieure d'Ingénieurs en Génie Electrique) em Rouen na França, em 1984, obteve o grau de Mestre em Instrumentação e Medidas pela Université de Rouen na França em 1984 e o grau de Doutor em Processamento e Análise de Imagens pela Université de Technologie de Compiègne na França em 1987.

Suas áreas atuais de interesse incluem processamento de imagens, segmentação de imagens, morfologia matemática e reconhecimento de padrões.

PREFÁCIO

O campo do processamento de imagens consiste em desenvolver técnicas de compreensão dos mecanismos de visão natural e artificial e em fornecer ferramentas cujo desempenho aproxima-se o mais possível das extraordinárias capacidades do sistema visual humano. A área de processamento de imagens tem meios de agir através dos recursos do computador cujo progresso em termos de velocidade e de memória deixa prever novas aplicações em um futuro próximo.

A meta da área de processamento e análise de imagens é simular experiências tais como criar movimento, distorções, iluminação de objetos em uma cena de maneira a resolver o problema proposto e/ou comparar entidades específicas com padrões previamente armazenados. O computador deve ter os meios de sentir e interagir com o espaço de investigação para construir uma base de conhecimento e de experiência. Sem interação física, não é possível capturar e armazenar modelos comportamentais que reflitam a complexidade de uma realidade. Dentro deste ponto de vista, a disciplina de processamento de imagens é claramente uma área multidisciplinar envolvendo áreas diversas como Eletrônica, Ótica, Inteligência Artificial, Ciência da Computação e outras.

Sumário

1	INTRODUÇÃO	1
1.1	A VISÃO HUMANA	1
1.2	CARACTERÍSTICAS DO SISTEMA VISUAL HUMANO	3
1.3	CONCEPÇÃO DE UM SISTEMA DE VISÃO ARTIFICIAL	4
1.4	EXEMPLOS DE PROCESSAMENTO DE IMAGENS	5
1.4.1	Tarefas Industriais	5
1.4.2	Reconhecimento de padrões	5
1.4.3	Reconstrução tridimensional	6
1.5	ARQUITETURA DE UM SISTEMA DE VISÃO ARTIFICIAL . . .	6
2	FUNDAMENTAÇÃO	9
2.1	INTRODUÇÃO	9
2.2	A IMAGEM	9
2.2.1	Introdução	9
2.2.2	Imagem da Câmera	9
2.2.3	Imagem Digital	10
2.3	PROPRIEDADES DE UM PIXEL	12
2.3.1	Vizinhança de um Pixel	12
2.3.1.1	4-Vizinhança	12
2.3.1.2	8-Vizinhança	12
2.3.2	Conectividade	12
2.4	TRANSFORMAÇÕES GEOMÉTRICAS	13
2.4.1	Translação	13
2.4.2	Transformação de escala	14
2.4.3	Rotação	14
2.4.4	Transformação perspectiva	15
3	REALCE DE IMAGENS	19
3.1	INTRODUÇÃO	19
3.2	MÉTODOS “PONTO POR PONTO”	20
3.2.1	NOÇÃO DE HISTOGRAMA	21
3.2.1.1	Modificação de histograma	22
3.2.1.2	Equalização de histograma	22
3.2.1.3	Especificação de histograma	24
3.2.2	Aprimoramento Local	25
3.2.3	Modificação de contraste	26
3.3	MÉTODOS DE VIZINHANÇA	26
3.3.1	Redução de ruído pela média de múltiplas imagens	27
3.3.2	Filtragem	28
3.3.2.1	Transformada de Fourier	28
3.3.2.1.1	Introdução	28

3.3.2.1.2	Filtragem pela Transformada de Fourier . . .	29
3.3.2.2	Filtragem por convolução	29
3.3.2.2.1	Filtragem passa-baixa	31
3.3.2.2.2	Filtragem passa-alta	32
3.3.2.2.3	Mediação local ou Filtragem da Média . . .	33
3.3.2.3	Filtragem não linear	34
3.3.2.3.1	Filtragem da ordem e Filtragem da Mediana	34
3.3.2.3.2	filtragem de Nagao-Matsuyama	36
3.3.2.3.3	filtragem homomórfica	37
3.3.2.4	Realce de bordas e detalhes	37
3.3.2.5	Operador de diferenciação	38
3.3.2.5.1	Gradiente	38
3.3.2.5.2	Operador de Prewitt	40
3.3.2.5.3	operador 1 de Sobel	40
3.3.2.5.4	operador 2 de Sobel	40
3.3.2.5.5	Laplaciano	41
4	SEGMENTAÇÃO DE IMAGENS	43
4.1	INTRODUÇÃO	43
4.2	SEGMENTAÇÃO POR REGIÃO	43
4.2.1	Limiarização	44
4.2.1.1	Introdução	44
4.2.1.2	Limiarização global	46
4.2.1.3	Limiarização local adaptativa	48
4.2.1.4	A escolha de um método de limiarização	48
4.2.2	Método de Divisão e Fusão	49
4.3	SEGMENTAÇÃO POR TEXTURA	50
4.3.1	Técnica Estatística	52
4.3.2	Técnica Estrutural	53
4.3.3	Técnica Espectral	54
4.4	SEGMENTAÇÃO POR CONTOURNO	54
4.4.1	Deteção de Pontos	55
4.4.2	Deteção de Bordas	55
4.5	ENLACE DE BORDAS E DETECÇÃO DE CONTORNOS	56
4.5.1	Perseguição de Contornos	57
4.5.2	Transformada de Hough	58
4.5.2.1	Deteção de retas pela Transformada de Hough . . .	58
4.5.2.2	Deteção de círculos pela Transformada de Hough .	59
4.5.2.3	Conclusão sobre a Transformada de Hough	61
4.5.3	Técnicas do grafo	62
5	MORFOLOGIA MATEMÁTICA	65
5.1	CONSIDERAÇÕES GERAIS	65
5.2	CONSIDERAÇÕES INICIAIS	66
5.2.1	Translação e Simetria	66
5.2.2	Operações de Minkowski	67
5.3	PROPRIEDADES BÁSICAS	67
5.3.1	Propriedade de Invariância à Translação	67
5.3.2	Propriedade de Invariância à Mudança de Escala	68
5.3.3	Conhecimento Local	68
5.3.4	Continuidade	69
5.4	OPERADORES MORFOLÓGICOS BINÁRIOS BÁSICOS	69
5.4.1	Aspecto digital	69
5.4.2	Operadores elementares	70

5.4.3	Erosão binária	70
5.4.3.1	Primeira Definição	70
5.4.3.2	Segunda Definição	72
5.4.3.3	Efeitos da erosão	74
5.4.4	Dilatação binária	75
5.4.4.1	Primeira Definição	75
5.4.4.2	Segunda Definição	77
5.4.4.3	Efeitos da dilatação	78
5.4.5	Propriedades da Erosão e da Dilatação binárias	79
5.4.6	Principais classes de elementos estruturantes	83
5.4.7	Deteção de bordas	84
5.4.8	Abertura binária	87
5.4.8.1	Definição da abertura binária	87
5.4.8.2	Efeitos da abertura binária	90
5.4.9	Fechamento binário	91
5.4.9.1	Definição do fechamento binário	91
5.4.9.2	Efeitos do fechamento binário	93
5.4.10	Propriedades da abertura e do fechamento	96
5.5	CONDICIONALIDADE E RECONSTRUÇÃO	99
5.5.1	Introdução	99
5.5.2	Erosão e Dilatação condicionais binárias	99
5.5.3	Exemplos de dilatação condicional binária	99
5.5.4	Aplicações da dilatação condicional binária	100
5.5.5	Reconstrução binária: teoria	101
5.5.6	Reconstrução binária: exemplo	102
5.6	EXERCÍCIOS	107

Lista de Figuras

1.1	As etapas da visão humana	2
1.2	Estrutura funcional de um sistema de visão artificial	7
2.1	Imagem quantificada com: a) 256 níveis de cinza b) 16 níveis de cinza c) 2 níveis de cinza	11
2.2	4-vizinhança	12
2.3	8-vizinhança	13
2.4	Ângulos das diversas rotações	14
2.5	Sistema de coordenadas pela transformação perspectiva	15
3.1	Exemplo de imagem e seu histograma associado	21
3.2	Exemplo de melhoria da dinâmica	22
3.3	Exemplo de uniformização de densidade	23
3.4	Exemplo de equalização de histograma: a) imagem original b) his- tograma correspondente c) imagem euqlizada d) histograma corre- spondente	24
3.5	Princípio da modificação do contraste	26
3.6	Exemplos de espectros de Fourier de: a) um ponto b) dois pontos c) uma linha horizontal d) um losângo	28
3.7	Exemplos de filtragem pela Transformada de Fourier: a) imagem original com linhas oblíquas, b) espectro de Fourier correspondente, c) imagem filtrada por remoção dos picos de frequência.	30
3.8	Função de transferência do filtro de Butterworth	31
3.9	Filtro uniforme circular	32
3.10	Filtro gaussiano	32
3.11	Filtro exponencial circular	32
3.12	Comparação dos filtro da média e do filtro da mediana: a) imagem original, b) média com uma janela 3×3 , c) mediana com uma janela 3×3 , d) média com uma janela 5×5 , e) mediana com uma janela 5	35
3.13	Respostas do filtro linear e da mediana a alguns sinais de teste	36
3.14	Representação das 9 máscaras do filtro não linear de Nagao-Matsuyama	37
3.15	Interpretação das informações extraídas pelo gradiente	39
3.16	Exemplo numérico da aplicação do operador Gradiente G_x de Prewitt	40
3.17	Exemplo de gradiente pelo operador 1 de Sobel: a) imagem original, b) módulo do gradiente, c) gradiente G_x , c) gradiente G_y	41
3.18	Exemplo numérico de realce de bordas pelo Laplaciano num sinal unidimensional	41
3.19	Máscara do Laplaciano	42
3.20	Exemplo numérico da aplicação do operador Laplaciano	42
3.21	Exemplo de Laplaciano	42

4.1	Influência do valor do limiar sobre a qualidade da limiarização: a) imagem original, b) histograma, c) limiarização com valor 50, d) limiarização com valor 150 e) limiarização com valor 235	45
4.2	Binarização a partir de funções de densidade de probabilidade Gaussianas.	47
4.3	Segmentação por divisão	49
4.4	Processo de divisão e fusão por Quadtree	50
4.5	Exemplo de texturas	51
4.6	Mixtura de texturas de tipo suave, irregular, granulada, paralela, periódica	52
4.7	Construção de texturas: (a) Primitiva de textura, (b) Regra $S \rightarrow aS$, (c) Uso de várias regras	54
4.8	Segmentação pelo operador de Moravec: (a) imagem original, (b) imagem binarizada dos pontos de Moravec	55
4.9	Qualidade do Gradiente: (a) imagem original, (b) módulo do gradiente 3x3, (c) módulo do gradiente 7x7	56
4.10	Pesquisa do contorno a partir dos códigos de Freeman	57
4.11	Exemplo de detecção de círculos pela Transformada de Hough: (a) imagem original em níveis de cinza, (b) tabela de acumulação (c) visualização tridimensional da tabela de acumulação (d) visualização bidimensional dos centros dos discos	60
4.12	Relação entre a direção do gradiente e a direção da tangente	61
4.13	Exemplo de detecção de círculos pela Transformada de Hough numa imagem complexa.	62
4.14	Elemento de borda entre os pixels $P(x, y)$ e $Q(x, y)$	63
4.15	Exemplo de detecção de contorno por grafo: a) imagem original b) contorno correspondendo ao caminho de menor custo	63
4.16	Grafo usado para a detecção de contorno da imagem da figura 4.15	64
5.1	Exemplos de inclusão do elemento estruturante numa imagem binária	65
5.2	Exemplo de interação entre o elemento estruturante e a imagem cinza	66
5.3	Translação e homotetia	68
5.4	Translação e simetria	68
5.5	Princípio da erosão binária	71
5.6	Interpretação da erosão binária pela subtração de Minkowski	73
5.7	Exemplos de erosão binária	74
5.8	Princípio da dilatação binária	76
5.9	Exemplos de dilatação binária	78
5.10	Exemplo de composição do elemento estruturante por dilatação	82
5.11	Exemplos de elementos estruturantes	84
5.12	Exemplos de dilatação a partir do mesmo elemento estruturante com um posicionamento diferente do ponto central	84
5.13	Detecção de bordas de uma imagem binária	85
5.14	Princípio da abertura binária	87
5.15	Exemplos de abertura binária	88
5.16	Princípio do fechamento binário	91
5.17	Exemplos de fechamento binário	95
5.18	Exemplo de reconstrução binária por dilatação condicional	103
5.19	Eliminação de conjuntos por reconstrução binária	104

Capítulo 1

INTRODUÇÃO

O sentido da visão pode ser considerado o meio mais eficiente de que o ser humano dispõe para captar as informações originadas no ambiente que o cerca. Algumas aplicações de imagens e alguns tipos de imagens requerem uma interação visual bastante intensa [Shi87]. A capacidade humana para processar e interpretar imensas quantidades de dados de natureza visual motiva o desenvolvimento de técnicas e dispositivos, de modo a estender esta sua capacidade e sensibilidade ainda mais. O conhecimento de como reage o sistema visual humano e de algumas técnicas disponíveis para melhor adequar a imagem à aplicação são importantes para poder explorar mais eficientemente os recursos de sistemas de processamento de imagens [FF87]. A utilização da imagem na forma digital torna possível o seu processamento computacional, aumentando sua qualidade.

1.1 A VISÃO HUMANA

O sistema visual humano refere-se ao complexo sistema biológico que permite ao homem a interpretação da radiação eletromagnética do espectro visível. Este abrange as operações fisiológicas do olho humano, seu processo de conversão desta radiação em sinais neuronais e operações de processamento de baixo, médio e alto nível que ocorrem em diversas áreas do cérebro (figura 1.1):

- **Emissão de sinais luminosos por reflexão da luz**

Os raios luminosos, provenientes de fontes externas, são refletidos pelas superfícies dos objetos, dando assim, origem a sinais que vão de encontro ao sistema visual.

- **Recepção dos sinais luminosos pelo sistema visual**

Para que esses sinais sejam recebidos adequadamente, são necessários ajustes de foco e de luminosidade, assim como a existência de uma superfície sensível à luz. Num sistema orgânico, o cristalino, o tamanho da pupila e a retina, respectivamente, fazem o papel requerido. De modo análogo ao de uma câmara fotográfica que focaliza a luz no plano do filme, a luz oriunda de uma cena é focalizada pelo cristalino do olho, formando uma imagem em receptores discretos de luz sobre a superfície da retina, células nervosas denominadas cones e bastonetes [Dou81].

Já em um sistema artificial, temos um conjunto de lentes, um mecanismo de movimentação das mesmas, um diafragma e uma superfície sensível que depende do sistema em utilização.

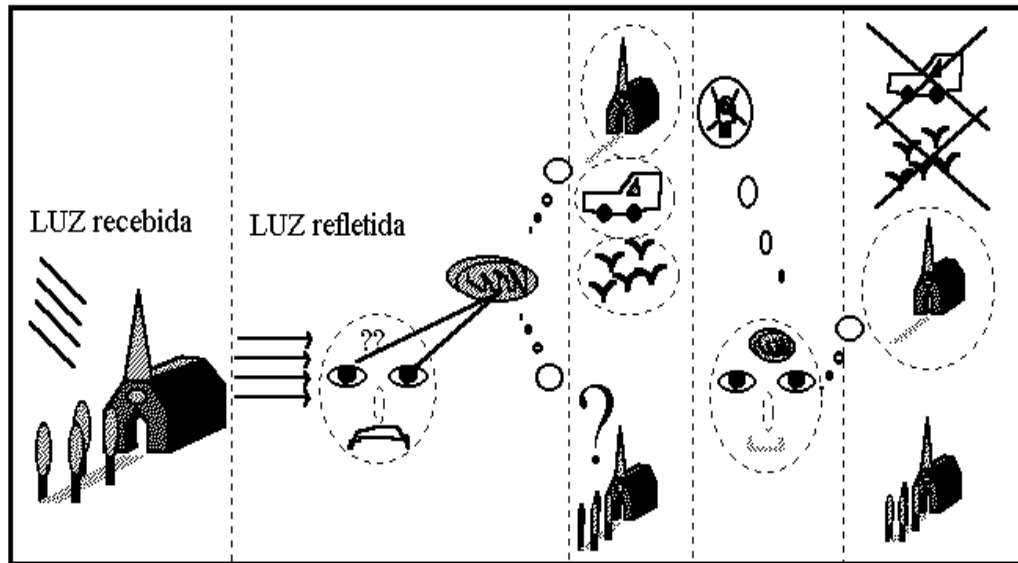


figura 1.1: As etapas da visão humana

- **Transformação dos sinais luminosos pela superfície sensível**

A retina é composta por células nervosas sensíveis à intensidade e ao comprimento de onda da luz. Essas células transformam as respectivas informações em sinais nervosos, os quais são remetidos ao cérebro pelo nervo óptico. A resolução de detalhes finos está principalmente ligada ao fato de que vários cones podem estar conectados a uma única terminação nervosa. A grande área de distribuição dos bastonetes na retina e o fato de que vários destes estão conectados a uma única terminação nervosa reduz a quantidade de detalhes discerníveis por estes receptores. Os cones são extremamente sensíveis a cores e estão concentrados em uma pequena região central, enquanto que o restante da retina consiste numa mescla de cones e bastonetes. Portanto, o olho possui amostragem espacial e características de resolução não lineares.

Para atender a exigências do mundo industrial, existe hoje uma diversidade muito grande de sensores que permitem a transformação da intensidade do sinal luminoso em um sinal elétrico. A evolução e a concepção dos mesmos acompanha de perto os circuitos eletrônicos, indo das válvulas (orticon, vidicon, plumbicon) aos circuitos de estado sólido (matriz de diodos e CCDs [DNBC87]).

- **O armazenamento da informação**

A informação é armazenada em uma memória compatível com o sistema em questão. Existe uma série de teorias sobre o funcionamento da memória animal assim como uma grande quantidade de tipos de memórias no mercado. O principal é que o sistema seja capaz de armazenar as informações por um determinado período de tempo e de ter acesso de modo eficiente a essas quando necessário [CDG92].

- **O processamento da informação**

A informação, uma vez recebida e armazenada, é processada de acordo com a configuração do sistema. A configuração está determinada pela evolução do organismo, assim como por processos de assimilação e/ou adaptação do mesmo. Existem várias teorias de compreensão do processamento da informação pelo cérebro. Uma abordagem é considerar que o cérebro leva em conta uma base

de dados para efetuar uma comparação entre a informação recebida e alguns padrões de situações semelhantes armazenados nessa base.

- **Resposta Nervosa**

De uma forma geral, o cérebro fornece para a entidade envolvida no processo (braço, perna, etc...) uma resposta dependente da semelhança encontrada ou não encontrada e prepara o comportamento adequado. Caso encontrar um padrão desconhecido ou novo, o cérebro tem a faculdade de classificá-lo na base de dados e então enriquecê-la.

1.2 CARACTERÍSTICAS DO SISTEMA VISUAL HUMANO

- **Assimilação da informação**

O número e a natureza de padrões que os sistemas visuais orgânicos são capazes de processar fazendo a devida assimilação e adaptação contínua do sistema é extremamente elevada. O atual conhecimento científico e tecnológico não permite tal desempenho e uma escolha com perda significativa de informação é em geral efetuada.

- **Eficiência do processamento**

O tempo com que os sistemas orgânicos desempenham o processamento da informação é o suficiente e o necessário para a sobrevivência dos organismos. Já a eficiência do processamento artificial impõe sérias restrições ao tempo de resposta necessário a muitas aplicações.

- **Diversidade do ambiente**

Enquanto um filme fotográfico possui um único nível de sensibilidade, a retina pode adaptar sua sensibilidade em uma vasta gama, por meio de adaptações químicas em seus receptores e pelo fato de que os bastonetes são especializados para baixas intensidades luminosas (visão scotópica) e os cones para altas intensidades (visão fotópica). A retina pode também se adaptar a variações locais que ocorrem numa imagem, mudando o limiar de sensibilidade de células receptoras individuais. Já um filme não tem essa flexibilidade.

A habilidade do olho em discriminar entre vários níveis de brilho é uma consideração importante para a apresentação dos resultados do processamento de imagens. A faixa dos níveis de intensidade luminosa ao qual o sistema visual humano pode se adaptar é da ordem de 1000. Há também considerável evidência experimental de que o brilho subjetivo, que é o brilho percebido pelo sistema visual humano, seja uma função logarítmica da intensidade da luz incidente no olho. Para o sistema visual humano poder operar sobre toda a faixa de intensidades, ocorre uma mudança da sua sensibilidade global, um fenómeno conhecido como adaptação de brilho. No caso de uma cena complexa, o sistema visual não se adapta a um nível único de intensidade, mas sim a um nível médio que depende das propriedades desta cena. À medida que o olho a percorre, o nível de adaptação instantâneo flutua em torno desta média. De modo a obter uma visualização suave para o olho, para uma grande classe de tipos de imagem, normalmente é necessária uma gama com mais de 100 níveis de intensidades.

Em função dos dados anteriores, podemos constatar que a capacidade dos organismos de se adaptarem a situações novas, e às vezes adversas, é extremamente boa quando comparada com os sistemas artificiais. Condições

deficientes de iluminação, informação imprecisa ou parcial, deformação da imagem e outros fatores exigem estruturas de tratamento da informação extremamente confiáveis para que o reconhecimento seja feito adequadamente.

1.3 CONCEPÇÃO DE UM SISTEMA DE VISÃO ARTIFICIAL

Os sistemas de visão artificial visam, com o auxílio do conhecimento de diversas áreas (biologia, medicina, comunicação visual, eletrônica, matemática, mecânica fina), obter um conjunto de técnicas e metodologias que possam dar suporte ao desenvolvimento de teorias e produtos suficientemente eficientes e confiáveis para aplicações práticas. Cita-se, como exemplo, a automatização dos processos de controle de qualidade, identificação e classificação de produtos e exploração de ambientes diversos.

O processamento e a análise de imagens é uma ciência que permite modificar, analisar e manipular imagens digitais, originalmente contínuas, a partir de um computador. Os algoritmos de processamento de imagens são procedimentos passo-a-passo para otimizar as operações de tratamento de imagens. A estação de processamento de imagens deve fornecer três facilidades: os meios de digitalizá-las, de visualizá-las e de manipulá-las [MV89].

Estas considerações supõem que o sistema seja dotado de ferramentas que permitam atingir esses objetivos. Um sistema genérico mínimo de processamento de imagens é constituído de captadores de visão, de uma parte para aquisição-digitalização de imagens, uma memória para as mesmas, um computador que pode ter acesso a essa memória e um conjunto de ferramentas que permita visualizar o conteúdo da memória. Então, o sistema adquire, processa e visualiza imagens.

Um sistema de visão é constituído de :

- **Sensores de Visão**

fornece uma projeção da cena de trabalho e realizam a aquisição de imagens. Na maioria dos casos, esses sensores são câmeras, sensores eletromecânicos, etc... No caso do uso de uma câmera, um sistema de iluminação pode permitir a obtenção de uma imagem de melhor contraste e de melhor qualidade, reduzindo portanto a quantidade de processamentos preliminares.

- **Um Hardware de Digitalização de Imagens**

cujo alvo é colocar a imagem do sensor na memória. O módulo de aquisição permite escrever de modo eficiente na memória que pode ser lida pelo computador e pelo módulo de visualização. O módulo de digitalização transforma as imagens contínuas em imagens digitais. O alvo dessa transformação é quantificar a qualidade, o contraste, as diferenças de cores da cena. O digitalizador cria uma imagem digitalizada, representada por uma matriz de números, cujos elementos são chamados pixels.

- **Um Computador**

executa os algoritmos de processamento de imagens permitindo flexibilidade e custos de processamento e de memória relativamente baixos. Os algoritmos são desenvolvidos para, em condições de iluminação a priori variáveis, possibilitar a escolha da informação adequada nas imagens, a ser interpretada e a partir da qual serão tomadas as decisões.

1.4 EXEMPLOS DE PROCESSAMENTO DE IMAGENS

A importância da visão na vida do ser humano é uma evidência para cada um de nós. A vontade e a necessidade crescente de automatização de atividades codificadas evidenciam um envolvimento crescente das ferramentas de processamento de imagens em um grande número de domínios. Os processos utilizados para tal, podem variar de acordo com o objetivo e a aplicação. Abrangem desde uma simples tarefa de aprimoramento de imagens adquiridas, até atingir a dificuldade de compreensão de um mecanismo. Aplicações de processamento e análise de imagens podem ser classificadas a partir de critérios como a possibilidade de um conhecimento inicial prévio do contexto ou não, a necessidade de processar em tempo real, etc... A seguir são apresentadas exemplos de aplicações que envolvem técnicas de processamento de imagens.

1.4.1 Tarefas Industriais

De forma geral, o desenvolvimento de um projeto industrial que exija a utilização de técnicas de processamentos de imagens supõe um importante conhecimento inicial do produto (normas de concepção, regras de produção, etc...) e exige técnicas rápidas para atender processos em tempo real.

- **Pilotagem de um robô**

Controlar um robô é uma das tarefas que requer rapidez e precisão. Um robô é usado para manipular peças que podem ser ordenadas, ou montadas com outros tipos de peças. Estas podem já estar acondicionadas em locais previamente conhecidos ou não. O alvo do robô, atingido com auxílio de visão computacional, pode envolver a escolha de uma peça entre várias outras peças, a definição da melhor pega ou a determinação de uma trajetória para evitar obstáculos.

- **Inspeção visual**

O controle de qualidade de um produto é uma tarefa essencial no domínio industrial. O processo de inspeção implica na medição de determinadas propriedades de um produto, como dimensões geométricas, superfícies, posição, orientação etc... Um dos alvos da visão computacional consiste em analisar, sem contato com o produto, os defeitos cuja detecção seria impossível ou muito difícil por outros métodos. A inspeção automatizada, mediante técnicas de processamentos de imagens possibilita a quantificação de propriedades e coleta de dados sobre o produto que um inspetor humano não é capaz de realizar, pois além de ser tedioso ou mesmo não realizável, permite uma realimentação constante e imediata do processo de manufatura.

1.4.2 Reconhecimento de padrões

A exigência crescente informação e automatização de tarefas humanas repetitivas e/ou cansativas, levou os pesquisadores a desenvolver ferramentas específicas para atender projetos cujo conhecimento inicial é pobre e/ou incompleto e onde um certo grau de inteligência é importante. O reconhecimento de um padrão inscreve-se neste tipo de tarefa. Dentro dos assuntos de pesquisa em reconhecimento de padrões, podemos citar:

- o reconhecimento de caracteres (OCR: “Optical Character Recognition”) cujos recentes progressos nas pesquisas permitem o desenvolvimento de sistemas

de leitura automática, já disponíveis no mercado. A dificuldade de uma tal pesquisa reside no grande número de tipos de letras usados e na necessidade de incluir o conhecimento sintático, semântico e pragmático do idioma em questão (árabe, japonês, latim). O que são problemas abertos não só em visão computacional, mas também em lingüística computacional.

- o reconhecimento de impressões digitais (AFIS: “Automatic Fingerprints Identification System”), cujo interesse para o bom funcionamento da sociedade é bastante útil, pode ser em parte automatizado. A dificuldade do confronto de uma impressão digital latente com um banco de dados reside na fragilidade da informação relevante contida na mesma, nas dificuldades de um levantamento cuidadoso do fragmento da impressão, e no tamanho e forma de organização do banco de dados que coloca restrições de desempenho no desenvolvimento de sistemas rápidos.
- o reconhecimento de assinaturas, cujo interesse para entidades bancárias e burocráticas faz-se necessário. Mas a facilidade que o ser humano tem de mudar a sua própria assinatura, de falsificar uma assinatura de outra pessoa, e/ou de conseguir reconhecê-la inviabilizou até agora sistemas automatizados de identificação.

1.4.3 Reconstrução tridimensional

A constante evolução de desempenho dos computadores viabiliza a percepção tridimensional do mundo a partir de imagens bidimensionais.

- vários projetos mundiais (Projeto Eurêka europeu por exemplo) tem como uma das metas a utilização de imagens bidimensionais para a melhoria da vida cotidiana de maneira a dotar o ser humano de ferramentas integradas de alto nível, como o carro de pilotagem automática, o robô limpador do metrô, etc... As dificuldades residem na complexidade das imagens do mundo real, na quantidade de imagens a serem processadas, e na necessidade de tratar os dados em tempo real.
- o aparecimento da tomografia de ressonância magnética, com imagens de alta definição e de nitidez nunca antes atingida até agora, permite esperar uma explosão de ferramentas de visualização tridimensionais de diferentes partes do corpo do ser humano e o desenvolvimento de sistemas de apoio ao diagnóstico médico.

1.5 ARQUITETURA DE UM SISTEMA DE VISÃO ARTIFICIAL

O processamento e a análise de imagens é uma ferramenta multidisciplinar e se presta à utilização para fins diversos. Os processamentos mudam muito em função da área de trabalho. Os tratamentos variam sobre as imagens segundo:

- **a natureza das imagens**
A natureza das imagens provém das áreas como medicina, do meio industrial ou de laboratório, onde as entidades encontradas e informações contidas são muito diversas.
- **a qualidade das imagens**
Os requisitos de qualidade de uma imagem variam de acordo com a aplicação. A resolução, que é o grau de detalhes discerníveis em uma imagem, é muito

dependente das dimensões da matriz da imagem, do número de níveis de cinza de cada pixel e do intervalo entre imagens (no caso de uma sequência de imagens). Características como o número e distribuição dos detalhes e a presença de áreas com variação gradual de brilho influem no grau subjetivo atribuído à qualidade da imagem. As condições de iluminação do ambiente estudado influenciam a qualidade das informações disponíveis (regiões, contornos, detalhes, contrastes etc...).

- **o conhecimento inicial do meio de estudo**

o conhecimento inicial de uma cena é elemento importante no processo de análise. No meio industrial, o conhecimento inicial quase sempre existente permite uma análise mais dirigida para uma solução rápida. Em casos onde o conhecimento inicial é fraco, como no caso de imagens adquiridas por satélites de planetas. O processo de análise deve ser então mais completo e mais rico para suprir a falta de conhecimento.

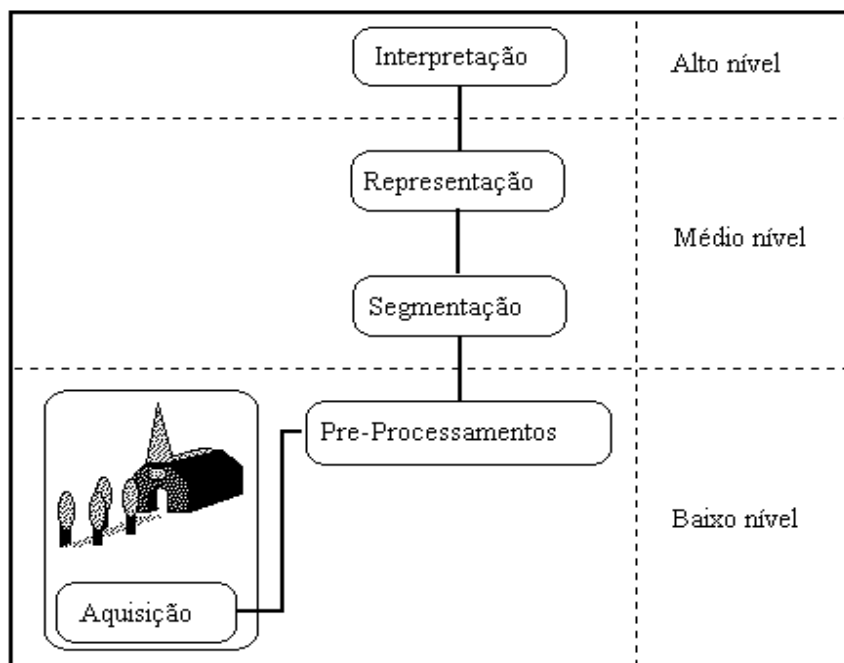


figura 1.2: Estrutura funcional de um sistema de visão artificial

O processamento digital de imagens pode ser classificado, quanto ao grau de abstração, em três níveis distintos: baixo, médio e alto [Duf86] [Mar82]. Ocorre uma redução progressiva da quantidade de informações manipuladas, à medida que se passa por níveis crescentes de abstração. No processamento de baixo nível, os dados de entrada são pixels da imagem original e os dados de saída representam propriedades da imagem, na forma de valores numéricos associados a cada pixel. No processamento de nível médio este conjunto de valores produz como resultado uma lista de características. O processamento de alto nível produz, a partir destas características, uma interpretação do conteúdo da imagem. Estas especificações supõem uma diversificação dos processamentos. Uma estrutura funcional completa de um sistema de processamento e análise de imagens pode ser descrito [Mar82] [Ros88] da seguinte forma (figura 1.2):

- **Aquisição e digitalização**

A imagem do sensor é transformada em uma imagem digital sobre a forma de

uma tabela de valores discretos inteiros chamados pixels.

- **Pré-processamento**

Essa etapa permite corrigir um certo número de defeitos e imperfeições aparecidos durante a aquisição da imagem, que podem ter como causa características físicas do sistema, as condições deficientes de iluminação, etc... O pré-processamento não é indispensável, mas, na maioria dos casos, necessário.

- **Segmentação**

A meta consiste em dividir uma imagem em partes constitutivas. Em uma imagem natural, a segmentação é efetuada pela detecção de descontinuidades (contornos) e/ou de similaridades (regiões) na imagem. A maioria dos processamentos é baseada na pesquisa dessas entidades que são armazenadas sobre uma forma adequada (segmentos ou primitivas).

- **Representação**

O alvo da representação é elaborar uma estrutura adequada, agrupando resultados das etapas precedentes e o armazenamento dos diversos padrões que contém o conhecimento a priori. A representação é efetuada após a obtenção das primitivas e, às vezes, ao mesmo tempo. Ela permite também medir as propriedades das “formas” resultantes da segmentação.

- **Interpretação**

É a parte mais “inteligente” do processo de visão por computador: ela representa o “alto nível” e permite obter a compreensão e a descrição final do fenômeno inicial. Ela faz uso do conhecimento a priori do caso estudado e o conhecimento adquirido durante as fases precedentes.

Capítulo 2

FUNDAMENTAÇÃO

2.1 INTRODUÇÃO

A imagem é a matéria prima indispensável para se resolver projetos de processamento de imagens e de visão computacional. Serão apresentados a seguir os conceitos de imagem, de pixel, de nível de cinza assim como os tipos de imagens existentes, e algumas transformações básicas de imagens.

2.2 A IMAGEM

2.2.1 Introdução

Uma imagem pode ser descrita por uma função $f(x, y)$ da intensidade luminosa, sendo seu valor, em qualquer ponto de coordenadas espaciais (x, y) , proporcional ao brilho da imagem naquele ponto. Já no caso de uma imagem que possui informações em intervalos ou bandas distintas de frequência, é necessário uma função $f(x, y)$ para cada banda. É o caso de imagens coloridas padrão *RGB*, que são formadas pela informação de cores primárias, como o vermelho "*Red*", verde "*Green*" e azul "*Blue*". Para o processamento da imagem digitalizada, é fundamental representar sua informação num formato adequado ao tratamento computacional. Uma imagem pode ser representada por uma matriz, em que os índices de linha e coluna referenciam o brilho médio amostrado no ponto correspondente da cena.

2.2.2 Imagem da Câmera

Uma imagem monocromática pode ser descrita por uma função matemática bidimensional $f(x, y)$ cujo valor indica a intensidade ao ponto (x, y) . A função $f(x, y)$ representa o produto da interação entre a iluminância $i(x, y)$ que ilumina o objeto e as propriedades de reflectância ou de transmitância próprias do objeto [GW87]. Essas propriedades podem ser representadas pela função $r(x, y)$ cujo valor exprime a fração de luz incidente que o objeto vai transmitir ou refletir ao ponto (x, y) .

$$f(x, y) = i(x, y) \times r(x, y) \quad (2.1)$$

com

$$I_1 < i(x, y) < I_2 \quad \text{unidade: candela / m}^2$$

$$0 < r(x, y) < 1$$

Exemplos :

$i(x, y)$	=	9000	tempo claro
	=	1000	tempo nublado
	=	100	iluminação média de escritório
	=	0,01	noite clara de lua cheia
$r(x, y)$	=	0,93	neve
	=	0,80	parede branco-fosca
	=	0,65	aço inoxidável
	=	0,01	veludo preto

2.2.3 Imagem Digital

De forma geral, os sensores fornecem um sinal analógico vídeo correspondente à amostragem seqüencial da imagem. Dado que a grande maioria das técnicas de processamento de imagens é realizada de forma numérica em um computador com o possível auxílio de processadores específicos, a discretização do sinal analógico faz-se necessário. A amostragem da imagem contida no sinal analógico vídeo é obtida por transformação via um conversor analógico-digital controlado por um sinal de relógio, seja ainda por um digitalizador [Fis84]. A discretização ou digitalização correta de um sinal e a minimização do erro necessitam a tomada de precauções ditadas pelo teorema de Shannon [Cia87]. A amostragem regular com um passo T de um sinal contínuo unidimensional $f(x)$ consiste em considerar os valores de $f(x)$ só nos pontos cuja abscissa é um múltiplo de T . Matematicamente, esta operação é efetuada por multiplicação do sinal $f(x)$ por uma função composta de Dirac. A condição de Shannon de boa amostragem expressa-se da seguinte forma:

para amostrar sem perda de informação um sinal cujo espectro é limitado, é preciso amostrar este sinal em uma frequência pelos menos igual ao dobro da maior frequência que o sinal contém.

Na prática, a condição de Shannon não é satisfeita para um sinal real por causa do ruído introduzido pelos diferentes elementos da cadeia de aquisição (cabos, conversores, etc...). Uma filtragem passa-baixa faz-se necessária para amenizar os efeitos das altas frequências do ruído.

A imagem digital é portanto o resultado da transformação da função matemática contínua $f(x, y)$ em uma outra função matemática $F(x, y)$ cujos valores são discretos.

A função $F(x, y)$ é definida como:

$$\begin{aligned}
 D_1 \times D_2 &\implies I = \{0, 1 \dots M\} \\
 x \in D_1 &= \{0, 1 \dots N_1\} \\
 y \in D_2 &= \{0, 1 \dots N_2\} \\
 (x, y) &\implies F(x, y) \\
 \text{se } d_i < f(x, y) < d_{i+1} &\implies F(x, y) = r_i
 \end{aligned}$$

$F(x, y)$ constitui o **tom de cinza** ou **nível de cinza**.

Um problema importante na concepção de um digitalizador é a escolha de uma boa quantificação.

A imagem pode ser representada por uma tabela $N_1 \times N_2$, onde cada elemento dessa tabela chama-se pixel (abreviação de “picture element”) e é codificado com R níveis de cinza.

Muitas vezes, N_1 , N_2 e R são potências de 2. Se $N_1 = 2^p$, $N_2 = 2^q$ e $R = 2^r$ então a imagem digital é uma matriz de $2^p \times 2^q$ valores inteiros.

O número de bits necessários para armazenar essa imagem é :

$$b = N_1 \times N_2 \times r \quad (2.2)$$

Os números N_1 e N_2 determinam a qualidade da imagem. A escolha de N_1 , N_2 e r depende da imagem. Para achar uma qualidade semelhante a de uma imagem de televisão, é preciso amostrar com 512×512 e codificar com 128 níveis de cinza. Em geral, 64 níveis de cinza são considerados como suficientes para o olho. A maioria dos sistemas de visão usa 256 níveis de cinza. A figura 2.1 ilustra bem a noção de qualidade ligada ao conceito de digitalização [Mor87a].

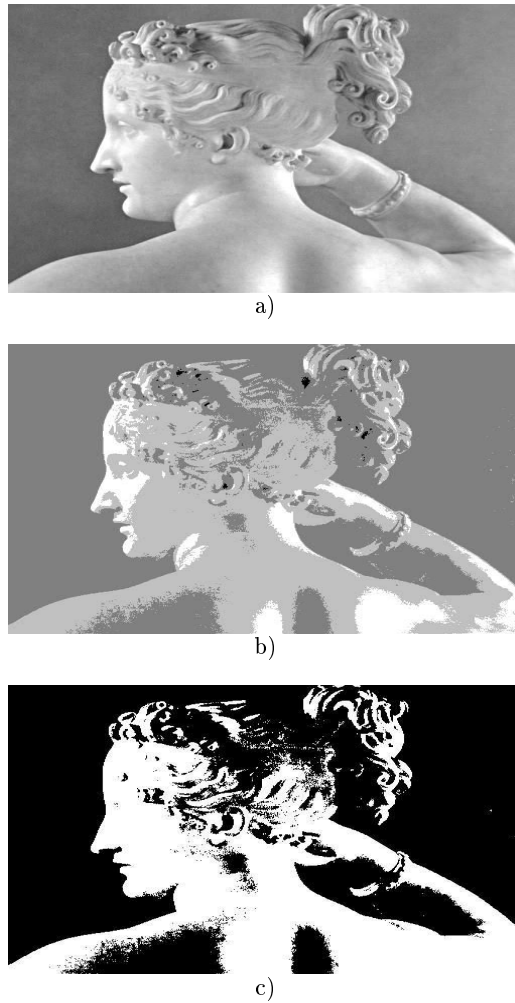


figura 2.1: Imagem quantificada com: a) 256 níveis de cinza b) 16 níveis de cinza c) 2 níveis de cinza

Não existe nenhum resultado especificando uma relação qualquer entre N_1 , N_2 e r . Quantificar a noção de qualidade de uma imagem é difícil.

Para se ter uma idéia geral dos fatores temporais da aquisição de uma imagem, veja o exemplo da digitalização de uma imagem fornecida por uma câmera de TV ordinária do tipo RS 170 de CCTV (Closed Circuit Television).

A imagem é fornecida linha por linha. Cada linha é constituída por um sinal de vídeo analógico entre 0 Volt (preto) e 1 Volt (branco). Cada linha dura $64 \mu s$ e a câmera fornece 50 meias imagens entrelaçadas de 262,5 linhas por segundo. Se cada sincronização de linha coincide com um relógio cujo período é $64/256 = 250$ ns, e se, para cada borda de subida, o sinal analógico de vídeo é convertido com 6 bits, guardando só 256 linhas sobre as 312,5 de cada meia imagem, uma imagem de

256×256 com 6 bits é obtida em 20 ms, ou seja: um débito de 19,66 megabauds, correspondendo aos conversores analógicos digitais de 4Mhz.

2.3 PROPRIEDADES DE UM PIXEL

2.3.1 Vizinhaça de um Pixel

2.3.1.1 4-Vizinhaça

A 4-vizinhaça $V_4(P)$ ou vizinhaça simples de um pixel P é definida como o conjunto dos pixels que ficam ao redor do ponto P e não leva em conta os pontos localizados nas diagonais passando pelo pixel P (figura 2.2). A 4-vizinhaça $V_4(P)$ é baseada na noção de uma distância D_4 dos quatro vizinhos mais próximos:

$$V_4(P) = \{Q | D_4(P, Q) \leq 1\} \quad (2.3)$$

onde

$$D_4[P(x_p, y_p), Q(x_q, y_q)] = \|x_p - x_q\| + \|y_p - y_q\| \quad (2.4)$$

e

$\|a\|$ é o módulo de a .

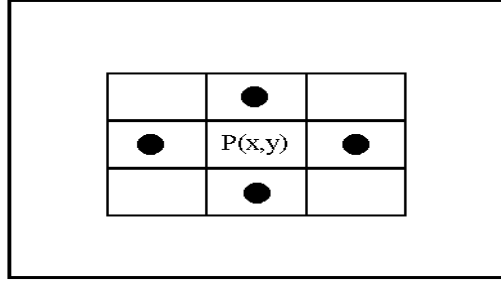


figura 2.2: 4-vizinhaça

2.3.1.2 8-Vizinhaça

A 8-vizinhaça $V_8(p)$ de um pixel P é definida como o conjunto de todos os pixels que ficam ao redor de P (figura 2.3). A 8-vizinhaça $V_8(P)$ é baseada na noção de uma distância D_8 dos oito vizinhos mais próximos:

$$V_8(P) = \{Q | D_8(P, Q) < 1\} \quad (2.5)$$

onde

$$D_8[P(x_p, y_p), Q(x_q, y_q)] = \max(\|x_p - x_q\|, \|y_p - y_q\|)$$

e

$\|a\|$ é o módulo de a .

2.3.2 Conectividade

Uma primeira definição de um conjunto conexo é que dois pontos quaisquer desse conjunto podem ser ligados por uma curva completamente contida nesse conjunto.

Uma outra definição mais completa é que dois pixels são ditos conexos se existe no mínimo uma sequência de pixels que os liga, tal que dois pixels consecutivos dessa sequência satisfaçam a condição de *conectividade imediata*.

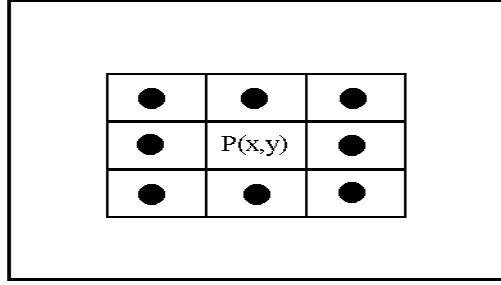


figura 2.3: 8-vizinhança

A conectividade imediata entre dois pixels traduz o fato de que os pixels têm características comuns e que eles são vizinhos. Essa definição é portanto condicionada pela noção de vizinhança (4- ou 8-vizinhança).

Os pixels conexos de uma região devem satisfazer o predicado associado que traduz o critério segundo o qual a região é construída. Esse predicado pode ser definido como uma função $F_c(\bullet)$ cujo domínio de definição pode variar muito e onde os valores dependem:

- do pixel $P(xp, yp)$ onde a função é calculada
- de uma vizinhança desse pixel P , $V(P)$
- do valor ao pixel P da função de conectividade definindo a região em formação, $G(P)$.

Essa função $G(\bullet)$ é o centro da função de conectividade $F_c(\bullet)$, ela exprime uma certa propriedade visual observada no pixel P . No caso de imagens binárias, o valor da função $G(\bullet)$ do pixel P é simplesmente o valor binário 0 ou 1 desse pixel.

Uma outra definição de um conjunto conexo pode ser enunciado como: se dois pontos quaisquer de um conjunto podem ser ligados por uma curva completamente contida nesse conjunto, então este é conexo.

2.4 TRANSFORMAÇÕES GEOMÉTRICAS

Serão detalhadas as transformações comuns usadas em processamento de imagens tais como: translação, transformação de escala, rotação e a transformação perspectiva [BB82] [GW87]. Um ponto será descrito no espaço pelas coordenadas X, Y, Z , e sua projeção sobre o plano imagem pelas coordenadas x, y .

2.4.1 Translação

Seja o vetor translação $T(X_O, Y_O, Z_O)$. O ponto $B(X', Y', Z')$, transladado do ponto $A(X, Y, Z)$, pode ser escrito como:

$$\begin{cases} X' = X + X_O \\ Y' = Y + Y_O \\ Z' = Z + Z_O \end{cases} \quad (2.6)$$

ou na forma matricial:

$$\begin{vmatrix} X' \\ Y' \\ Z' \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} X \\ Y \\ Z \\ 1 \end{vmatrix} \quad (2.7)$$

Para normalizar a escrita e poder concatenar as diferentes transformações possíveis, introduzir-se-á uma quarta coordenada e trabalhar-se-á em coordenadas homogêneas. As coordenadas homogêneas são redundantes porque um ponto num espaço $\mathbb{R}^d$ em coordenadas homogêneas é representado por uma linha em $\mathbb{R}^{d+1}$.

$$\begin{vmatrix} X^* \\ Y^* \\ Z^* \\ 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & X_0 \\ 0 & 1 & 0 & Y_0 \\ 0 & 0 & 1 & Z_0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} X \\ Y \\ Z \\ 1 \end{vmatrix} \quad (2.8)$$

Nesse sistema de coordenadas, a translação é definida pela matriz de translação T :

$$T = \begin{vmatrix} 1 & 0 & 0 & X_0 \\ 0 & 1 & 0 & Y_0 \\ 0 & 0 & 1 & Z_0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \quad (2.9)$$

2.4.2 Transformação de escala

A matriz de transformação de escala S é definida por:

$$S = \begin{vmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \quad (2.10)$$

2.4.3 Rotação

De modo geral, é necessário convencionar o sistema de eixos (fig 2.4). As diversas rotações são efetuadas no sentido horário quando a origem é vista pela parte positiva do eixo. A matriz de rotação de um ângulo γ segundo o eixo Z é :

$$R_z = \begin{vmatrix} \cos \gamma & \sin \gamma & 0 & 0 \\ -\sin \gamma & \cos \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \quad (2.11)$$

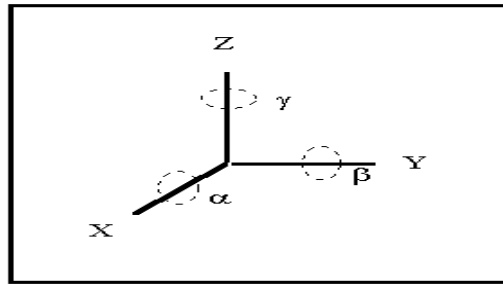


figura 2.4: Ângulos das diversas rotações

Da mesma forma, as matrizes respectivas de rotação de ângulos α e β segundo os eixos X e Y são :

$$R_x = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha & 0 \\ 0 & -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \quad (2.12)$$

e

$$R_y = \begin{vmatrix} \cos \beta & 0 & -\sin \beta & 0 \\ 0 & 1 & 0 & 0 \\ \sin \beta & 0 & \cos \beta & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \quad (2.13)$$

Uma transformação complexa pode ser descrita pelo produto das matrizes das transformações elementares correspondentes. É importante lembrar que esses produtos não são comutativos.

2.4.4 Transformação perspectiva

A imagem é o resultado da projeção pontual da cena sobre um plano: a transformação perspectiva resultante é uma transformação não linear, portanto diferente das transformações precedentes (figura 2.5).

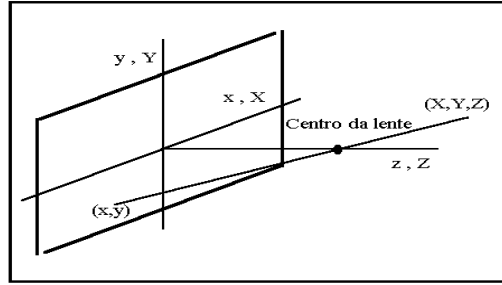


figura 2.5: Sistema de coordenadas pela transformação perspectiva

O sistema de coordenadas da câmera (x,y,z) é o seguinte: o plano imagem coincide com o plano xy e o eixo óptico passando pelo centro da lente com o eixo z . O centro do plano imagem é a origem e o centro da lente é o ponto $(0,0,f)$, onde f é a distância focal da lente. Nessas condições, o sistema de coordenadas da câmera é idêntico ao sistema de coordenadas do espaço (X,Y,Z) .

Seja (X,Y,Z) as coordenadas de um ponto do espaço com $Z > f$ e (x,y) sua projeção sobre o plano imagem.

Geometricamente, é fácil verificar que:

$$\begin{aligned} \frac{x}{f} &= -\frac{X}{Z-f} \\ &= \frac{X}{f-Z} \end{aligned} \quad (2.14)$$

e

$$\frac{y}{f} = \frac{Y}{f-Z} \quad (2.15)$$

Portanto, as coordenadas no plano imagem são dados por:

$$x = f \times \frac{X}{f-Z} \quad (2.16)$$

$$y = f \times \frac{Y}{f-Z} \quad (2.17)$$

Esta transformação não linear pode ser expressa sob a forma matricial utilizando as coordenadas homogêneas. As coordenadas de um ponto (X,Y,Z) são definidas por $(k \times X, k \times Y, k \times Z, k)$, onde k é uma constante diferente de zero.

A matriz de transformação perspectiva P é definida:

$$P = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{f} & 1 \end{vmatrix} \quad (2.18)$$

Seja um ponto W com coordenadas clássicas e W_h com coordenadas homogêneas.

$$W = \begin{vmatrix} X \\ Y \\ Z \end{vmatrix} \quad (2.19)$$

e

$$W_h = \begin{vmatrix} k \times X \\ k \times Y \\ k \times Z \\ k \end{vmatrix} \quad (2.20)$$

Tem-se:

$$C_h = P \times W_h \Rightarrow C_h = \begin{vmatrix} k \times X \\ k \times Y \\ k \times Z \\ -k \times \frac{Z}{f} + k \end{vmatrix} \quad (2.21)$$

As componentes de C_h são das coordenadas homogêneas. As coordenadas clássicas são obtidas por dividir as três primeiras componentes pela quarta.

$$C = \begin{vmatrix} f \times \frac{X}{f-Z} \\ f \times \frac{Y}{f-Z} \\ f \times \frac{Z}{f-Z} \\ f \times \frac{Z}{f-Z} \end{vmatrix} \quad (2.22)$$

As duas primeiras componentes são as coordenadas no plano imagem. A transformação perspectiva inversa é definida pela matriz P^{-1}

$$W_h = P^{-1} \times C_h \quad (2.23)$$

$$P^{-1} = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{f} & 1 \end{vmatrix} \quad (2.24)$$

Seja um ponto C_h no plano imagem :

$$W = \begin{vmatrix} k \times X \\ k \times Y \\ 0 \\ k \end{vmatrix} \quad (2.25)$$

O ponto W_h do espaço é deduzido em coordenadas homogêneas, e o ponto W :

$$W = \begin{vmatrix} X \\ Y \\ Z \end{vmatrix} = \begin{vmatrix} x \\ y \\ 0 \end{vmatrix} \quad (2.26)$$

e

$$W_h = \begin{vmatrix} k \times x \\ k \times y \\ 0 \\ k \end{vmatrix} \quad (2.27)$$

Evidentemente, o ponto não fica em $Z = 0$. É importante salientar o fato que é impossível reconstruir a partir de uma projeção. O ponto (x, y) corresponde ao conjunto dos pontos de $\mathbb{R}^3$ da reta passando para $(x, y, 0)$ e $(0, 0, f)$. As equações dessa reta são:

$$X = x \times \frac{f - Z}{f} \quad (2.28)$$

e

$$Y = y \times \frac{f - Z}{f} \quad (2.29)$$

Reconstruir o ponto exige portanto o conhecimento de Z .

A partir dessa observação, pode-se formular a perspectiva inversa de um modo diferente, utilizando a componente z como componente indeterminada :

$$C_h = \begin{vmatrix} k \times x \\ k \times y \\ k \times z \\ k \end{vmatrix} \quad (2.30)$$

e

$$W = \begin{vmatrix} k \times X \\ k \times Y \\ k \times z \\ k \times \frac{z}{f+k} \end{vmatrix} \quad (2.31)$$

Ou em coordenadas cartesianas:

$$W = \begin{vmatrix} X \\ Y \\ Z \end{vmatrix} = \begin{vmatrix} f \times \frac{x}{f+z} \\ f \times \frac{y}{f+z} \\ f \times \frac{y}{f+z} \end{vmatrix} \quad (2.32)$$

Expressar z em função de Z permite escrever :

$$X = x \times \frac{f - Z}{f} \quad (2.33)$$

e

$$Y = y \times \frac{f - Z}{f} \quad (2.34)$$

De novo chega-se na conclusão que construir o ponto exige então o conhecimento de Z .

æ

Capítulo 3

REALCE DE IMAGENS

3.1 INTRODUÇÃO

As imagens monocromáticas são, de forma geral, representadas matematicamente por uma função f , onde o valor $f(x, y)$ é uma medida do nível de cinza ou intensidade no ponto $P(x, y)$. A imagem digital é uma imagem que sofreu um processo de digitalização tanto das coordenadas espaciais quanto da intensidade, onde a digitalização das coordenadas espaciais refere-se à amostragem da imagem e a digitalização da intensidade refere-se à quantificação dos níveis de cinza. Uma imagem qualquer sofre sempre degradações diversas. Para a obtenção de uma imagem de boa qualidade pode-se considerar alguns problemas ligados à:

- correção de defeitos criados pela aquisição
- avaliação do sinal desejado a partir das observações disponíveis
- construção de uma imagem bidimensional a partir de dados unidimensionais
- compressão dos dados

Pode-se considerar vários critérios, conforme:

- as degradações sejam fracas ou fortes
- os padrões de degradação sejam conhecidos ou desconhecidos

As técnicas de melhoria da qualidade de imagens podem ser divididas em duas famílias: as de realce e as de restauração de imagens. Quando a melhoria é usada para combater um processo de degradação conhecido ou avaliado por métodos da teoria da filtragem, a palavra usada é restauração de imagens. A restauração difere de realce pelo fato de que a primeira procura obter a imagem “real” tendo, se possível, um conhecimento a priori da degradação. Sua ênfase está na modelagem de degradações e na recuperação de imagem por inversão do processo de degradação. Sua meta é a obtenção de uma imagem com entidades restauradas que diferem o menos possível daqueles da imagem “real”.

Os métodos de realce diferem dos de restauração pelo fato de que os primeiros são essencialmente projetados para manipular a imagem a partir das características psicofísicas do sistema de referência que é a visão humana.

As degradações a combater podem ser classificadas como espaciais, pontuais ou uma combinação de ambas. Degradações espaciais agem na imagem de maneira a afetar sua resolução, deixando as transições entre regiões da imagem difusas. Exemplos disto são: a ausência de foco, aberrações do sistema óptico, o movimento

de entidades da imagem, etc.. Por outro lado, as degradações pontuais, como por exemplo, ruído aditivo ou multiplicativo, distorcem os níveis de cinza mas não nublam a imagem. A luz dispersa, os reflexos de luz, o sombreamento das lentes são exemplos que podem causar degradações pontuais.

Neste capítulo apresentar-se-á os métodos de realce de imagens, permitindo enfatizar alguma característica de interesse da imagem para transformá-la em uma imagem mais “inteligível”.

As técnicas de realce de imagens são relacionadas com expansão de contraste, realce de bordas e suavização. Alguns dos objetivos principais do realce de uma imagem são a obtenção de:

- imagem com um bom contraste
- contornos nítidos
- detalhes importantes
- pouco ruído

Os filtros tradicionais no domínio espacial concentram-se, na maioria das vezes, em procedimentos em frequência (Fourier ou Walsh por exemplo) que são conhecidos como sendo filtros de grande consumo em termos de tempo computacional, apesar das técnicas de transformadas rápidas que foram desenvolvidas. Existem técnicas estatísticas que fracassam no objetivo de determinar um modelo estatístico pelo simples fato de que é quase impossível descrever adequadamente uma imagem por um processo aleatório simples.

Existem outros tipos de correção no domínio espacial como por exemplo da suavização que é a aplicação de um operador local através de uma máscara. Esse tipo de método requer algum critério de processamento obtido, ou por comparação local da vizinhança, ou por um conhecimento a priori na escolha do tamanho da máscara, dos pesos, etc...

A maioria das técnicas de realce de imagens são heurísticas e orientadas para aplicações específicas. Um critério de avaliação universal ainda não existe. O sucesso da aplicação depende do julgamento subjetivo do analista.

Técnicas digitais para realce de imagens podem ser divididas em duas categorias principais que são:

- os métodos “ponto por ponto”
 - modificação de histograma
 - modificação de contraste
- os métodos de vizinhança ou locais:
 - suavização do ruído
 - realce de contornos e de detalhes

3.2 MÉTODOS “PONTO POR PONTO”

Esta classe de métodos introduz as técnicas que modificam a distribuição dos níveis de cinza de uma imagem preservando os contornos. São operações pontuais sobre cada pixel que fazem referência à filtragem espacial.

3.2.1 NOÇÃO DE HISTOGRAMA

Seja uma imagem ideal contínua onde a intensidade x pode tomar todos os valores de um intervalo $[a, b]$. Em termos de probabilidade, pode-se definir:

FUNÇÃO DE REPARTIÇÃO $F(x)$: a função de repartição de x , $F(x)$ é a probabilidade P que x tenha um valor inferior ou igual a α , seja

$$P[x < \alpha] = F(\alpha)$$

PROBABILIDADE DO INTERVALO $[\alpha, \beta]$ é a diferença dos valores tomados pela função de repartição em α e β , seja:

$$P(\alpha, \beta) = P[x \in (\alpha, \beta)] = F(\beta) - F(\alpha)$$

No intervalo $\Delta x = (\alpha, \beta)$, a probabilidade $P(\Delta x)$ escreve-se:

$$P(\Delta x) = F(x + \Delta x) - F(x)$$

Usando a noção de limite, pode-se dizer que:

$$\lim_{\Delta x \rightarrow 0} \frac{(F(x + \Delta x) - F(x))}{\Delta x} = F'(x)$$

$F'(x)$ representa a função de densidade de probabilidade de x e denota-se $p(x)$.

O histograma na primeira ordem de uma imagem contínua é a função de densidade de probabilidade $F'(x)$ e o histograma acumulado da mesma imagem é a integral:

$$\int_0^{x_m} F'(x) dx$$

Seja o caso de uma imagem digitalizada com 2^m níveis de cinza que contém n pixels. Cada um deles toma um valor aleatório de nível de cinza r_k dentro dos k valores existentes. A função de densidade de probabilidade chama-se nesse caso p_k e expressa-se com base na função de repartição $F(r_k)$ como:

$$p_k(r_k) = \frac{(F(r_k) - F(r_k - 1))}{n} = \frac{n_k}{n}$$

onde n_k é o número de pixels para o nível de cinza r_k , e $p_k(r_k)$ representa o histograma discreto.

Pode-se notar que o histograma é uma tabela que dá, para cada nível de cinza, o número de pixels correspondentes na imagem (fig. 3.1). Na medida que o histograma fornece, para um nível de cinza dado, somente o número de pixels e não a localização desses, ele permite dar uma descrição global da imagem.

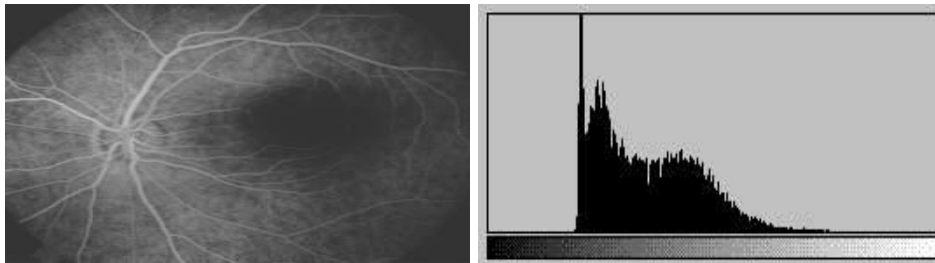


figura 3.1: Exemplo de imagem e seu histograma associado

O histograma de uma imagem contém uma informação global sobre os “objetos” nela contidos. Se todos os pontos da imagem são do mesmo objeto, o histograma dá

uma idéia da probabilidade condicional $p(r/objeto)$ de um ponto possuir um dado nível de cinza r . No caso de um único objeto, o histograma tem uma distribuição “gaussiana”. No caso de imagens constituídas de dois objetos (por exemplo o objeto e o fundo), o histograma vai ter, com freqüência, uma distribuição bimodal. De forma geral, a cena contém vários objetos distintos, o histograma fornece nesse caso uma soma ponderada das probabilidades condicionais.

O histograma é usado para melhorar a dinâmica limitada de certos captores, uma iluminação mal controlada, contrastes ruins (figura 3.2).

3.2.1.1 Modificação de histograma

De forma geral, as imagens digitais têm um espectro de luminância estreito e próximo da faixa dos pretos; os pixels têm na maioria uma luminância mais fraca que a média natural. É necessário muitas vezes aprimorar o contraste e a dinâmica da imagem. Nesse caso, por operar pontualmente sobre pixels a partir do histograma, são possíveis correções artificiais eficientes e, às vezes, espetaculares.

Vão ser abordadas ferramentas que auxiliam na manipulação e no arranjo da distribuição dos níveis de cinza.

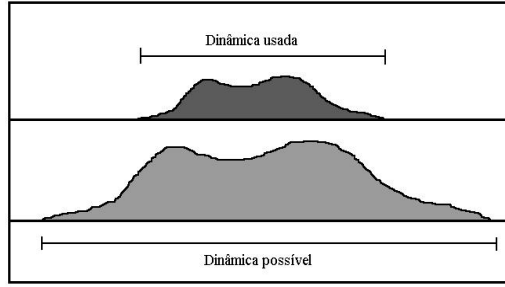


figura 3.2: Exemplo de melhoria da dinâmica

3.2.1.2 Equalização de histograma

Seja r , a variável representando o nível de cinza suposto normalizado entre 0 e 1 (0=preto e 1=branco), e uma transformação $T(\bullet)$:

$$s = T(r)$$

Supõe-se que a transformação $T(\bullet)$ satisfaça as condições:

- $T(r)$ é uma função monótona estritamente crescente no intervalo $[0,1]$
- $0 < T(r) < 1$ para $r \in [0,1]$

A transformação inversa $T^{-1}(\bullet)$ supostamente verifica as mesmas condições.

Suponha-se r e s duas variáveis aleatórias contínuas caracterizadas pelas densidades de probabilidades respectivas $p(r)$ e $p(s)$. A transformação $T(\bullet)$ define uma função de variáveis $r \rightarrow s$. Portanto a função de probabilidade $p(s)$ é definida como:

$$p(s) = [p(r) \frac{dr}{ds}]_{r=T^{-1}(s)} \quad (3.1)$$

Admita-se que se escolha como função $T(r)$ a função de distribuição da variável aleatória r , isto é:

$$s = T(r) = \int_0^r p_r(w) dw \quad (3.2)$$

A função de repartição da variável é:

$$\frac{ds}{dr} = p(r)$$

Aplicando-se a lei de s , pode-se obter:

$$\begin{aligned} p(s) &= \left(\frac{ds}{dr} \times \frac{dr}{ds} \right)_{r=T^{-1}(s)} \\ &= 1 \end{aligned} \quad (3.3)$$

A variável aleatória s é portanto uma variável aleatória uniformemente repartida sobre $[0, 1]$. Essa transformação chama-se “uniformização de densidade” (figura 3.3).

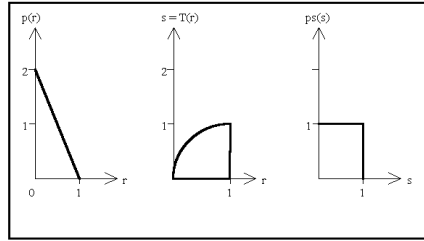


figura 3.3: Exemplo de uniformização de densidade

No caso de uma imagem original, as frequências são dadas por:

$$p_r(r_i) = \frac{n_i}{n} \quad (3.4)$$

onde i representa o nível de cinza e $i \in [0, L - 1]$ e n_i corresponde ao número de pixels do nível de cinza i e n corresponde ao número total dos pixels.

A transformação associada à uniformização é:

$$s_k = T(r_k) = \sum_{j=0}^k \frac{n_j}{n} = \sum_{j=0}^k p_k(r_k) \quad (3.5)$$

A transformação inversa é:

$$r_k = T^{-1}(s_k) \quad (3.6)$$

Ao processo geral denomina-se de equalização do histograma. Por causa dos problemas de aproximação de densidade, o histograma obtido não corresponde exatamente ao histograma de uma lei uniforme (fig 3.4).

Haralick propõe um outro algoritmo de equalização do histograma que tem a vantagem de fixar o número K de níveis de cinza. Fixando K menor que o valor de origem permite aproximar-se mais da lei uniforme. A filosofia do algoritmo é a seguinte:

Sejam $F(\bullet)$ a função de repartição de r e $Q(\bullet)$ uma função de quantificação equiprovável de K níveis de cinza.

$$Q(r) = K \Leftrightarrow q_{k-1} < r < q_k$$

O problema é obter $q_0, q_1, \dots, q_k$. O algoritmo define estes valores de uma forma iterativa. Parte-se de um valor inicial q arbitrário, por exemplo zero.

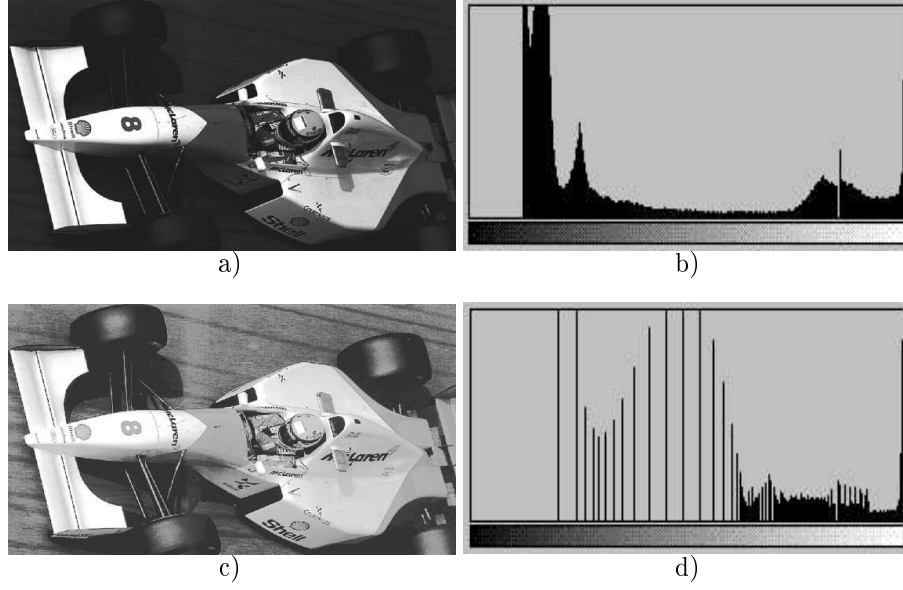


figura 3.4: Exemplo de equalização de histograma: a) imagem original b) histograma correspondente c) imagem euqlizada d) histograma correspondente

Supomos conhecidos $q_0, q_1, \dots$ até q_k . Na k -ésima iteração, $F(q_{k-1})$ é distribuída com $(k-1)$ níveis de cinza. Falta então distribuir $[1 - F(q_{k-1})]$. Esse resto deve ser repartido de uma forma uniforme sobre os níveis de cinza restantes.

Portanto, cada nível de cinza obterá:

$$\frac{[1 - F(q_{k-1})]}{[K - (k - 1)]}$$

Então:

$$F(q_k) = F(q_{k-1}) + \frac{[1 - F(q_{k-1})]}{[K - k + 1]}$$

$F(\bullet)$ é uma função em escada, não tem-se a certeza de encontrar um q_k que satisfaça a igualdade. Portanto o q mais próximo de:

$$F(q_{k-1}) + \frac{[1 - F(q_{k-1})]}{[K - k + 1]}$$

é procurado.

3.2.1.3 Especificação de histograma

Pode ser útil, para certas aplicações, escolher o tipo de densidade que devem seguir os pixels. No caso de um sinal contínuo, por exemplo, considera-se $p_z(z)$ a densidade desejada. Uma equalização é feita:

$$s = T(r) = \int_0^r p_r(w) dw$$

A densidade desejada pode ser uniformizada pela função $G(\bullet)$:

$$\nu = G(z) = \int_0^z p_z(w) dw \quad (3.7)$$

A transformação inversa $z = G^{-1}(\nu)$ dá a saída desejada. As densidades $p_s(s)$ e $p_\nu(\nu)$ são iguais. Portanto, ao invés de usar os valores de ν no processo inverso, podem-se usar os valores de s obtidos a partir da imagem original. A imagem desejada pode ser obtida por um mapeamento composto dado por:

$$z = G^{-1}[T(r)] \quad (3.8)$$

De onde o princípio:

- determinar a função, que equaliza o histograma da imagem original, a partir da equação (3.2.1.3);
- uma vez especificada a função densidade de probabilidade desejada, obter a função $G(\bullet)$ através da equação (3.7);
- aplicar a equação (3.8) aos níveis da primeira etapa.

No caso discretizado onde o número de níveis de cinza é pequeno, $G^{-1}(\bullet)$ é construído como uma tabela.

A formulação discreta pode ser especificada por:

$$s_k = T(r) = \sum_{j=0}^k p_r(r_k) \quad (3.9)$$

$$G(z) = \sum_{j=0}^k p_z(z_j) \quad (3.10)$$

$$z_k = G^{-1}(s_k) \quad (3.11)$$

onde $p_r(r_k)$ é computado pela imagem original e $p_z(z_j)$ pela imagem desejada.

3.2.2 Aprimoramento Local

Os métodos precedentes são globais porque a intensidade de cada “pixel” é modificada em função da distribuição dos níveis de cinza da imagem inteira. Para conservar certos detalhes, recorre-se às transformações ditas locais que agem sobre as vizinhanças dos “pixels”.

As modificações do histograma podem ser usadas localmente em janelas que se deslizam sobre a imagem.

Um outro exemplo de modificação local é a transformação seguinte, que não é baseada sobre a noção de histograma.

$$f(x, y) \implies g(x, y) \quad (3.12)$$

com

$$g(x, y) = L(x, y)[f(x, y) - m(x, y)] + m(x, y) \quad (3.13)$$

onde

$$L(x, y) = k \left[\frac{M}{\sigma(x, y)} \right] \text{ para } 0 < k < 1 \quad (3.14)$$

onde $m(x, y)$ e $\sigma(x, y)$ são, respectivamente, a média e a variância das intensidades em uma vizinhança de (x, y) , M é a média geral e k uma constante. O termo $L(x, y)$ amplia as variações locais, $m(x, y)$ é somado na expressão da nova intensidade para restituir a média local.

3.2.3 Modificação de contraste

Para compensar os efeitos do sensor, é necessário aumentar a dinâmica. Uma modificação do contraste é uma aplicação $t(\bullet)$ tal como:

$$z' = t(z) \text{ onde } z_{min} < z, z' < z_{max} \quad (3.15)$$

Para aumentar o contraste, com $z \in [a, b] \subset [z_{min}, z_{max}]$

$$z' = \left[\frac{(z_{max} - z_{min})}{(b - a)} \right] (z - a) + z_{min} \quad (3.16)$$

Existe também uma outra possibilidade quando a maioria dos níveis de cinza estão entre a e b :

$$z' = \begin{cases} z_{min} & \text{se } z < a \\ \left[\frac{(z_{max} - z_{min})}{(b - a)} \right] (z - a) + z_{min} & \text{se } a < z < b \\ z_{max} & \text{se } b < z \end{cases}$$

Essa transformação quase linear comprime os intervalos $[z_{min}, a]$ e $[b, z_{max}]$. Isto é aceitável quando tem-se poucos níveis de cinza nesses intervalos (figura 3.5).

Geralmente, transformações não lineares podem ser usadas, a única restrição é que o resultado deve ficar no intervalo $[z_{min}, z_{max}]$. Uma transformação $t'(\bullet)$ desse tipo pode ser definida:

$$t' = \left[\frac{(z_{max} - z_{min})}{(t_2 - t_1)} \right] (t(z) - t_1) + z_{min} \quad (3.17)$$

onde:

$$t_1 = \min t(z)$$

$$t_2 = \max t(z)$$

com $z_{min} < z < z_{max}$

$t'(\bullet)$ satisfaz a condição:

$$z_{min} < t'(z) < z_{max}$$

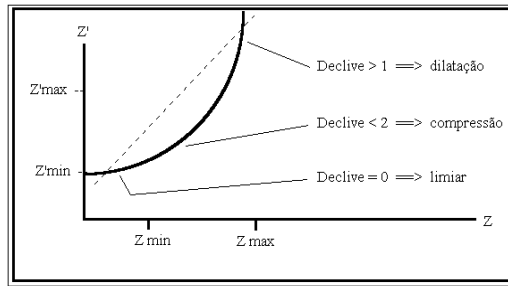


figura 3.5: Princípio da modificação do contraste

3.3 MÉTODOS DE VIZINHANÇA

Pode-se encontrar vários tipos de ruído em uma imagem. Apresentar-se-á alguns métodos simples para suavizar uma imagem. Suponha-se inicialmente que o ruído é aditivo.

3.3.1 Redução de ruído pela média de múltiplas imagens

Uma imagem adquirida por um sensor é freqüentemente ruidosa. Esse ruído pode ser reduzido pela média de sucessivas tomadas da mesma cena.

Suponha o modelo:

$$g(x, y) = f(x, y) + n(x, y) \quad (3.18)$$

onde $n(x, y)$ representa um ruído de média nula e de variância σ^2 .

Tomam-se N sucessivas imagens da mesma cena, isto é:

$$g_i(x, y) = f_i(x, y) + n_i(x, y); i \in [0, N]$$

A média das múltiplas imagens é dada pela expressão:

$$\bar{g}(x, y) = \frac{1}{N} \sum_{i=1}^N g_i(x, y) \quad (3.19)$$

Com a hipótese de média nula do ruído, a esperança estatística de $g(x, y)$ é:

$$E[\bar{g}(x, y)] = \frac{1}{N} \sum_{i=1}^N E[f(x, y) + n(x, y)] = f(x, y)$$

A variância de $g(x, y)$ é dada por:

$$\sigma_g^2 = E[g^2(x, y)] - E^2[\bar{g}(x, y)]$$

Omitindo-se as coordenadas (x, y) por simplificação de notação, tem-se:

$$\sigma_g^2 = E\left[\frac{1}{N^2} \sum_i \sum_j g_i g_j\right] - f^2$$

Considerando-se que os ruídos n_i, n_j e f não são correlacionados, resulta:

$$E\left[\frac{1}{N^2} \sum_i \sum_j f \times n_j\right] = 0$$

e

$$E\left[\frac{1}{N^2} \sum_i \sum_j n_i f\right] = 0$$

e para $i \neq j$

$$E[n_i n_j] = E[n_i] E[n_j] = 0$$

O que permite obter o resultado final muito importante:

$$\sigma_g^2 = E\left[\frac{1}{N^2} \sum_i \sum_j n_i^2\right] = \frac{1}{N} \sigma^2 \quad (3.20)$$

A variância do ruído da média é dividida pelo número de amostras, o que resulta numa redução do nível de ruído da imagem média em relação a cada imagem original.

O maior problema na aplicação desta técnica é o seu alto custo para armazenar várias amostras da mesma imagem. Além disso, o método não é realista no domínio industrial.

3.3.2 Filtragem

Numa filtragem, o processamento de um nível de cinza de um pixel $P(x, y)$ depende dos valores de nível de cinza desse pixel $P(x, y)$ e de seus pixels vizinhos. Em geral, na vizinhança, os pixels mais próximos de $P(x, y)$ contribuem mais na definição do novo valor de nível de cinza de $P(x, y)$ do que os pixels mais afastados.

As operações de filtragem podem ser divididas em duas classes: a filtragem linear, obtida através de uma operação de convolução, e a filtragem não linear.

3.3.2.1 Transformada de Fourier

3.3.2.1.1 Introdução :

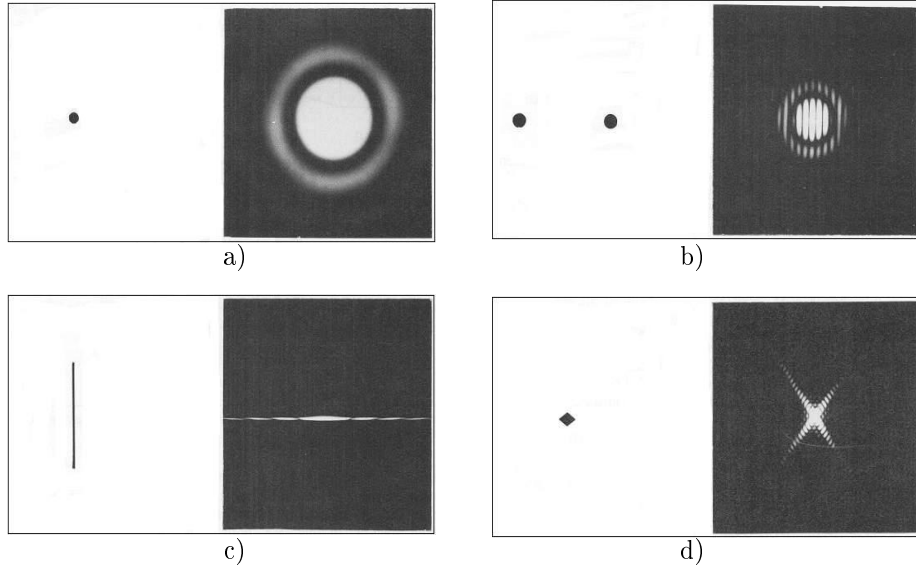


figura 3.6: Exemplos de espectros de Fourier de: a) um ponto b) dois pontos c) uma linha horizontal d) um losângulo

As variáveis de espaço sendo x e y , e as variáveis de frequência u e v , a Transformada de Fourier $\mathcal{TF}(u, v)$ da função $f(x, y)$ escreve-se:

$$\mathcal{TF}(f(x, y)) = \mathcal{TF}(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-2\pi i(ux+vy)} dx dy \quad (3.21)$$

e a transformada inversa escreve-se:

$$\mathcal{TF}^{-1}(f(u, v)) = f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{F}(u, v) e^{2\pi i(ux+vy)} du dv \quad (3.22)$$

Se $f(x, y)$ é uma função real, $\mathcal{TF}(u, v)$ é uma função complexa que aparece na forma de uma soma de termos real $\mathcal{R}(\bullet)$ e imaginário $\mathcal{I}(\bullet)$:

$$\mathcal{TF}(u, v) = \mathcal{R}(u, v) + i\mathcal{I}(u, v) \quad (3.23)$$

A função $\mathcal{TF}^2(u, v) = (\mathcal{R}^2(u, v) + \mathcal{I}^2(u, v))$ representa o espectro de potência do sinal $f(x, y)$, a função $\psi(u, v) = \tan^{-1}(\mathcal{I}(u, v)/\mathcal{R}(u, v))$ representa o ângulo da fase e a função $E(u, v) = \mathcal{TF}^2(u, v)$ representa o espectro de energia.

A figura 3.6 ilustra exemplos de imagens com os seus espectros de Fourier correspondentes.

No nosso caso de imagens digitais quadradas de $N \times N$ pixels, a transformada de Fourier escreve-se:

$$\mathcal{TF}(f(x, y)) = \mathcal{TF}(u, v) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \epsilon^{-2\pi \frac{i(u x + v y)}{N}} \quad (3.24)$$

e tem a propriedade interessante de separabilidade, isto é:

$$\mathcal{TF}(u, v) = 1/N \sum_{x=0}^{N-1} f(x, y) \epsilon^{-2\pi \frac{i u x}{N}} \sum_{y=0}^{N-1} f(x, y) \epsilon^{-2\pi \frac{i v y}{N}} \quad (3.25)$$

seja ainda:

$$\mathcal{TF}(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} N \mathcal{TF}(x, v) \epsilon^{-2\pi \frac{i u x}{N}}$$

onde

$$\mathcal{TF}(x, v) = \left(\frac{1}{N} \sum_{y=0}^{N-1} f(x, y) \epsilon^{-2\pi \frac{i v y}{N}} \right)$$

Essa propriedade de separabilidade é explorada na implementação do algoritmo *FFT Fast Fourier Transform*; após calcular $\mathcal{TF}(x, v)$ a partir de uma *FFT* monodimensional sobre os valores y de $f(x, y)$, é calculado $\mathcal{TF}(u, v)$ pela *FFT* monodimensional sobre os valores de x de $\mathcal{TF}(x, v)$.

3.3.2.1.2 Filtragem pela Transformada de Fourier :

Uma filtragem eficiente por ser realizada usando a Transformada de Fourier quando a imagem apresenta ruído regular e/ou periódico. Essa regularidade ou periodicidade vai se traduzir pela presença no espectro de Fourier de um pico de frequência específica. Removendo esse pico do espectro e efetuando a Transformada de Fourier inversa, será removido a causa desse um pico de frequência e portanto o ruído regular e/ou periódico.

A figura 3.7 ilustra um exemplo de filtragem de imagens pela Transformada de Fourier. A imagem usada corresponde a uma imagem ruidosa de um rosto com linhas oblíquas. O espectro de Fourier correspondente coloca em evidência dois picos de frequência, um acima e a esquerda e outro abaixo e a direita. Removendo esse picos e efetuando a Transformada de Fourier inversa, é gerada uma imagem menos ruidosa sem linhas oblíquas.

3.3.2.2 Filtragem por convolução

Em geral, a convolução discreta de uma imagem $I(x, y)$ com outra imagem $F(x, y)$ (filtro), $x \in [0, M-1]$ e $y \in [0, N-1]$ é uma outra imagem $R(x, y)$, $x \in [0, M-1]$ e $y \in [0, N-1]$, dada por:

$$R(x, y) = I(x, y) ** F(x, y) \quad (3.26)$$

onde o símbolo $**$ representa o produto de convolução e onde

$$R(x, y) = \left(\frac{1}{MN} \right) \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(m, n) F(x - m, y - n) \quad (3.27)$$

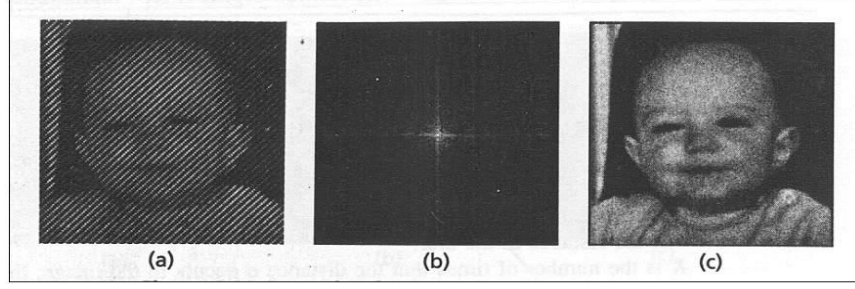


figura 3.7: Exemplos de filtragem pela Transformada de Fourier: a) imagem original com linhas oblíquas, b) espectro de Fourier correspondente, c) imagem filtrada por remoção dos picos de frequência.

e os valores fora dos intervalos $[0, M - 1]$, $[0, N - 1]$ são:

$$A(x, -y) = A(x, N - y)$$

$$A(-x, y) = A(M - x, y)$$

onde A pode ser I ou F .

Pelas propriedades de simetria da Transformada de Fourier, as fórmulas são equivalentes considerando uma imagem $A(x, y)$ definida em $[0, M - 1]$, $[0, N - 1]$ como definida em $[-M, M - 1]$ e $[-N, N - 1]$. De um modo geral, as três imagens $I(x, y)$, $F(x, y)$, $R(x, y)$ são definidas para todo x e y inteiros e se repetem em intervalos de M pixels na vertical e de N pixels na horizontal.

O resultado da convolução de $I(x, y)$ por $F(x, y)$ num pixel $P(x, y)$ é, na realidade, uma média ponderada dos pontos de $I(x, y)$, onde os pesos são dados pela imagem $F(x, y)$.

A resposta impulsiva $F(x, y)$ do filtro tem sempre dimensões finitas, e usando a notação matricial:

$$[F] = \begin{bmatrix} F(1,1) & F(2,1) & \cdots & F(M,1) \\ F(1,2) & F(2,2) & \cdots & F(M,2) \\ \vdots & \vdots & \ddots & \vdots \\ F(1,N) & \cdots & \cdots & F(M,N) \end{bmatrix}$$

A convolução de imagem $I(x, y)$ com a imagem $F(x, y)$ pode ser computada tomando-se as transformadas de Fourier $\mathcal{TF}(I)$ e $\mathcal{TF}(F)$.

A transformada discreta de Fourier da imagem $I(x, y)$ é:

$$\begin{aligned} \mathcal{I}(u, v) &= \mathcal{TF}[I(x, y)] \\ &= \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(m, n) \epsilon^{-2i\pi(\frac{mu}{M} + \frac{nv}{N})} \end{aligned} \quad (3.28)$$

A transformada de Fourier está relacionada com a convolução. A transformada de Fourier da convolução $R(x, y)$ das duas imagens $I(x, y)$ e $F(x, y)$ é igual ao produto das transformadas de Fourier de $I(x, y)$ e $F(x, y)$ como indicado:

$$\mathcal{TF}(R) = \mathcal{TF}(I(x, y) * F(x, y)) = \mathcal{TF}(I) \mathcal{TF}(F) \quad (3.29)$$

A convolução da imagem I é então a transformada inversa do produto, ou seja:

$$R(x, y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \mathcal{R}(u, v) \epsilon^{-i2\pi(xmu/M + ynv/N)}$$

Assim, conhecendo a transformada de Fourier de um filtro $R(x, y)$, pode-se prever o efeito deste quando aplicado a uma imagem qualquer.

Os filtros podem ser divididos em duas grandes classes:

- passa-baixa
- passa-alta

3.3.2.2.1 Filtragem passa-baixa :

Um filtro passa-baixa tem valores próximos de zero para as altas frequências. Portanto, o efeito visual desse tipo de filtro é o de “suavização” (smoothing) da imagem, uma vez que as altas frequências que correspondem às transições abruptas, são atenuadas. A suavização tende também, pelas mesmas razões, a minimizar o efeito do ruído em imagens. Infelizmente, o filtro passa-baixa tem, por outro lado, o efeito indesejado de diminuir a definição e a nitidez da imagem.

No domínio de Fourier, pode-se escrever:

$$\mathcal{R}(u, v) = \mathcal{I}(u, v)\mathcal{H}(u, v) \quad (3.30)$$

onde $\mathcal{H}(u, v)$ chama-se de função de transferência do filtro.

Um filtro passa-baixa ideal tem com função de transferência:

$$\mathcal{H}(u, v) = \begin{cases} 1 & \text{se } D(u, v) \leq D_0 \\ 0 & \text{se } D(u, v) > D_0 \end{cases} \quad (3.31)$$

com $D(u, v) = [u^2 + v^2]^{1/2}$ onde D_0 chama-se de “frequência de corte”. Na prática, este filtro não pode ser realizado, entretanto, pode-se obter uma aproximação com um filtro de Butterworth, cuja resposta impulsiva é (figura 3.8):

$$\mathcal{H}(u, v) = \frac{1}{1 + \left(\frac{D(u, v)}{D_0}\right)^{2n}} \quad (3.32)$$

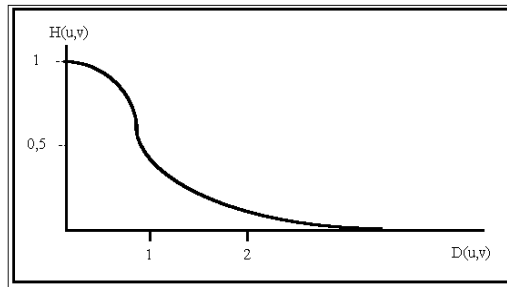


figura 3.8: Função de transferência do filtro de Butterworth

Trabalhando no domínio cartesiano, pode-se definir três classes de filtros:

- a janela uniforme circular (figura 3.9):

$$h(x, y) = \begin{cases} \frac{1}{\pi R^2} & \text{se } \sqrt{x^2 + y^2} < R \\ 0 & \text{caso contrário} \end{cases} \quad (3.33)$$

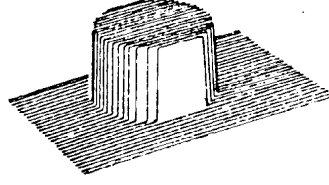


figura 3.9: Filtro uniforme circular

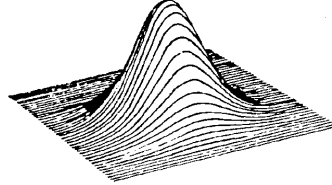


figura 3.10: Filtro gaussiano

- a janela gaussiana (figura 3.10):

$$h(x, y) = \frac{k e^{-\frac{x^2 + y^2}{2\sigma^2}}}{2\pi\sigma^2} \quad (3.34)$$

- a janela exponencial (figura 3.11):

$$h(x, y) = \frac{k e^{-(|x| + |y|)/a}}{4a^2} \quad (3.35)$$

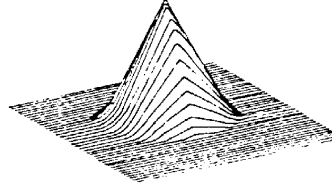


figura 3.11: Filtro exponencial circular

3.3.2.2 Filtragem passa-alta :

Um filtro passa-alta tem valores próximos de zero para as baixas frequências. Entretanto, o efeito visual desse tipo de filtro é o de “agudização” (sharpening) da imagem. As transições entre diferentes regiões da imagem tornam-se mais nítidas. O efeito indesejado é o de enfatizar o ruído que possa existir na imagem.

Um filtro passa-alta ideal pode ser também definido. A resposta impulsiva é do tipo:

$$H(u, v) = \begin{cases} 1 & \text{se } D(u, v) \geq D_0 \\ 0 & \text{se } D(u, v) < D_0 \end{cases} \quad (3.36)$$

com $D(u, v) = \sqrt{u^2 + v^2}$ onde D_0 chama-se de novo a frequência de corte. Esse filtro, não realizável na prática, pode ser aproximado com um filtro de Butterworth

cujas resposta impulsiva é:

$$H(u, v) = \frac{1}{1 + (\frac{Do}{D(u, v)})^{2n}} \quad (3.37)$$

3.3.2.2.3 Mediação local ou Filtragem da Média :

A mediação local, feita a partir de uma imagem, é um exemplo de filtro passa-baixa. Mas, em geral, o produto de convolução não é usado diretamente. O filtro da Média pode ser representado de forma matricial por máscara que vai deslizando sobre a imagem a processar. No caso do filtro da Média, a máscara usada [F2] tem todos os pesos iguais a $\frac{1}{9}$. Cada pixel é substituído pelo valor médio dos vizinhos dele. O tamanho da máscara depende da imagem e da quantidade de ruído nela contido.

Se o ruído depende do sinal, pode ser usada uma vizinhança variável com cada pixel.

Algumas máscaras típicas pela vizinhança constante 3x3 são:

$$[F1] = \frac{1}{5} \begin{vmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{vmatrix}$$

$$[F2] = \frac{1}{9} \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix}$$

$$[F3] = \frac{1}{10} \begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{vmatrix}$$

$$[F4] = \frac{1}{12} \begin{vmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{vmatrix}$$

As máscaras [F1], [F2] efetuam respectivamente a suavização pela média sobre a base de 4-vizinhança e 8-vizinhança. As máscaras [F3], [F4] introduzem uma ponderação variando com a distância e a orientação.

De uma forma geral, o tamanho da vizinhança está relacionado com os valores dos pixels vizinhos. Escolher-se-á a vizinhança grande quando os níveis de cinza dos pixels vizinhos forem de valor fraco, senão optar-se-á por uma vizinhança pequena. Para evitar uma forte mudança da média percebida pela observação visual de uma imagem mais clara ou obscura, o nível de cinza médio do ponto central da vizinhança é de maneira geral inversamente proporcional à soma dos níveis de cinza nessa vizinhança.

A vantagem desse tipo de filtragem é de poder realizar operações rápidas. A maior desvantagem é de obter depois do processamento contornos mais borrados (fig. 3.12-b)). Se o ruído numa imagem aparece como erro aditivo, aleatório ou não correlacionado, os pixels afetados são, pelo filtro da Média, substituídos pela média local que reduzirá as variações dos níveis de cinza mas espalharão o efeito do ruído sobre os níveis de cinza dos outros pixels vizinhos. Para evitar de suavizar muito os contornos, o ideal é realizar operações de suavização somente sobre zonas particulares da imagem. Um método possível é:

- comparar a intensidade de um ponto com a média local;
- substituir esta intensidade pela intensidade média quando a intensidade ultrapassa a média de um limiar.

3.3.2.3 Filtragem não linear

Com os filtros lineares, o ruído aditivo, aleatório ou não correlacionado pode ser reduzido mas este ruído afeta os pixels vizinhos. Mostrar-se-á como, a partir de filtros não-lineares, os efeitos indesejados vistos nos parágrafos anteriores podem ser minimizados.

3.3.2.3.1 Filtragem da ordem e Filtragem da Mediana :

Os filtros da ordem [NR78] [HEY82] tem como objetivo substituir o pixel central de uma janela $w \times w$ pelo K -ésimo valor dos níveis de cinza dessa janela arranjados de acordo com seu valor em ordem crescente. Na prática, um filtro da ordem pode ser implementado da seguinte forma:

Seja uma seqüência $P(1), P(2) \dots, P(N)$ dos $N = w^2$ valores de nível de cinza da janela $w \times w$ e seja a seqüência ordenada $O(1), O(2) \dots, O(N)$ dos $N = w^2$ valores de níveis de cinza da mesma janela $w \times w$. A função definida por:

$$R_{N,k}[imagem] = O(k) \quad (3.38)$$

representa o filtro da ordem $R_{N,k}$.

Os casos mais conhecidos desses arranjos são:

- $k = 1$ conhecido como o operador min
- $k = N$ conhecido como o operador max
- $k = \frac{N+1}{2}$ conhecido como o filtro da **Mediana**.

Os operadores min e max são generalizações de operadores de compressão e de expansão para aplicação em imagens. O uso consiste em aplicar o operador min repetidamente e então aplicar o operador max repetidamente.

Estudos mostraram que filtros de ordem maior que o filtro da **Mediana** aumentam na fronteira regiões claras da imagem proporcionalmente a sua ordem. Assim, objetos claros e pequenos vistos com dificuldade na imagem original aparecem mais nítidos pelo filtro max que aumenta a área. De forma similar, os filtros de ordem menor que $\frac{N+1}{2}$ aumentam regiões escuras e pequenas.

Numa forma prática, a mediana é calculada a partir das intensidades dos pixels contidos em uma vizinhança de um ponto $P(x, y)$. No caso de uma vizinhança 3×3 , $P(x, y)$ tem 8 vizinhos. Todos os nove pontos são ordenados, e é tomado como novo valor de $P(x, y)$, o valor da mediana dessa ordenação. Pode-se verificar que, quando aparecem valores fora do domínio da ordenação dos valores na vizinhança, a mediana da distribuição não é alterável.

Em geral, esse tipo de filtro destrói pontos isolados e, às vezes, linhas respeitando as transições. Além disso, ele é eficiente para ruídos do tipo pulso. Se o ruído numa imagem aparece como erro aditivo, aleatório ou não correlacionado, os pixels afetados são, pelo filtro da **Mediana**, corrigidos com a vantagem de evitar espalhar o efeito do ruído para os pixels vizinhos. Com o aumento do tamanho da janela, os resultados do filtro da média continuam péssimos, a perda de nitidez é maior ainda. Enquanto o filtro da mediana tem tendência a eliminar pequenas regiões sem diminuir a nitidez da imagem. A figura 3.12) ilustra bem as diferenças de eficiência

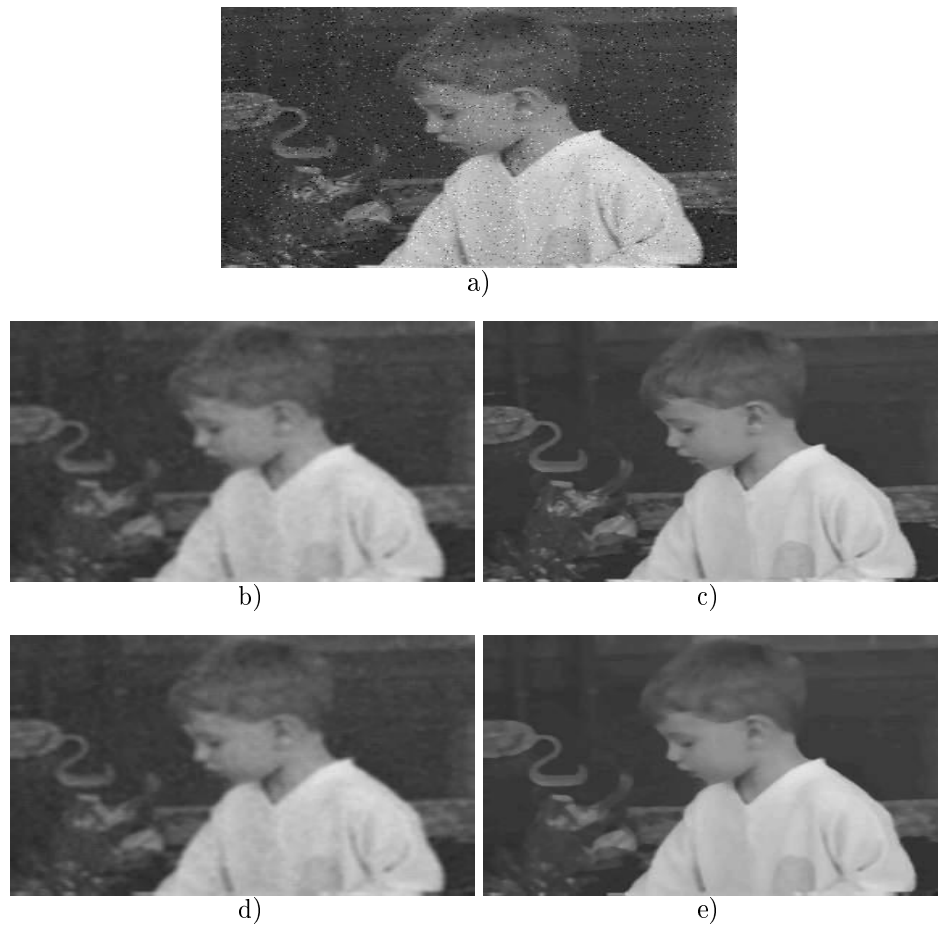


figura 3.12: Comparação dos filtro da média e do filtro da mediana: a) imagem original, b) média com uma janela 3×3 , c) mediana com uma janela 3×3 , d) média com uma janela 5×5 , e) mediana com uma janela 5

entre o filtro da média e o da mediana.

Na figura 3.13 é ilustrada a ação de um filtro linear e de um filtro da **Mediana** monodimensional sobre alguns sinais teste. É possível notar que o filtro da **Mediana** respeita melhor o sinal de entrada, o que demonstra que representa um método muito poderoso que permite diminuir o problema da falta de nitidez.

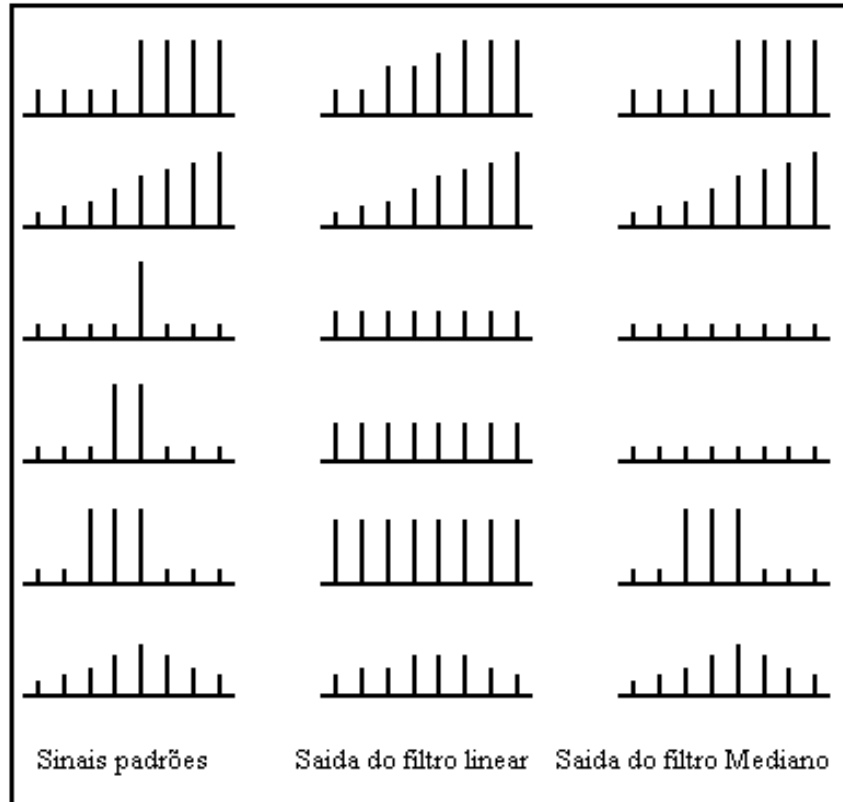


figura 3.13: Respostas do filtro linear e da mediana a alguns sinais de teste

Quando a preocupação de manter as transições nítidas faz-se mais forte e o filtro clássico da **Mediana** não permite resolver o problema, o uso de vizinhanças mais específicas pode auxiliar a respeitar melhor as transições. Por exemplo, o uso de uma vizinhança cruciforme permitem preservar melhor as linhas verticais e horizontais.

3.3.2.3.2 filtragem de Nagao-Matsuyama :

[NM79] propuseram um método de suavização que associa a cada ponto da imagem $P(x,y)$ a ser filtrado, a partir de máscaras localizadas, o nível de cinza médio dos elementos numa vizinhança 5×5 centrada no ponto $P(x,y)$. A partir de 9 máscaras (figura 3.14), das quais 8 são assimétricas, são avaliados 9 médias e 9 desvios. O princípio desse filtro é substituir o pixel central $P(x,y)$ da vizinhança por um novo pixel cujo valor é a média da máscara que fornece o desvio mínimo, seja ainda a menor variância.

Esse filtro realiza o compromisso difícil de suprimir os pontos isolados, reduzir o ruído e ao mesmo tempo realçar as transições e a nitidez da imagem original.

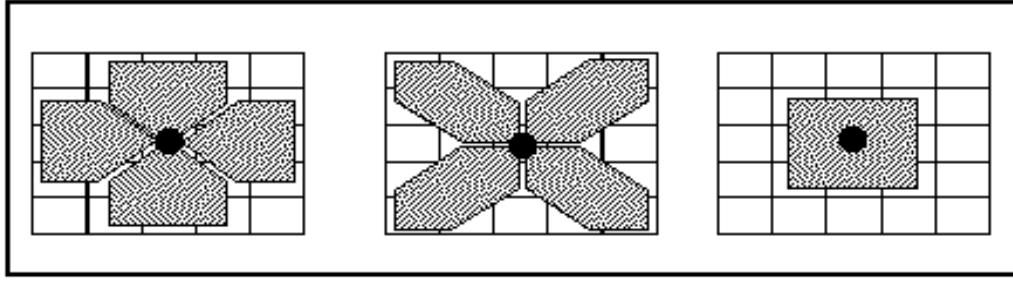


figura 3.14: Representação das 9 máscaras do filtro não linear de Nagao-Matsuyama

3.3.2.3.3 filtragem homomórfica :

A intensidade de uma imagem pode ser expressa em termos do produto de uma componente proveniente da iluminância da cena por uma outra componente proveniente da reflectância dessa cena:

$$f(x, y) = i(x, y) \times r(x, y) \quad (3.39)$$

onde $i(x, y)$ representa o termo de iluminância com uma grande dinâmica e uma componente de baixa frequência e $r(x, y)$ representa o termo de reflectância com uma fraca dinâmica. Ele fornece a informação própria aos objetos e varia bastante na junção de objetos diferentes.

Seja a função $z(x, y) = \ln f(x, y)$, então

$$z(x, y) = \ln i(x, y) + \ln r(x, y)$$

Usando a transformada de Fourier:

$$\mathcal{Z}(u, v) = \mathcal{I}(u, v) + \mathcal{R}(u, v)$$

Efetuada uma transformação pelo filtro de resposta impulsiva $\mathcal{H}(u, v)$:

$$\begin{aligned} s(x, y) &= \mathcal{F}(\mathcal{H}(u, v)\mathcal{Z}(u, v)) \\ &= i'(x, y) + r'(x, y) \end{aligned}$$

Obte-se afinal $g(x, y)$ definida da seguinte maneira:

$$g(x, y) = \epsilon^{s(x, y)} = iO(x, y)rO(x, y) \quad (3.40)$$

Essa técnica chama-se de filtragem homomórfica. A reflectância pode ser extraída por meio de uma filtragem passa-altas. Em geral, $\mathcal{H}(u, v)$ é escolhido de maneira a trabalhar com baixas e altas frequências (do tipo filtragem passa-faixa). O tratamento da informação é feito segundo a não linearidade da retina no sistema visual humano.

3.3.2.4 Realce de bordas e detalhes

Qualquer captor introduz uma falta de nitidez mais visível especificamente pelas bordas e detalhes, que representam componentes de alta frequência. Um método de realce das bordas e detalhes é usar um operador de diferenciação que vai amplificar as altas frequências. Não deve-se esquecer que o ruído é também um componente de alta frequência. Portanto, um tratamento preliminar é necessário quando a imagem é ruidosa para evitar que a diferenciação amplifique ao mesmo tempo o ruído e as bordas.

3.3.2.5 Operador de diferenciação

Um operador de diferenciação é baseado sobre a noção de derivada. Uma imagem $I(x, y)$ é uma seqüência de regiões de tamanhos diversos, onde os níveis de cinza são constantes, e de fronteiras, encontros de várias regiões, marcando uma transição entre elas. Do ponto de vista matemático, as regiões são caracterizadas por derivadas suaves e pouco acentuadas e as fronteiras por derivadas fortes e acentuadas. O princípio é derivar a imagem para acentuar as bordas contidas na imagem. As derivadas de $f(x)$ apresentam

Um ponto muitas vezes importante no processo de diferenciação (capítulos da segmentação e da interpretação) é construir operadores que sejam invariantes pelas transformações $t(\bullet)$ do tipo translação, rotação e mudança de escala. Um operador $O(\bullet)$ é dito invariante pela função $t(\bullet)$ se e somente se aplicar o operador depois da função conduz ao mesmo resultado que realizar a função depois de ter aplicado o operador. Para uma imagem $I(x, y)$, tem-se:

$$t[O(I)] = O[t(I)] \quad (3.41)$$

O interesse desse tipo de operador é permitir movimentar-se em qualquer direção depois do realce das bordas. Pode-se mostrar que:

- tais operadores diferenciais lineares não podem ter unicamente derivadas pares
- no caso de quaisquer operadores, as derivadas ímpares podem estar presentes na condição de serem elevadas com potências pares.

Os operadores de diferenciação são invariantes pelas transformações de translação e de troca de escala. Não são invariantes frente à rotação. Uma rotação é definida:

$$\begin{aligned} u &= x \cos \alpha + y \sin \alpha \\ v &= -x \sin \alpha + y \cos \alpha \end{aligned}$$

onde (x, y) e (u, v) representam respectivamente as coordenadas iniciais e as coordenadas após a rotação.

As derivadas parciais de uma imagem I são:

$$\begin{aligned} \frac{\partial I(x, y)}{\partial x} &= \frac{\partial I(x, y)}{\partial u} \cos \alpha + \frac{\partial I(x, y)}{\partial v} \sin \alpha \\ \frac{\partial I(x, y)}{\partial y} &= \frac{\partial I(x, y)}{\partial u} \sin \alpha + \frac{\partial I(x, y)}{\partial v} \cos \alpha \end{aligned}$$

Estas derivadas não são invariantes pela rotação. Entretanto, se utilizarmos derivadas de segunda ordem (Laplaciano), a invariância é verificada:

$$\left(\frac{\partial I(x, y)}{\partial u}\right)^2 + \left(\frac{\partial I(x, y)}{\partial v}\right)^2 = \left(\frac{\partial I(x, y)}{\partial x}\right)^2 + \left(\frac{\partial I(x, y)}{\partial y}\right)^2$$

3.3.2.5.1 Gradiente :

Verificou-se que a derivada em uma direção qualquer é função das derivadas parciais nas direções x e y . Pode-se determinar, num ponto dado $P(x, y)$, a direção α_0 onde a derivada de $I(x, y)$ atinge um máximo. Considerando que o novo sistema de referência é centrado em $P(x, y)$, $(P(x, y), u, v)$, e no qual os eixos são obtidos por rotação dos eixos de referência (x, y) em relação ao ângulo α , tem-se :

$$\begin{aligned} u &= x \cos \alpha + y \sin \alpha \\ v &= -x \sin \alpha + y \cos \alpha \end{aligned}$$

onde (x, y) e (u, v) representam respectivamente as coordenadas iniciais e as coordenadas após a rotação por α .

Derivando $\partial I(x, y)/\partial u$ por α , tem-se:

$$\frac{\partial I(x, y)}{\partial u} = 0 \rightarrow -\frac{\partial I(x, y)}{\partial x} \sin \alpha + \frac{\partial I(x, y)}{\partial y} \cos \alpha = 0$$

Se α_0 representa a direção do máximo da derivada de $I(x, y)$, então α_0 é dado por:

$$\alpha_0 = \arctan\left(\frac{\partial I(x, y)/\partial y}{\partial I(x, y)/\partial x}\right)$$

$\alpha + \pi$ é também uma solução. Pode-se verificar que em uma direção, tem-se um máximo, na outra um mínimo.

$$\begin{aligned} \cos \alpha_0 &= (1 + \tan^2 \alpha)^{-\frac{1}{2}} \\ \sin \alpha_0 &= (1 - \tan^2 \alpha)^{-\frac{1}{2}} \end{aligned}$$

O valor máximo de $\partial I(x, y)/\partial u$ é obtido pela equação:

$$\frac{\partial I(x, y)}{\partial x} \cos \alpha_0 + \frac{\partial I(x, y)}{\partial y} \sin \alpha_0 = \sqrt{\left(\frac{\partial I(x, y)}{\partial x}\right)^2 + \left(\frac{\partial I(x, y)}{\partial y}\right)^2}$$

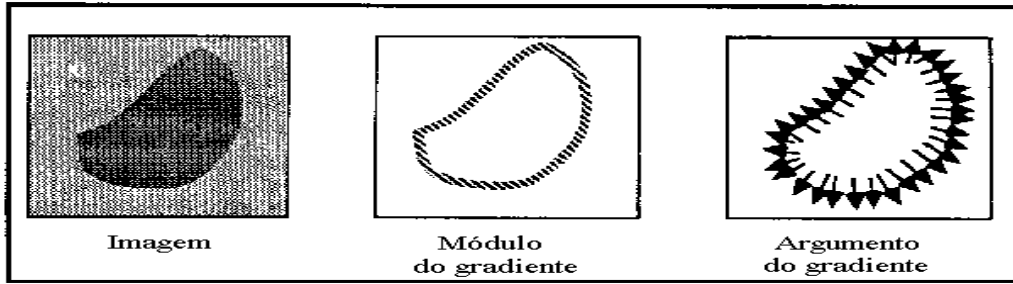


figura 3.15: Interpretação das informações extraídas pelo gradiente

O vetor desta amplitude na direção α chama-se vetor gradiente G num ponto $P(x, y)$ da imagem $I(x, y)$. Esse vetor G representa a direção da maior variação da descontinuidade ao ponto de cálculo (figura 3.15). O vetor G nesse ponto é caracterizado pelo módulo e pelo argumento :

- **módulo:**

$$\begin{aligned} ||G[P(x, y)]|| &= (G_x^2 + G_y^2) \\ &= \sqrt{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2} \end{aligned} \quad (3.42)$$

- **argumento:**

$$\begin{aligned} \alpha &= \arctan\left(\frac{G_y}{G_x}\right) \\ &= \arctan\left[\left(\frac{\partial I}{\partial y}\right)/\left(\frac{\partial I}{\partial x}\right)\right] \end{aligned} \quad (3.43)$$

No caso de uma imagem digital, é preciso aproximar as derivadas parciais num ponto de uma vizinhança usando os níveis de cinza. Existem vários operadores de avaliação das derivadas parciais. Os operadores que dão uma boa aproximação levam à 8-vizinhança, por exemplo:

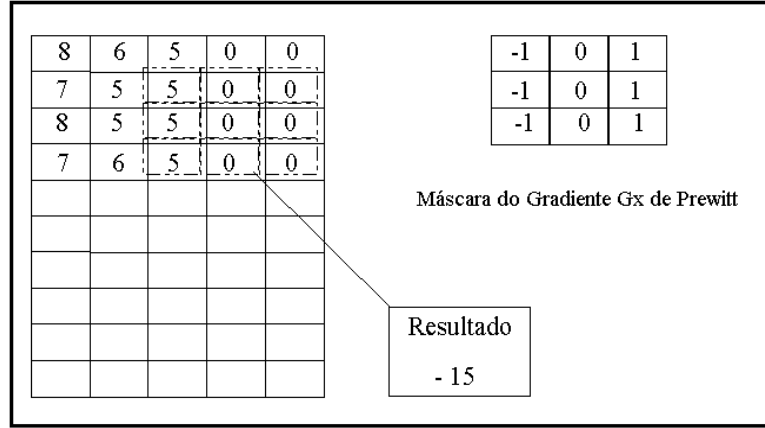


figura 3.16: Exemplo numérico da aplicação do operador Gradiente Gx de Prewitt

3.3.2.5.2 Operador de Prewitt :

$$[G_x] = \begin{vmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{vmatrix}$$

$$[G_y] = \begin{vmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{vmatrix}$$

De forma prática, para construir uma “*imagem gradiente*” da imagem original, a operação consiste em deslizar sobre toda a imagem uma janela do tamanho da máscara. Num pixel dado, os valores da máscara são multiplicados pelos valores dos níveis de cinza dos pontos correspondentes vizinhos desse pixel. A soma dessas multiplicações representa o valor do operador aplicado ao pixel central dessa vizinhança (figura 3.16).

3.3.2.5.3 operador 1 de Sobel :

$$[G_x] = \begin{vmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{vmatrix}$$

$$[G_y] = \begin{vmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{vmatrix}$$

3.3.2.5.4 operador 2 de Sobel :

$$[G_x] = \begin{vmatrix} 2 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -2 \end{vmatrix}$$

$$[G_y] = \begin{vmatrix} 0 & 1 & 2 \\ 1 & 0 & -1 \\ -2 & -1 & 0 \end{vmatrix}$$

A construção da imagem gradiente pelos operadores 1 e 2 de Sobel segue a mesma filosofia que no caso do operador de Prewitt.

A figura 3.17 ilustra o gradiente numa imagem em níveis de cinza com operador 1 de Sobel.

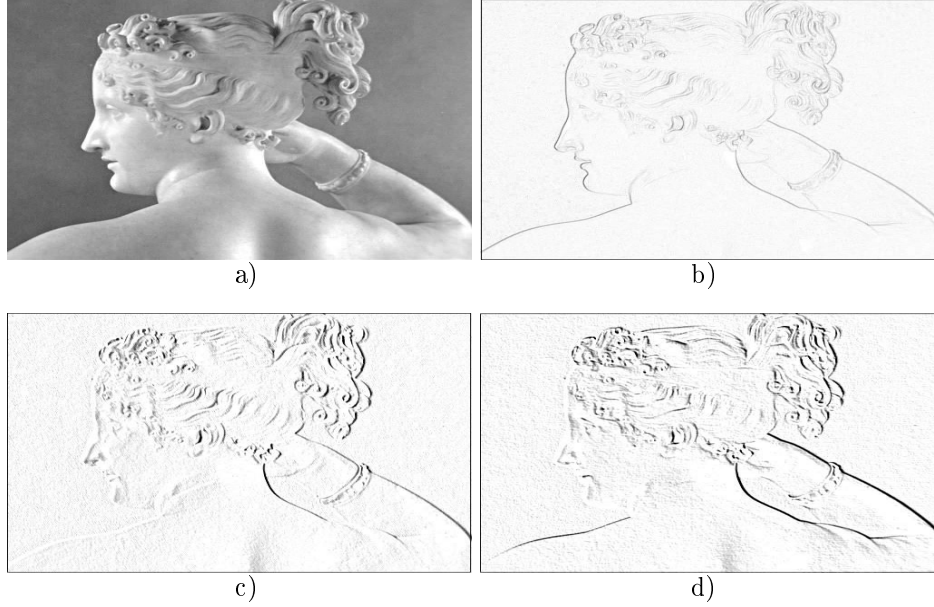


figura 3.17: Exemplo de gradiente pelo operador 1 de Sobel: a) imagem original, b) módulo do gradiente, c) gradiente G_x , c) gradiente G_y

3.3.2.5.5 Laplaciano :

O Laplaciano é:

$$\Delta^2 I(x, y) = L[I(x, y)] = \frac{\partial^2 I(x, y)}{\partial x^2} + \frac{\partial^2 I(x, y)}{\partial y^2} \quad (3.44)$$

A intensificação das bordas pode ser realizada por subtração da imagem $I(x, y)$ com a derivada segunda de $I(x, y)$. É uma operação bem conhecida em fotografia e chamada “unsharp masking” ($I - \Delta^2 I$) ilustrada no exemplo da figura 3.18 em um sinal unidimensional:

I	:	0	0	0	1	2	3	4	5	5	5	5	5	5	6	6	6	6	3
$\Delta^2 I$	:	0	0	1	-1	0	0	1	-1	0	0	0	0	1	-1	0	0	0	3
$I - \Delta^2 I$	:	0	0	-1	2	2	3	3	6	5	5	5	5	4	7	6	6	6	0

figura 3.18: Exemplo numérico de realce de bordas pelo Laplaciano num sinal unidimensional

Para uma imagem digital, o Laplaciano calculado em um ponto pode ser aproximado da seguinte maneira:

$$\begin{aligned} \Delta^2 I(x, y) &= L[I(x, y)] \\ &= \frac{\partial^2 I(x, y)}{\partial x^2} + \frac{\partial^2 I(x, y)}{\partial y^2} \end{aligned}$$

$$= I(x+1, y) + I(x-1, y) + I(x, y+1) + I(x, y-1) - 4I(x, y)$$

O que, de forma prática, faz-se usando as máscaras definidas da figura 3.19.

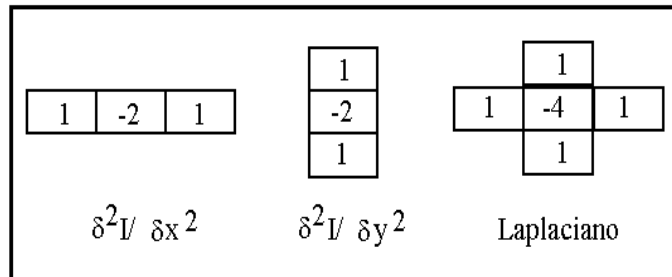


figura 3.19: Máscara do Laplaciano

De forma prática, para construir uma “*imagem laplaciana*” a partir da imagem original, reitera-se o processo utilizado para avaliar o Gradiente. Uma janela é centrada sobre o pixel, e os valores da máscara são multiplicados pelos valores dos pontos correspondentes vizinhos do pixel. A soma dessas multiplicações representa o valor do operador aplicado ao pixel centro dessa vizinhança. Um exemplo numérico do cálculo do Laplaciano é ilustrado na figura 3.20.

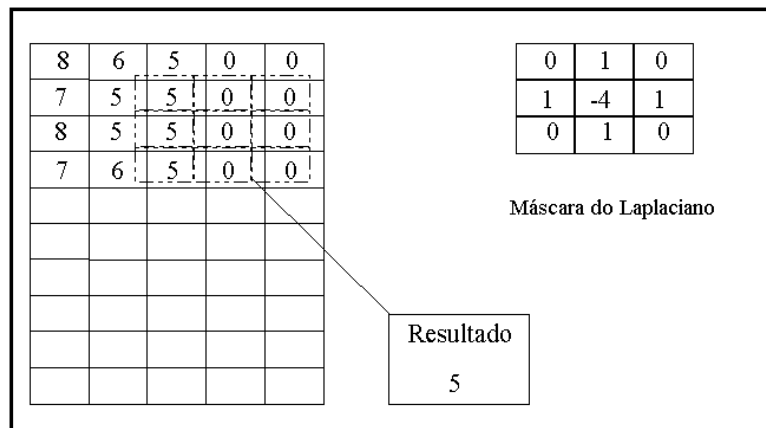


figura 3.20: Exemplo numérico da aplicação do operador Laplaciano

A figura 3.21 ilustra o Laplaciano numa imagem em níveis de cinza.

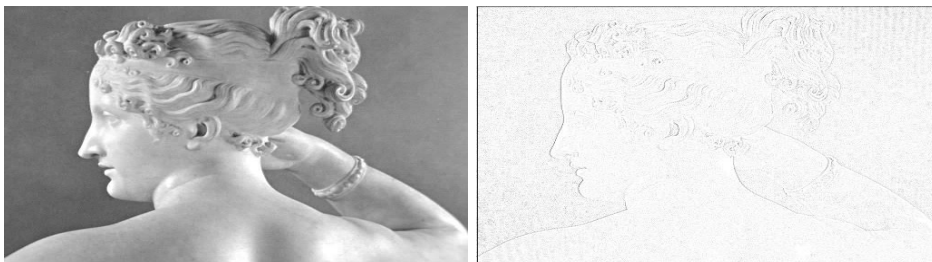


figura 3.21: Exemplo de Laplaciano

Capítulo 4

SEGMENTAÇÃO DE IMAGENS

4.1 INTRODUÇÃO

Alguns psicólogos alemães no início do século (Khler, Wertheimer, Koffka) introduziram o princípio da segmentação. Eles mostraram que o ser humano, a nível da visão, efetua agrupamentos sobre o que ele percebe, baseados sobre a proximidade, a similaridade, e a continuidade. Segundo alguns estudos baseados na teoria da visão (Gibson), a decomposição de uma cena visual em “unidades significantes” constitui uma etapa decisiva na compreensão da imagem.

Como em Robótica, onde alguns robôs têm algumas características antropomórficas (ombro - braço - pulso - dedos de preensão), a visão por computador tenta adotar um funcionamento análogo ao da visão humana. A segmentação traduz esse esforço.

O alvo da segmentação é obter, a partir de uma imagem digitalizada pré-processada, um conjunto de “primitivas ou segmentos significativos” que contém a informação semântica relativa à imagem de origem.

Em geral, as primitivas usadas são “naturais”: contornos e regiões. Os diferentes tratamentos usados levam à noção de contorno, de região ou a uma noção mista. Essa noção é preferível, mas é mais complicada porque a passagem região-contorno ou contorno-região é mais fácil em teoria que na prática. Muitas vezes, a noção unilateral é escolhida em razão dos tempos de execução mais curtos.

4.2 SEGMENTAÇÃO POR REGIÃO

A detecção de regiões numa imagem pode ser feita com um dos dois objetivos:

- extrair uma determinada região ou
- dividir (particionar) a imagem num conjunto de regiões disjuntas cuja união representa a imagem inteira.

Uma região de numa imagem é um conjunto de pontos “ligados” onde, de qualquer ponto da região pode-se chegar a qualquer outro ponto por um caminho inteiramente contido nessa região. As regiões que se deseja detectar, em geral, são regiões “homogêneas”, ou seja, apresentam alguma propriedade local aproximadamente constante. Comumente esta propriedade é a continuidade no nível de cinza.

Um exemplo de função exprimindo essa propriedade de continuidade sobre uma região $R(x, y)$ pode ser:

- $R(x, y)$ é homogênea se e somente se, para qualquer dupla de pontos, a diferença entre os níveis de cinza dos dois pontos não for maior que uma constante.

Os diferentes métodos para particionamento em regiões são:

- limiarização ou binarização (threshold em inglês)
- divisão e fusão (split and merge em inglês)

4.2.1 Limiarização

4.2.1.1 Introdução

O processo de digitalização gera uma imagem composta pixels cujos níveis de cinza vão de 0 até 255. De forma geral, lida-se com imagens contendo objetos, células, texto, gráficos, todo tipo de informação. A partir de agora, chamar-se-á o primeiro plano de objetos (foreground em inglês) as regiões de interesse. E de fundo (background em inglês) todas as outras regiões consideradas como não interessantes. Para ilustrar esses conceitos, seja o exemplo de uma imagem de cheque bancário. Um documento do tipo cheque bancário contém informações importantes (como número do banco, da agência, da conta etc..) sobrepostos a um fundo geralmente complexo para evitar cópias fraudulentas do cheque. As informações importantes constituem a região do primeiro plano enquanto o resto constitui o segundo plano ou fundo. É óbvio que em função da imagem processada, essa noção de primeiro plano e de fundo pode mudar.

O princípio da limiarização consiste em separar o primeiro plano do fundo. O objetivo da binarização é marcar os pixels que pertencem às verdadeiras regiões de primeiro plano com um valor simples (ON) e as regiões de fundo (segundo plano) com uma intensidade diferente (OFF). E assim gerar o que se chama de imagem binária que é uma imagem preto-branca apresentando somente duas classes, o primeiro plano de cor preta sobre o fundo de cor branca ou o contrário.

A limiarização é baseada na utilização do histograma (figura 3.1 no parágrafo 3.2.1). No caso do primeiro plano ser bem diferenciado em relação ao fundo, o histograma apresenta distintamente duas classes na forma de dois picos separados por um “vale”.

Limiarizar consiste, no caso de um objeto mais claro que o fundo, em escolher no histograma um valor T , tal que para um pixel $P(x, y)$:

$$\begin{array}{ll} \text{se } I[P(x, y)] > T & \implies P(x, y) \in \text{objeto} \\ \text{senão} & \implies P(x, y) \in \text{fundo} \end{array}$$

Esse limiar T fica em algum lugar do vale. A figura 4.1 ilustra a influência do valor do limiar sobre a qualidade da limiarização.

Este processo é muitas vezes usado para dissociar o fundo das identidades presentes (células, objetos, nós de madeira etc..). A maior vantagem é que o princípio é simples e precisa de pouca memória. Mas a escolha do limiar é delicada e fica muito empírica. Ela depende muito da qualidade dos contrastes, da qualidade e da natureza da imagem. A qualidade da imagem é função do fator de refletância e do

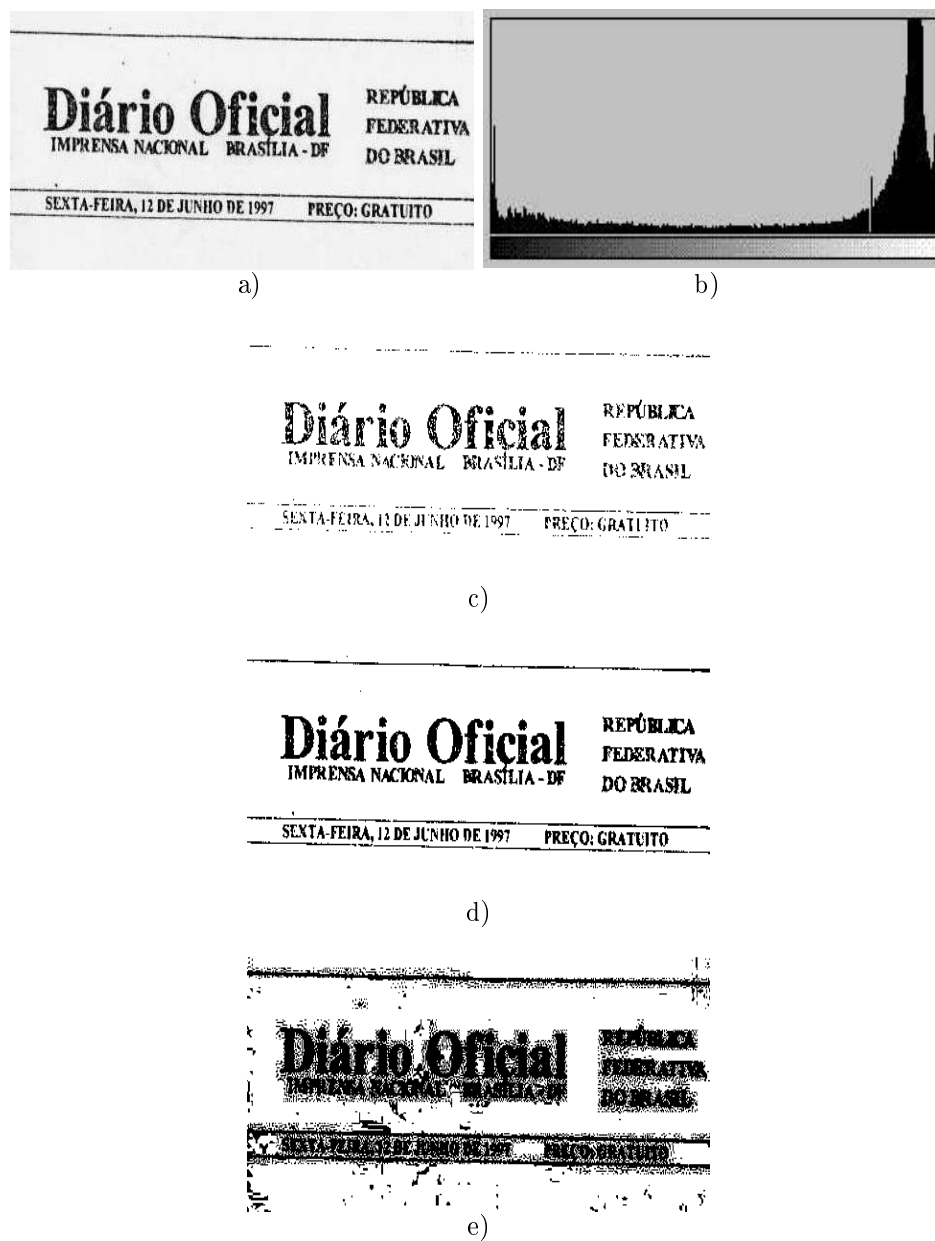


figura 4.1: Influência do valor do limiar sobre a qualidade da limiarização: a) imagem original, b) histograma, c) limiarização com valor 50, d) limiarização com valor 150 e) limiarização com valor 235

fator de iluminação. Esse último influencia fortemente a natureza dos picos do histograma. No caso de uma iluminação não uniforme sobre toda a cena, a separação dos picos e o estabelecimento do limiar são mais difíceis.

Se os valores dos níveis de cinza dos pixels do primeiro plano e aqueles do fundo são razoavelmente consistentes em seus respectivos valores na imagem inteira, então um único valor de limiar pode ser encontrado para esta imagem. Nesse caso chama-se de limiarização global. Na literatura podem ser encontrados muitos métodos de limiarização global.

Contudo, para muitas aplicações, devido a problemas de iluminação e de falta de qualidade da imagem, um único valor de limiar não permite resultar numa limiarização de boa qualidade. Acontece também que, devido a complexidade da imagem, valores diferentes de limiar são requeridos para regiões diferentes. O que exige buscas diferenciadas. Este tipo de limiarização é chamado de limiarização adaptativa ou ainda local.

4.2.1.2 Limiarização global

Como já foi explicado anteriormente, o modo mais direto para selecionar automaticamente um valor global de limiar é através do uso de um histograma das intensidades dos pixels na imagem. O histograma de intensidade traça o número de pixels com valores de cada nível de intensidade. Para uma imagem com intensidades de primeiro e segundo planos bem diferenciados, o histograma terá dois picos distintos. O vale entre estes picos pode ser encontrado como o mínimo entre dois máximos, e o valor da intensidade aí é escolhido como o limiar que melhor separa os dois picos. Já que os métodos de detecção de vales e picos estão propensos a localizar o local máximo e o mínimo diferentes dos desejados, um método [Rosenfeld e De La Torre [RT83]] ajusta um envelope convexo ao histograma bimodal e então detecta o vale como a máxima distância entre a cobertura e a altura da caixa.

Existe um número de desvantagens para a seleção do limiar global baseada na forma da distribuição da intensidade. A primeira é que as imagens nem sempre contêm intensidades de primeiro e segundo planos bem diferenciados como um resultado de um pobre contraste e ruído. A segunda é que, especialmente para uma imagem de componentes de primeiro plano esparsos, tais como a maioria das imagens gráficas, o pico representando estas intensidades será muito menor do que o pico das intensidades de segundo plano (fundo). Esta diferença freqüentemente torna difícil encontrar o vale entre estes dois picos. Além disso, métodos de detecção de vales e picos confiáveis são problemas separados entre si.

Uma maneira de melhorar este enfoque é compilar um histograma das intensidades dos pixels que são ponderados pelo inverso dos seus valores de comprimento das bordas [Mason [Mas75]]. As intensidades de regiões de pixels com valores de bordas baixos serão mais ponderados do que as intensidades das bordas e os pixels de ruído com valores das bordas mais altos, cortando, deste modo os picos do histograma causados por estas regiões e facilitando a detecção do limiar entre eles.

Uma técnica análoga é ponderar altamente estas intensidades de pixels com valores altos de bordas e então escolher o limiar no pico do histograma, correspondendo a transição entre as regiões [Katz [Kat65]] [Watanabe [Wat74]] [Weszka, Nagel e Rosenfeld [WNR74]]. Este método requer a detecção do pico de um simples máximo, o qual é freqüentemente mais fácil do que a detecção de vales entre

dois picos. Este enfoque também reduz o problema de uma grande discrepância em termos de tamanho entre os picos das regiões de primeiro e segundo planos, porque os pixels de borda estão acumulados no histograma ao invés dos pixels de região; a diferença entre uma área de tamanho grande e pequena é uma quantidade linear de bordas versus uma quantidade muito maior para as regiões.

Um outro método utiliza a ponderação Laplaciana. O Laplaciano é um operador de segunda derivada, o qual ponderará altamente as transições das regiões em bordas (a primeira derivada pondera altamente as bordas). Este método irá ponderar altamente os pixels de borda de ambas as regiões de primeiro plano e seus fundos (segundo plano) circunvizinhos; assim, o histograma terá dois picos de áreas similares. Embora estas técnicas de histograma ofereçam a vantagem de que as detecções pico e vale são intuitivas, a detecção de picos é suscetível ao erro devido ao ruído e a regiões pobremente separadas. Além disso, quando a classe de primeiro plano e a de fundo consistem em várias regiões estreitas - tais como - textos e bordas, a medida do Laplaciano pode ser pobre por causa de transições muito abruptas (bordas estreitas) entre o primeiro e o segundo plano.

Um número de técnicas determinam as classes através do reconhecimento de padrões formal que otimiza algumas medidas de separação (figura 4.2). Um enfoque é o limiar do erro mínimo [Kittler e Illingworth [SSW88]] [Ye e Danielson [HD88]].

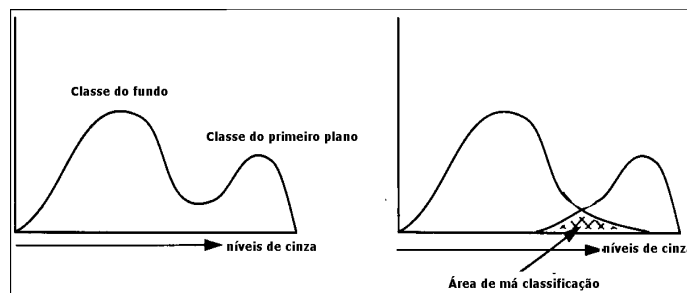


figura 4.2: Binarização a partir de funções de densidade de probabilidade Gaussianas.

Aqui, as distribuições de intensidade do primeiro e do segundo plano são modeladas como funções de densidade de probabilidade normal (Gaussiana ou forma de sino). Para cada valor de intensidade (de 0 a 255, ou uma faixa muito pequena se o limiar é conhecido e é limitada por ele), as médias e as variâncias são calculadas para as classes do primeiro e do segundo plano, e o limiar é escolhido tal que o erro da não-classificação entre as duas classes seja minimizado. O último método é classificado como uma técnica paramétrica pela suposição que a distribuição da escala de nível de cinza pode ser modelada como uma função densidade de probabilidade. Este é um método popular para várias aplicações em visão computacional, mas alguns experimentos indicam que certas imagens não seguem bem este modelo. Deste modo, os resultados com este método são mais pobres do que com abordagens não paramétricas [Abutaleb [Abu89]].

Uma abordagem bem conhecida e eficiente não paramétrica é o método de [Otsu [SSW88]] [Reddi, Rudin e Keshavan [RRK84]]. Primeiro os cálculos são feitos da relação entre as classes de variância (interna e externa) para cada valor de limiar em potencial. As classes aqui são os pixels de primeiro e segundo planos, e o objetivo

é encontrar o limiar que maximiza a variância das intensidades entre duas classes e as minimiza dentro de cada classe. Esta relação (razão) é calculada para todos os níveis de limiar em potencial. Nível este no qual a razão é máxima se é o limiar escolhido.

Um outro enfoque similar ao de Otsu emprega uma medida teórica de informação, chamada de entropia, a qual é a medida da informação de uma imagem expressada como a média do número de bits requeridos para representar a informação [Abutaleb [Abu89]] [Johannsen e Bille [JB82]] [Sahoo e Wong [SSW88]] [Pun [Pun81]]. Aqui a entropia para duas classes é calculada para cada limiar em potencial e o limiar cuja soma de duas entropias é o maior, é escolhido como o melhor.

4.2.1.3 Limiarização local adaptativa

Uma forma comum de executar a limiarização adaptativa é analisando as intensidades de níveis de cinza dentro de uma janela local sobre a imagem para determinar limiares locais. [Bernsen [Ber86]] [Casey e Wong [CW94]] [Kamel e Zhao [KZ93]] [Wong [Won78]]. [White e Rohrer [WR83]] descrevem um algoritmo para a tal limiarização onde o limiar é alterado continuamente através da imagem estimando o nível de fundo como uma média móvel bidimensional dos valores locais dos pixels tomados para todos estes pixels nessa imagem.

[Mitchell e Gillies [MG89]] descrevem um método similar de limiarização onde uma normalização de nível branco no fundo é feita pela primeira estimativa de nível de branco e subtraindo este nível da imagem original. Então, a segmentação de caracteres é realizada aplicando uma faixa de limiares e selecionando a imagem resultante com um conteúdo mínimo de ruído. Este conteúdo é medido como a soma das áreas ocupadas pelos componentes mais pequenos e finos do que empiricamente através de parâmetros determinados. Olhando para os resultados da binarização da fig. 1 anteriormente mostrados, pode-se observar que a escolha do melhor limiar produz um ruído mínimo visível.

O problema principal com qualquer técnica de binarização adaptativa é a escolha do tamanho da janela. O tamanho escolhido deve ser grande o suficiente para garantir que o número de pixels de fundo (segundo plano) incluídos sejam também grandes o suficiente para obter uma boa estimativa do valor da média, mas não tão grande como a média sobre as intensidades de segundo plano (fundo) não uniformes. Porém, as características na imagem freqüentemente variam em tamanho tal que existem problemas com o tamanho da janela fixo. Para remediar este problema, a informação dependente de domínio pode ser usada para verificar que os resultados da binarização geram as características esperadas (por exemplo, uma bolha grande de uma região com valor ON não é esperada na página dos símbolos pequenos). Se o resultado é inesperado, então o tamanho da janela pode ser modificado e a binarização é aplicada outra vez.

4.2.1.4 A escolha de um método de limiarização

Se os métodos de limiarização global e adaptativo são usados para binarização, não se pode nunca esperar resultados perfeitos. Dependendo da qualidade do original, pode haver buracos nas linhas, bordas rompidas na região limítrofes e regiões de pixels estranhas com valores *ONe OFF*. A suposição de que os resultados do processamento não serão perfeitos é geralmente verdadeiro com outros métodos de

processamento da imagem em geral.

O procedimento recomendado é processar tanto quanto possível cada passo, mas adiar decisões que não necessitam ser executadas até passos recentes, para evitar erros irreparáveis. Nessas últimas etapas, existe mais informação do resultado do processamento do que aquele ponto e esta informação provê o maior contexto e descrições de alto nível para auxiliar na tomada de decisões corretas e por conseguinte, o reconhecimento. O adiamento, quando possível, é um princípio apropriado para todos os passos da análise de documentos (exceto, é claro, o último).

Um número de métodos diferentes de limiarização foram apresentados nesta seção. Nenhum método é melhor para todos os tipos de imagens e aplicações. Para problemas mais simples, onde as características da imagem não variam muito dentro dela ou entre diferentes imagens, os métodos mais simples serão suficientes. Para problemas mais difíceis de ruído ou características da imagem variantes, métodos mais complexos geralmente serão requeridos.

A melhor maneira de escolher um método neste momento é primeiro estreitando as escolhas através de descrições de métodos e então apenas ir experimentando-os com os outros e examinando seus resultados.

4.2.2 Método de Divisão e Fusão

O método de Divisão e Fusão permite agrupar pixels para criar regiões sobre as quais um predicado de homogeneidade H é verificado [CMVM86]. A imagem será manipulada através de divisões e fusões até se obter regiões que satisfaçam o critério H (mesmo nível de cinza, mesma média etc...).

Esse método é muitas vezes associado a uma estrutura *QUADTREE* [CA84] que permite, de uma forma elegante, decompor e agrupar partes de uma imagem. Construir um quadtree consiste em dividir uma imagem $I(x, y)$ em quatro regiões iguais e continuar a dividir cada região da mesma maneira até encontrar regiões satisfazendo a um critério H (figura 4.3). O quadtree é uma estrutura arborescente constituída, a): de uma raiz, que corresponde à imagem original, b): de nós, onde cada nó representa um quadrante da imagem e pode ter quatro ou zero descendentes, os filhos.

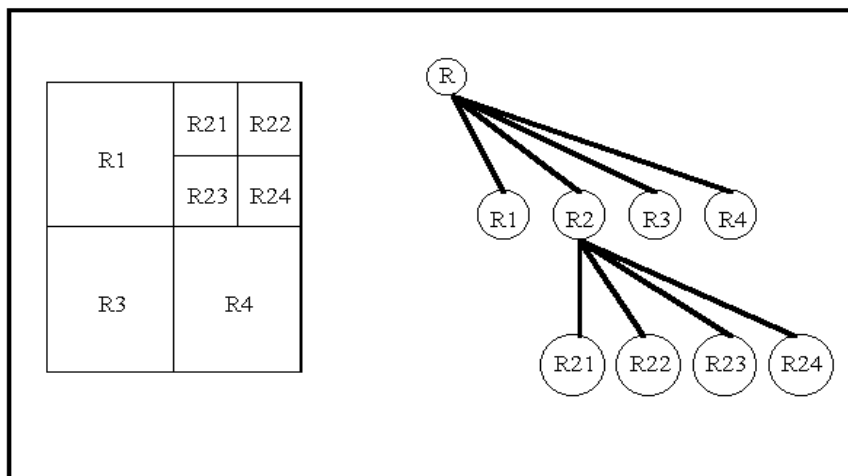


figura 4.3: Segmentação por divisão

O princípio, a partir de uma partição R da imagem, é ilustrado na figura 4.4.

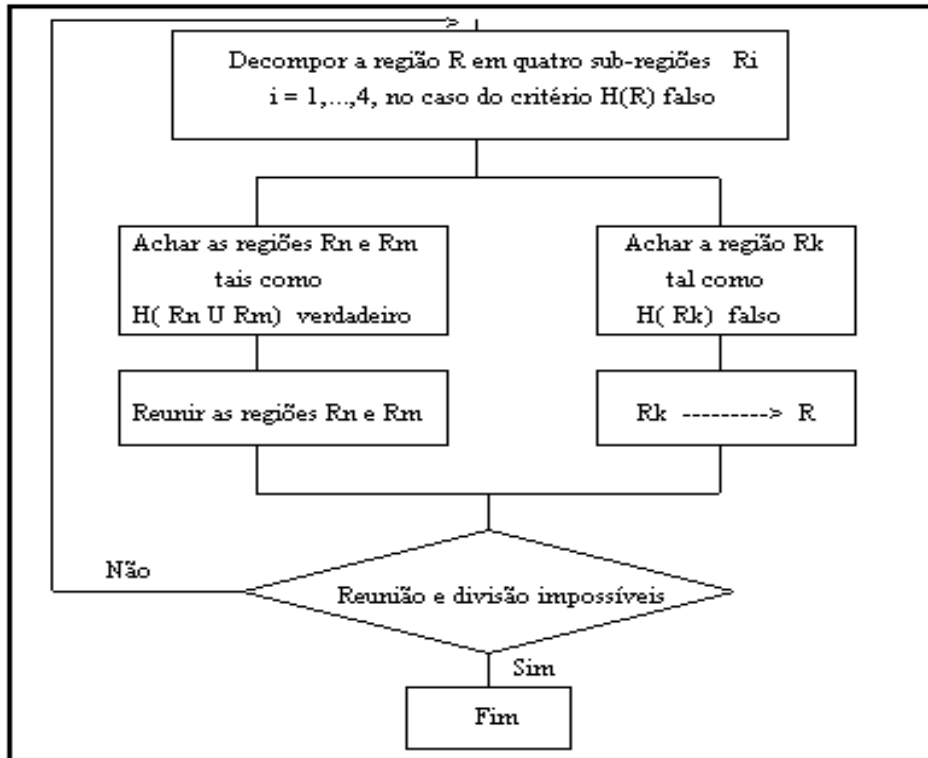


figura 4.4: Processo de divisão e fusão por Quadtree

4.3 SEGMENTAÇÃO POR TEXTURA

A segmentação por textura deve permitir a colocação em evidência dos motivos da imagem considerando-se as propriedades de regularidade e de repetição. A noção de textura supõe que:

- o motivo seja um agrupamento aleatório ou não aleatório de subconjuntos da imagem;
- as entidades detectadas sejam uniformes e tenham aproximadamente as mesmas dimensões em qualquer lugar da imagem
- o motivo local seja repetido numa região comparativamente larga em relação ao tamanho desse motivo.

Na realidade, apesar dessa definição parecer razoável e clara, a noção própria de textura fica mal definida. A figura 4.5 exemplifica vários exemplos de textura. Nessa figura, é possível averiguar o quanto a noção de textura é complexa. Alguns pesquisadores consideram que ela pode revestir-se de um aspecto anárquico e homogêneo (exemplos da erva, da casca de árvore etc...) enquanto outros consideram que ela se reveste mais de um aspecto determinístico e estruturado (exemplos do tecido, do muro etc...).

Os três métodos principais para segmentar uma imagem pela textura são: estatísticos, estruturais e espectrais. As técnicas estatísticas permitem caracterizar

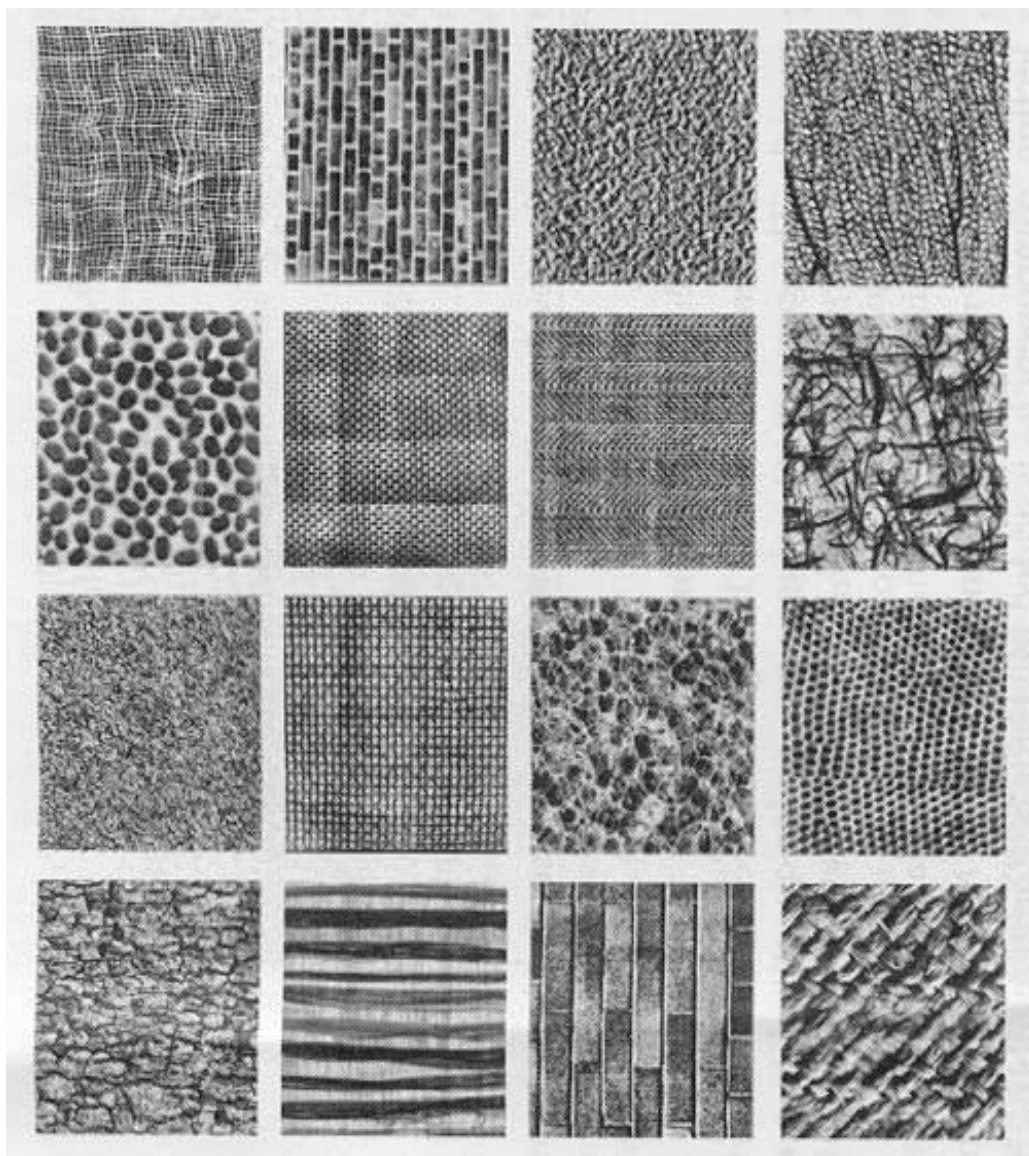


figura 4.5: Exemplo de texturas

uma textura de tipo suave, irregular, granulada etc... As técnicas estruturais tratam da disposição das primitivas numa imagem, por exemplo, textura baseada sobre linhas paralelas dispostas regularmente. As técnicas espectrais levam às propriedades do espectro de Fourier e permitem detectar a periodicidade global, identificando altas energias). A figura 4.6 exemplifica uma mistura de texturas de tipo suave, irregular, granulada, paralela e periódica.

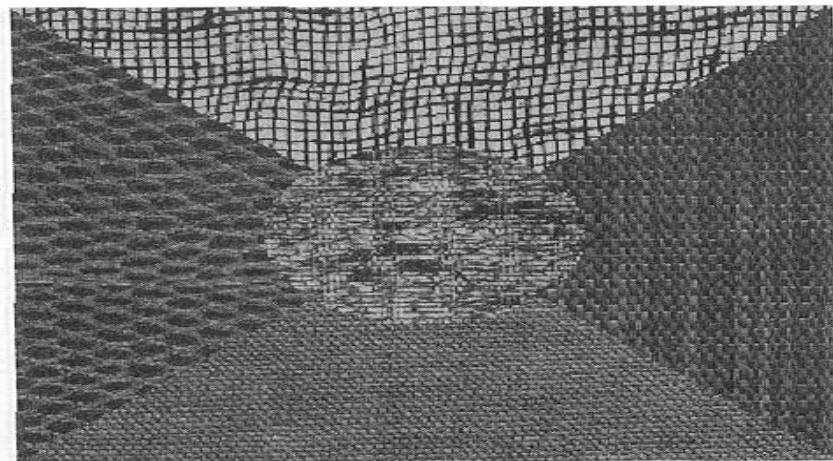


figura 4.6: Mistura de texturas de tipo suave, irregular, granulada, paralela, periódica

4.3.1 Técnica Estatística

Um bom método é considerar a distribuição das intensidades numa imagem e as posições dos pixels com níveis de cinza iguais ou parecidos.

Considere-se um critério de posição P e uma matriz $A(i, j)$ cujo elemento a_{ij} representa o número de vezes que pixels com nível de cinza z_i respeitam o critério P com pixels de nível de cinza z_j . Para facilitar a compreensão desses critérios, considere-se o exemplo de uma imagem com três níveis de cinza, $z_0 = 0$, $z_1 = 1$ e $z_2 = 2$, como segue:

$$\begin{vmatrix} 0 & 0 & 0 & 1 & 2 \\ 1 & 1 & 0 & 1 & 1 \\ 2 & 2 & 1 & 0 & 0 \\ 1 & 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{vmatrix}$$

Se o critério de posição P é “um pixel à direita e um pixel em baixo”, a matriz $A(i, j)$ é então:

$$A(3, 3) = \begin{vmatrix} 4 & 2 & 1 \\ 2 & 3 & 2 \\ 0 & 2 & 0 \end{vmatrix} \quad (4.1)$$

Seja n_p o número total de pares de pixels da imagem que satisfazem ao critério P . Definindo-se $C(i, j)$ como a matriz formada a partir dos elementos de $A(i, j)$ divididos por n_p , então $C(i, j)$ é uma avaliação da probabilidade que o par de pixels satisfazendo o critério P , tenha os valores (z_i, z_j) . Essa matriz chama-se matriz de co-ocorrência de nível de cinza. Como $C(i, j)$ depende de P , é possível detectar uma textura dada, escolhendo um bom critério P . De modo geral, o problema consiste

em analisar uma dada matriz $C(i, j)$ para classificar a textura da região sobre a qual $C(i, j)$ é computada. Haralick propõe um conjunto de descritores:

- o máximo de probabilidade

$$\max(c_{ij}) \quad (4.2)$$

- o momento de diferença de ordem k

$$\sum_i \sum_j (i - j)^k c_{ij} \quad (4.3)$$

- o momento inverso de diferença de ordem k

$$\sum_i \sum_j \frac{c_{ij}}{(i - j)^k} \quad (4.4)$$

- a entropia

$$\sum_i \sum_j c_{ij} \lg(c_{ij}) \quad (4.5)$$

- a uniformidade

$$\sum_i \sum_j c_{ij}^2 \quad (4.6)$$

Estes descritores permitem caracterizar o “conteúdo” da matriz $C(i, j)$. O primeiro descritor dá uma indicação da melhor resposta através de P , o segundo tem um valor que diminui quando os altos valores de $C(i, j)$ estão próximos da diagonal principal (as diferenças $(i - j)$ diminuem), o terceiro tem o efeito contrário ao segundo, o quarto dá uma medida da estrutura aleatória da textura, e o quinto é o menor quando os c_{ij} são todos iguais.

4.3.2 Técnica Estrutural

A idéia é construir, a partir de uma primitiva de textura simples, texturas mais complexas, usando regras que permitem limitar o número de arranjos possíveis da(s) primitiva(s).

Uma regra pode ser definida como:

$$S \Rightarrow aS \quad (4.7)$$

O símbolo S pode ser escrito, então, aS . Três aplicações dessa regra fornecem $aaaS$.

Considere o exemplo onde a representa um quadrado e a regra aS representa o conceito “*quadrados à direita de*”. Esta regra permite gerar a textura da figura 4.7.(b).

Extensões desta regra podem ser do tipo:

$$S \Rightarrow bA \quad (4.8)$$

$$A \Rightarrow cA \quad (4.9)$$

$$A \Rightarrow c \quad (4.10)$$

$$S \Rightarrow a \quad (4.11)$$

onde, b significa “*quadrado em baixo*”, e c “*quadrado à esquerda*”. Pode-se criar a textura da figura 4.7.(c) a partir da estrutura gerada pelas regras acima citadas.

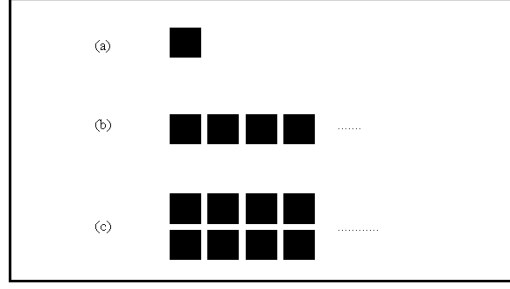


figura 4.7: Construção de texturas: (a) Primitiva de textura, (b) Regra $S \rightarrow aS$, (c) Uso de várias regras

4.3.3 Técnica Espectral

Uma ferramenta muito utilizada é o espectro de Fourier, que permite detectar a direcionalidade de motivos periódicos numa imagem. Considera-se as características do espectro de Fourier:

- picos proeminentes no espectro fornecem a direção principal dos motivos de textura
- a localização desses picos no plano das frequências fornece o período fundamental espacial dos motivos
- eliminando os componentes periódicos pela filtragem, ficam os componentes não periódicos que podem ser analisados pelas técnicas estatísticas.

A detecção e a interpretação das características do espectro são simplificadas quando o espectro $S(r, \vartheta)$ pode ser expresso a partir das coordenadas polares r e ϑ , variáveis neste sistema de coordenadas. Em cada direção ϑ , $S(r, \vartheta)$ pode ser considerada como uma função unidimensional $S(r)$. De uma forma equivalente, para cada frequência r , $S(\vartheta)$ é uma função unidimensional. A análise de $S(r)$ para um valor fixo de ϑ fornece o comportamento do espectro, a presença dos picos segundo a direção radial a partir da origem, e a análise de $S(\vartheta)$ para um valor fixo de r fornece o comportamento segundo o círculo centrado na origem.

A partir dos descritores $S(r)$ e $S(\vartheta)$ pode-se determinar outras primitivas:

$$\sum_{\vartheta=0}^n S_{\vartheta}(r) \quad (4.12)$$

$$\sum_{r=1}^R S_r(\vartheta) \quad (4.13)$$

onde R é o raio do círculo centrado na origem (para um espectro $N \times N$, R é escolhido igual $N/2$).

$S(r)$ e $S(\vartheta)$ constituem o par de valores para cada par de coordenadas (r, ϑ) . Variando essas coordenadas, pode-se gerar duas funções unidimensionais $S(r)$ e $S(\vartheta)$ que caracterizam a energia espectral da textura.

4.4 SEGMENTAÇÃO POR CONTORNO

Um contorno é uma mudança brusca do nível de cinza entre duas regiões relativamente homogêneas. Ele pode aparecer como uma sequência de pontos, uma linha, um segmento, uma curva ou uma forte variação do nível de cinza médio.

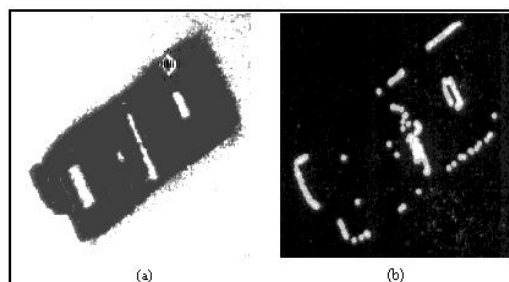


figura 4.8: Segmentação pelo operador de Moravec: (a) imagem original, (b) imagem binarizada dos pontos de Moravec

4.4.1 Detecção de Pontos

A forma mais resumida de segmentar uma imagem sob a forma de contornos é de segmentá-la na forma de pontos particulares relevantes. Isto supõe muitas vezes eliminar de forma eficiente o ruído para evitar detecções falsas. Esses pontos particulares relevantes pesquisados são pontos isolados que devem conter muita informação e não devem ser deteriorados por um ruído. Essa categoria de pontos é muito usada no processo de reconhecimento de objetos e na análise do movimento.

Um exemplo de segmentação de uma imagem sob a forma de pontos particulares é usar o operador de Moravec [Mor87b] (figura 4.8). Ele representa um algoritmo interessante de detecção de pontos que traduzem zonas locais particulares ou que podem constituir o centro de regiões de grande variância segundo algumas direções. O operador descrito por Moravec permite extrair pontos de alta variância segundo várias direções. Segundo esse operador, um ponto é altamente caracterizável se e somente se ele possuir uma variância alta segundo quatro direções: vertical, horizontal e as duas diagonais principais.

A partir de uma janela (normalmente 5x5 pixels), o operador avalia as variâncias segundo as quatro direções computando os quadrados das diferenças dos valores dos níveis de cinza. O centro da janela é então substituído pelo valor mínimo das quatro variâncias. Escolher esse valor mínimo como novo ponto da janela evita levar em consideração pontos aberrantes com níveis de cinza ruidosos.

4.4.2 Detecção de Bordas

Uma das técnicas de detecção de bordas mais usada consiste no processamento de uma imagem a partir de um operador de derivada local. Foram apresentados no capítulo precedente o gradiente $G(x,y)$ e o laplaciano $L(x,y)$, operadores de derivada de primeira e segunda ordem, para realçar o contraste. Esses operadores são também muito usados na pesquisa de contornos.

A avaliação do gradiente através de máscaras não é única, mas depende dos coeficientes e do tamanho das máscaras. A escolha das máscaras adequadas depende de vários critérios (quantidade de ruído, tipo de ruído, contraste, etc..) cujo principal e mais significativo é o tamanho da máscara. Este define a importância da suavização nos cálculos das diferenças (princípio da derivada local). Quanto maior a máscara, maior será a suavização, maior será a eliminação do ruído e, infelizmente, dos detalhes finos e mais demorado será o processo (figura 4.9).

A imagem gradiente não fornece diretamente os contornos. Em prática, ela fornece traços grossos e irregulares. Para obter traços mais precisos e mais nítidos, usa-se o processo de limiarização da imagem do módulo do gradiente. A precisão dos traços depende do valor da limiarização e da exatidão dos cálculos das derivadas: um

valor alto usado na limiarização fornece traços finos mas muitas vezes interrompidos, enquanto um valor fraco proporciona traços mais completos porém mais grossos e produz linhas “parasitas”.

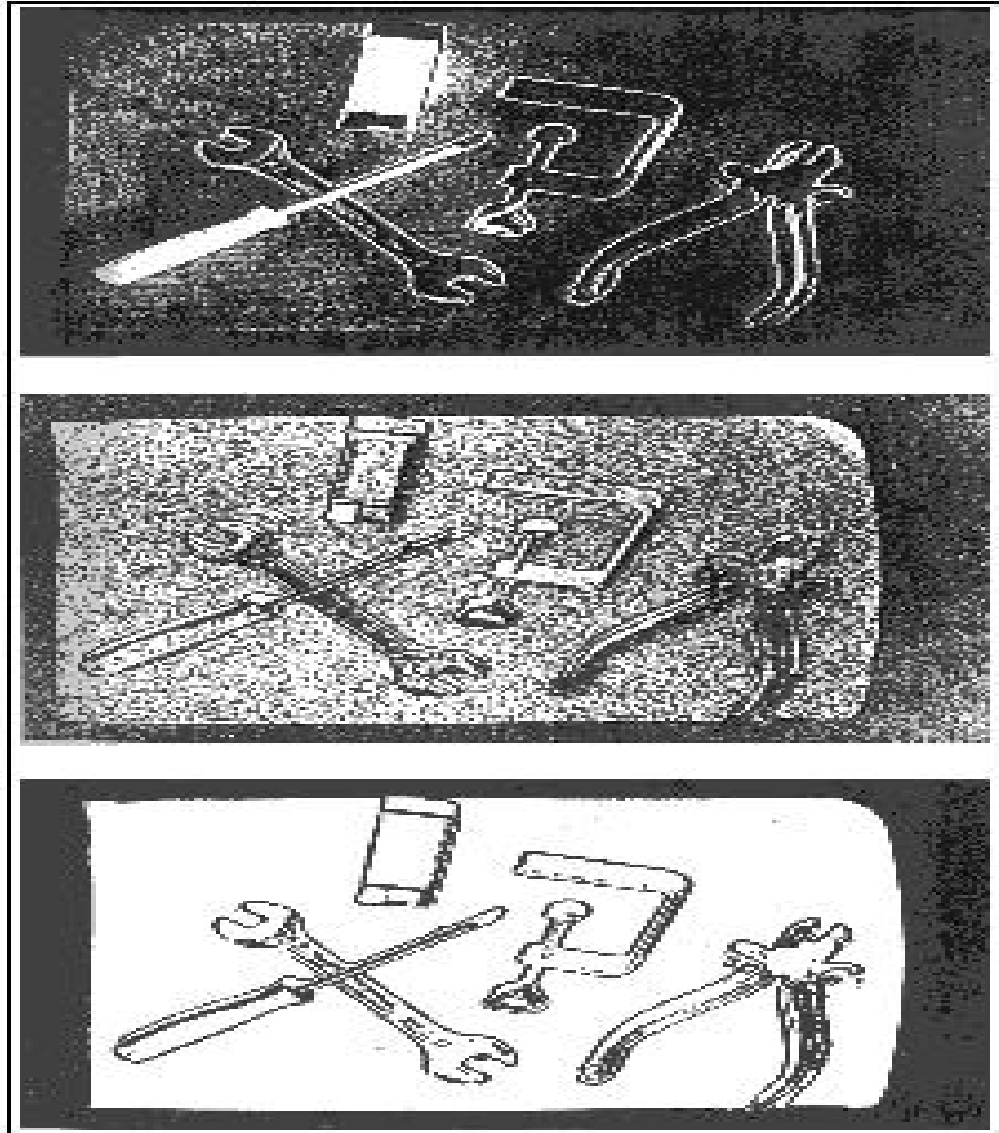


figura 4.9: Qualidade do Gradiente: (a) imagem original, (b) módulo do gradiente 3x3, (c) módulo do gradiente 7x7

4.5 ENLACE DE BORDAS E DETECÇÃO DE CONTORNOS

As técnicas precedentes permitem detectar descontinuidades nos níveis de cinza e fornecer contornos entre as diferentes regiões. Mas raramente o conjunto de pixels traduz contornos completos. Por causa do ruído, dos cortes no contorno devido à iluminação não uniforme e de outras perturbações, o contorno aparece com sendo um conjunto descontínuo.

A partir desses fatos, alguns procedimentos de enlace de bordas são usados junto aos algoritmos de detecção de contornos para reunir pixels de bordas em um conjunto significativo de contornos de regiões.

4.5.1 Perseguição de Contornos

Uma técnica apropriada para enlaçar pixels de um contorno é dispor de uma ferramenta que permite analisar as características dos pixels que tem a possibilidade de pertencer ao contorno e escolher os mais relevantes [Pug83]. Um algoritmo eficiente de perseguição de contornos consiste em usar a codificação de Freeman [Fre61]. A vantagem dessa codificação é permitir representar simplesmente uma seqüência de pontos a partir de oito direções, numeradas de 0 até 7, que são os eixos verticais, horizontais e as duas diagonais.

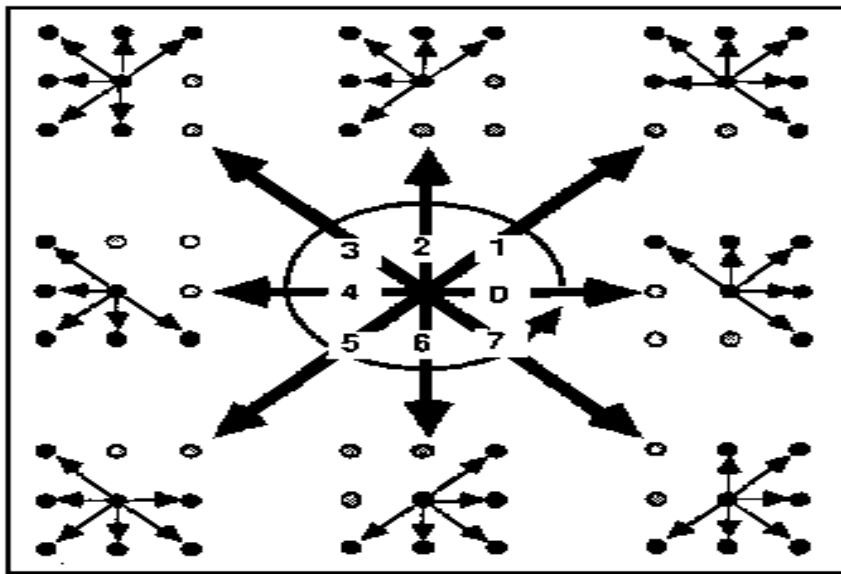


figura 4.10: Pesquisa do contorno a partir dos códigos de Freeman

O princípio do algoritmo de perseguição de contornos reside na busca iterativa de pontos similares do contorno, cada um desses pontos respeitando um critério de pertinência ao contorno e escolhido da vizinhança do ponto anterior já retido. O grande interesse desse algoritmo vem do fato que a busca é otimizada por pesquisa orientada segundo direções preferenciais (figura 4.10).

O critério de pertinência depende exclusivamente do tipo de imagem processada, mas, de forma geral, está ligada à noção de nível de cinza. Ele pode ser, por exemplo, o valor de gradiente caso a imagem tiver sido derivada por um operador de diferenciação do tipo Gradiente, ou ainda simplesmente um dos dois níveis de cinza (preto ou branco) caso a imagem seja binária.

Escolhendo o caso de imagens binárias (dois níveis de cinza) onde a pesquisa do contorno é de um objeto preto (nível de cinza 0) sobre o fundo branco (nível de cinza 1), o algoritmo de perseguição de contorno, detalhado na figura 4.10, mostra que a busca de um ponto preto P_{n+1} do contorno é função do ponto preto anterior P_n do contorno já definido. Em realidade, o ponto anterior P_n participa na busca de P_{n+1} através do valor do vetor de Freeman F_n segundo o qual P_n foi detectado. Essa característica do processo permite limitar o leque de possíveis buscas e então otimizar a detecção de P_{n+1} . Por exemplo, caso o ponto P_n foi descoberto segundo a direção de Freeman $F_n = 6$, então é inútil voltar a procurar o próximo ponto

P_{n+1} na direção oposta, no nosso caso $F_{n+1} = 2$. Um exemplo das possíveis futuras buscas é procurar em uma das cinco direções $F_{n+1} = 5, 6, 7, 0$ ou 1 . Na prática, a flexibilidade do algoritmo permite modificar e adaptar a definição das direções preferenciais em função da imagem tratada (imagem médica, industrial, etc..).

4.5.2 Transformada de Hough

Um problema freqüentemente encontrado na extração de primitivas é a detecção de curvas analíticas do tipo segmento de reta, círculos, elipses, etc... Essas curvas C constituem uma família F representada da forma seguinte:

$$f(X, \alpha) = 0 \quad (4.14)$$

onde X é vetor representando um ponto $M(x, y)$ do espaço e α representando um vetor de parâmetros característicos de C .

Dados um número de pontos discretos $Mi(xi, yi)$, o objetivo é determinar se existe um sub-conjunto de pontos Mi que pertencem a uma curva C específica com o vetor parâmetro α a definir. Um método brutal consiste em testar todos os pontos $Mi(xi, yi)$ dois a dois, o que conduz a um número exagerado de iterações. Hough e Rosenfeld [Bal81] propuseram um método para os pontos do plano $(0, X, Y)$ que consiste em calcular para cada ponto $Mi(xi, yi)$ do contorno o conjunto dos parâmetros α que verificam $f(xi, yi, \alpha) = 0$. Isso consiste em avaliar, no espaço dos parâmetros α , a curva apresentada pelos parâmetros de todas as curvas $f(xi, yi, *)$ que passam por $Mi(xi, yi)$.

O método da transformada de Hough representa um método robusto e genérico de detecção de curvas. A transformada de Hough é aplicável quando se possui informações precisas acerca da forma da curva mas não depende de informações precisas sobre a localização da curva. Os dados de base da transformada de Hough são geralmente pontos de uma imagem obtidos através das transformações de gradiente. **A idéia da transformada de Hough é aplicar na imagem uma transformação tal que todos os pontos pertencentes a uma mesma curva sejam mapeados num único ponto de um novo espaço de parametrização da curva procurada.** A figura 4.11 ilustra as diferentes etapas do método no caso da detecção de círculos. A transformada de Hough é um método de acumulação de requisitos muito geral. Ela permite detectar qualquer curva, mesmo pouco visível ou fortemente ruidosa.

A seguir serão apresentados a transformada de Hough aplicada a detecção de retas e de círculos assim que os conceitos envolvidos nessa abordagem como tabela de acumulação, novo espaço de parametrização, entre outros.

4.5.2.1 Detecção de retas pela Transformada de Hough

Considere o exemplo de um conjunto de pontos $P(x, y)$ no qual tente-se detectar segmentos de reta. A reta pode ser parametricamente definida em coordenadas cartesianas como em coordenadas polares. No caso da transformada de Hough, é mais judicioso optar pela segunda formulação. Uma equação da reta definida em coordenadas polares é dada por:

$$\rho = x \cos \vartheta + y \sin \vartheta \quad (4.15)$$

No plano da imagem, qualquer ponto $P(x, y)$ está situado na reta $R(\rho, \vartheta)$ cujos parâmetros ρ e ϑ são constantes. No plano de Hough (ou ainda espaço de

parametrização), cada ponto $P(x, y)$, no espaço da transformada de Hough, gera uma senóide $S(x, y)$. Portanto, no plano de Hough, qualquer ponto da transformada de Hough pertence a uma senóide $S(x, y)$ cujos parâmetros x e y são constantes.

A propriedade mais importante da transformada de Hough é que todos os pontos $P(x, y)$ de uma mesma curva geram no espaço de Hough senóides $S(x, y)$ que se interseccionam num único ponto. Ou ainda que as curvas do plano de Hough que correspondem aos pontos de uma mesma curva passam por um único ponto.

No caso de detecção de retas, não poderia ser diferente, e as curvas do plano de Hough que correspondem aos pontos de uma mesma reta $R(\rho, \vartheta)$ passam também por um único ponto. Na prática pelo fato de usar imagens digitais, a solução opera a partir de uma discretização do espaço de parametrização. A detecção de retas pode ser decomposta da seguinte maneira:

- Para cada ponto $P(x, y)$ são definidas todas as retas de parâmetros ρ e ϑ regidas pela equação 4.15 que passam por $P(x, y)$;
- O espaço de parametrização, definido aqui pelos parâmetros ρ e ϑ , é então discretizado “em uma **tabela de acumulação**” $A(\rho, \vartheta)$, inicializado em zero, de duas dimensões, uma para cada parâmetro;
- Cada elemento da tabela $A(\rho, \vartheta)$ corresponde a um intervalo de ρ e de ϑ .
- De todos os parâmetros presentes na equação 4.15, são conhecidos x e y fornecidos por cada ponto $P(x, y)$;
- O valor do parâmetro ϑ é obtido pela informação do gradiente $G(x, y)$, e mais precisamente pelo argumento do gradiente $G(x, y)$ ao ponto $P(x, y)$;
- Para cada ponto $P(x, y)$, a partir da equação 4.15, é computado o valor de ρ ;
- O elemento $A(\rho, \vartheta)$ da tabela de acumulação é então incrementado;
- Após a determinação dos parâmetros de todos os pontos, a busca de valores fortes da tabela de acumulação permite obter as coordenadas polares das retas mais significantes presentes na imagem.

No caso de reta, o método acima descrito pode ser melhorado, utilizando a equação $\rho = \frac{(xG_x + yG_y)}{\sqrt{G_x^2 + G_y^2}}$ que evita computar os valores de ϑ , $\cos \vartheta$ e $\sin \vartheta$.

4.5.2.2 Detecção de círculos pela Transformada de Hough

O desenvolvimento do método de detecção de círculos pela Transformada de Hough é genérico e pode ser aplicado para outras curvas de parametrização conhecida.

Uma parametrização possível de um círculo em coordenadas cartesianas é: $(x - x_c)^2 + (y - y_c)^2 = R^2$, onde (x_c, y_c) e R representam respectivamente as coordenadas do centro e o raio do círculo.

Para iniciar a busca do novo espaço de parametrização de Hough, sabe-se que a informação primordial é o gradiente. Para introduzir essa tal informação, precisa-se derivar na primeira ordem, respectivamente a x , a equação do círculo.

A derivada primeira da equação do círculo é:

$$(x - x_c) + (y - y_c) \frac{\partial y}{\partial x} = 0 \quad (4.16)$$

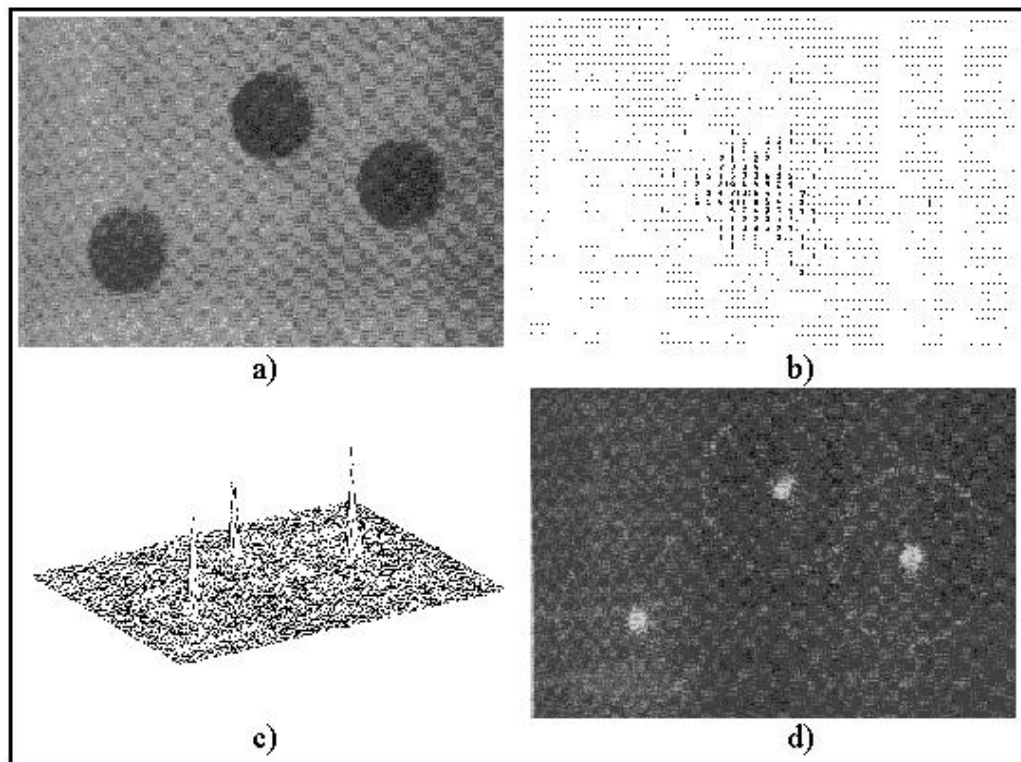


figura 4.11: Exemplo de detecção de círculos pela Transformada de Hough: (a) imagem original em níveis de cinza, (b) tabela de acumulação (c) visualização tridimensional da tabela de acumulação (d) visualização bidimensional dos centros dos discos

Sabe-se que $\frac{\partial y}{\partial x}$ representa na realidade a tangente da derivada no ponto (x, y) . Sendo ϕ o ângulo dessa tangente, tem-se que

$$\frac{\partial y}{\partial x} = \tan \phi$$

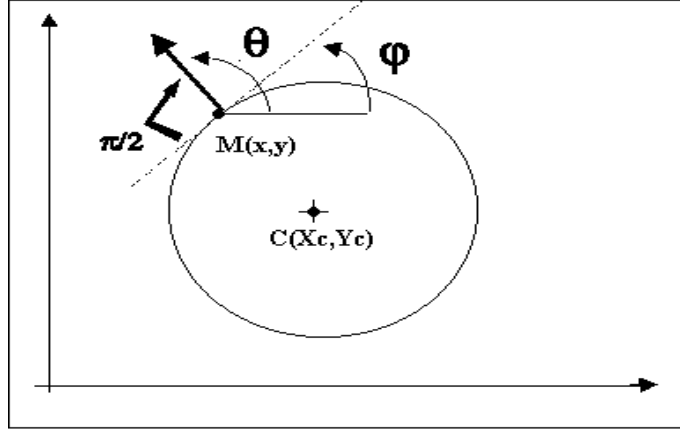


figura 4.12: Relação entre a direção do gradiente e a direção da tangente

Para introduzir a informação do gradiente, sabe-se que a direção do gradiente de ângulo θ é perpendicular à curva, portanto perpendicular à direção da tangente (figura 4.12), seja:

$$\theta = \phi + \frac{\pi}{2}$$

Associando essas duas últimas equações com a equação 4.16, tem-se que:

$$(x - x_c) \tan \phi = (y - y_c)$$

Tem-se afinal um sistema de duas equações, a primeira proveniente do último cálculo e a segunda da parametrização do círculo:

$$\begin{aligned} (x - x_c) \tan \phi &= (y - y_c) \\ (x - x_c)^2 + (y - y_c)^2 &= R^2 \end{aligned}$$

A resolução desse sistema fornece como solução final:

$$\begin{aligned} x_c &= x - R \cos \phi \\ y_c &= y - R \sin \phi \end{aligned} \quad (4.17)$$

Dessa solução, é possível concluir que o novo espaço de parametrização de Hough é o espaço formado pelas abscissas e ordenadas dos círculos. Isto é possível somente se considerar o raio do círculo como sendo um dado conhecido. As figuras 4.11 e 4.13 ilustram respectivamente as diferentes etapas do método no caso da detecção de círculos e a detecção de círculos numa imagem complexa.

4.5.2.3 Conclusão sobre a Transformada de Hough

A transformada de Hough constitui um método muito genérico que pode ser generalizado para qualquer tipo de curva. As vantagens dessa abordagem são:

- Cada ponto pode ser tratado independentemente, o que permite processamento paralelo e aplicações em tempo real;

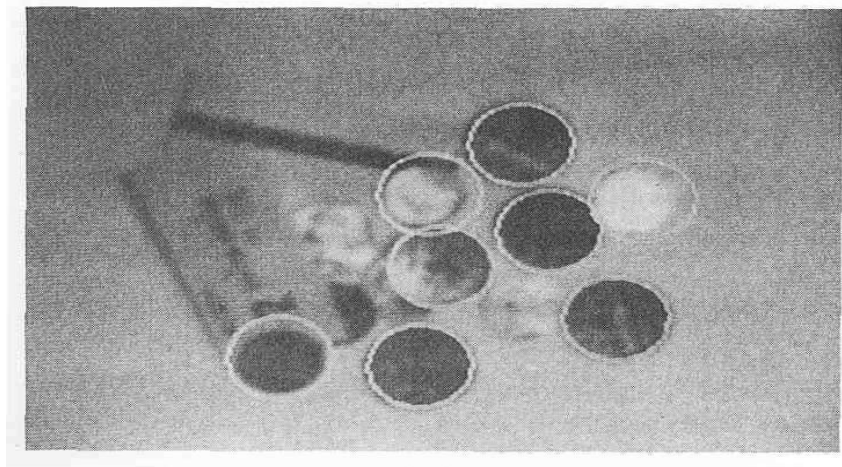


figura 4.13: Exemplo de detecção de círculos pela Transformada de Hough numa imagem complexa.

- A transformada de Hough permite detectar qualquer curva, mesmo fortemente ruidosa;
- A transformada de Hough é um dos poucos métodos que permite a detecção de curvas parcialmente oclusas e/ou ligeiramente deformadas;
- A transformada de Hough permite evidenciar várias curvas de um mesmo tipo numa mesma imagem.

A maior desvantagem da transformada de Hough reside no esforço computacional em termos de cálculo e de armazenamento. Quando mais complexa for a curva, maior será o número de parâmetros e também o esforço computacional. Uma outra desvantagem é a precisão da discretização dos parâmetros. O risco, no caso de um passo de discretização muito fraco, é não detectar as curvas. Ao contrário, o passo de discretização muito alto cria erros de interpretação do acumulador ao nível da precisão do lugar das curvas bem como ao nível do número das retas.

4.5.3 Técnicas do grafo

Os métodos precedentes levam em consideração a operação de gradiente que, pela natureza de derivação, podem fornecer resultados bons exclusivamente no caso de imagens pouco ruidosas. Ver-se-á agora o método global baseado na representação de segmentos de reta na forma de uma estrutura de grafo e na pesquisa do menor custo que corresponde a segmentos significativos. Esse procedimento é muito mais complicado mas tem a vantagem de fornecer resultados satisfatórios em presença de ruído.

Um grafo $G = (N, A)$ é um conjunto finito de nós N associado ao conjunto A de pares não ordenados de elementos distintos de N . Cada par (n_i, n_j) de A chama-se “*arco*”. Um grafo onde os arcos são direcionados chama-se “*grafo direcionado*”. Se um arco é direcionado do nó n_i para o nó n_j , o nó n_j chama-se de “*sucedor*” do nó pai n_i . O processo de identificar os sucessores de um nó chama-se de “*expansão*” do nó.

Supondo-se um conjunto de pontos, o processo de procurar contornos a partir de um grafo composto dos nós $n_1, \dots, n_i, \dots, n_j, \dots, n_k$ consiste em encontrar o melhor caminho entre os dois nós n_1 e n_k associando uma função de custo $c(n_i, n_j)$ a cada

arco (n_i, n_j) . A seqüência dos nós $n_1, \dots, n_{i-1}, n_i, \dots, n_k$, onde n_i é o sucessor de n_{i-1} chama-se trajetória de n_1 até n_k e o custo é dado pela equação:

$$\sum_{i=2}^k c(n_{i-1}, n_i) \quad (4.18)$$

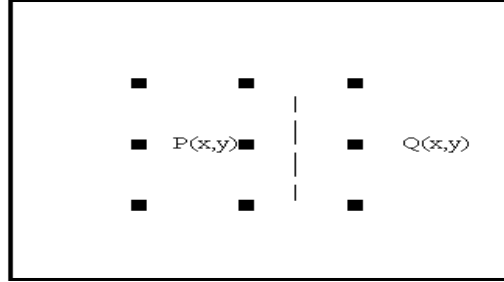


figura 4.14: Elemento de borda entre os pixels $P(x,y)$ e $Q(x,y)$

Um elemento da borda é definido como a borda entre dois pixels $P(x,y)$ e $Q(x,y)$ tal que $P(x,y)$ e $Q(x,y)$ são 4-vizinhos (figura 4.14). O contorno é uma seqüência de elementos de bordas.

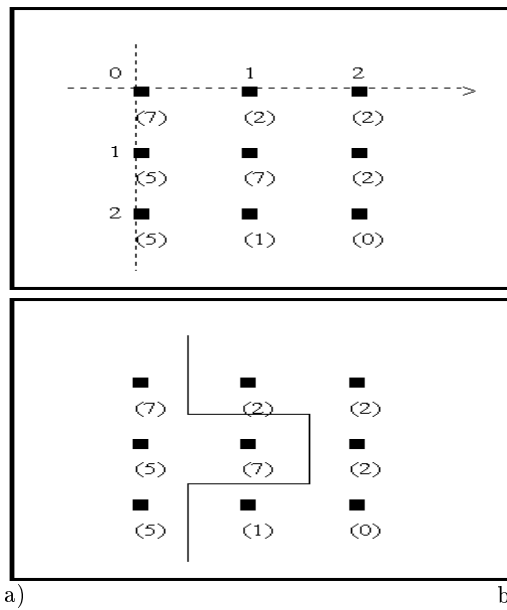


figura 4.15: Exemplo de detecção de contorno por grafo: a) imagem original b) contorno correspondendo ao caminho de menor custo

No exemplo da (figura 4.15), os números entre parênteses são o nível de cinza, e os outros números representam as coordenadas do pixel.

A função custo é definida:

$$c(P, Q) = H - [I(P) - I(Q)] \quad (4.19)$$

onde H é o maior valor de nível de cinza (7 nesse exemplo), $I(R)$ o nível de cinza de um ponto R .

No grafo obtido (figura 4.16), cada nó corresponde a um elemento do contorno e um arco existe entre dois nós se os dois elementos correspondentes podem fazer parte de um contorno. O custo de cada elemento é mostrado em cima de cada arco.

Com o primeiro elemento o par $[(0, 0), (0, 1)]$ ou $[(0, 1), (0, 2)]$ e o último $[(2, 0), (2, 1)]$ ou $[(2, 1), (2, 2)]$, a trajetória de mínimo custo é apresentada pela linha pontilhada.

O contorno correspondente é mostrado na figura 4.15-b).

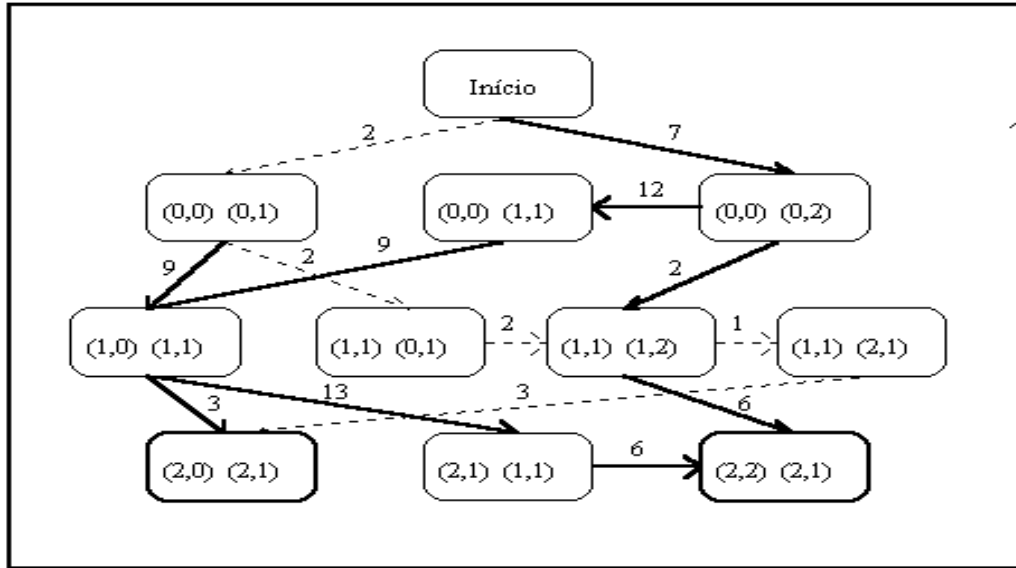


figura 4.16: Grafo usado para a detecção de contorno da imagem da figura 4.15

Capítulo 5

MORFOLOGIA MATEMÁTICA

5.1 CONSIDERAÇÕES GERAIS

Como em biologia onde a morfologia refere-se ao estudo da estrutura dos animais e das plantas, a morfologia matemática, elaborada inicialmente por Georges Matheron e Jean Serra, concentra seu esforço no estudo da estrutura geométrica das entidades presentes numa imagem.

O princípio básico da morfologia matemática consiste em extrair uma informação relativa à geometria e à topologia de um conjunto desconhecido de uma imagem pela transformação a partir de um outro conjunto completamente definido chamado *elemento estruturante*. A partir do seu formato e tamanho, tem-se a possibilidade de testar e de quantificar de que maneira o elemento estruturante “*está ou não está contido*” na imagem. Na figura 5.1, tem-se uma imagem binária e um exemplo de elemento estruturante. Pode-se verificar que, em função de sua posição, esse elemento estruturante inclui-se ou não no conjunto desconhecido da imagem. Marcando os resultados das posições onde o elemento estruturante inclui-se na imagem, tem-se uma primeira resposta sobre a estrutura geométrica da entidade dessa imagem. O tipo e a natureza da informação extraída depende necessariamente do tipo de elemento estruturante e do tipo de imagem estudada.

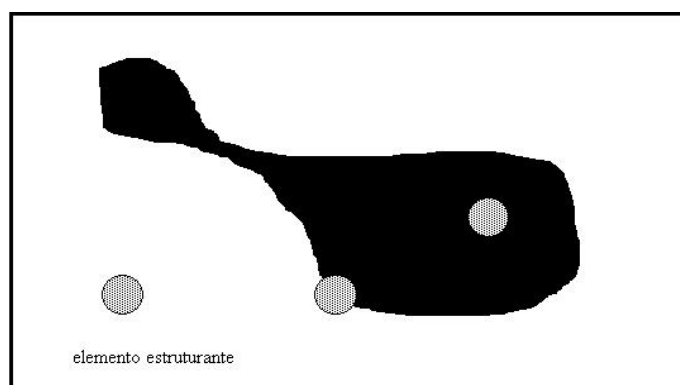


figura 5.1: Exemplos de inclusão do elemento estruturante numa imagem binária

De forma geral, existem dois tipos de morfologia matemática: a *morfologia binária* que se aplica sobre imagens binárias e a *morfologia em níveis de cinza* que

se aplica sobre imagens em níveis de cinza. Na morfologia binária, na vizinhança de cada pixel da imagem original, é procurada uma configuração de pontos pretos e brancos. Quando a configuração é encontrada, ao pixel correspondente da imagem resultante é dado o rótulo “verdadeiro”; senão, o pixel resultante recebe o rótulo “falso”. Uma operação morfológica binária é portanto completamente determinada a partir da vizinhança examinada ao redor do ponto central, da configuração de pontos pretos e brancos nessa vizinhança e do algoritmo.

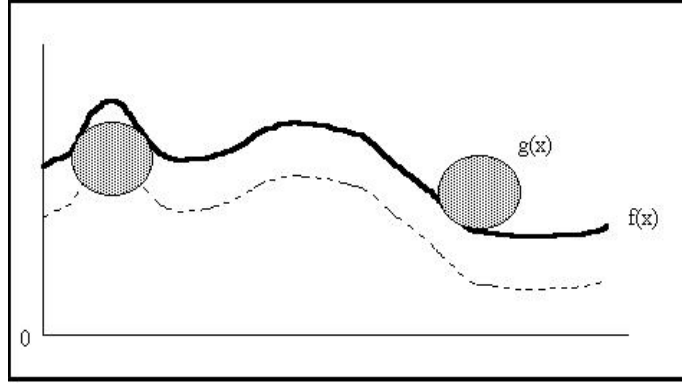


figura 5.2: Exemplo de interação entre o elemento estruturante e a imagem cinza

Na morfologia cinza (figura 5.2), na vizinhança de cada pixel ou numa parte da vizinhança da imagem original, é necessário conhecer o valor do pixel mais escuro MIN , o valor do pixel mais claro MAX . O valor do pixel resultante corresponde a uma combinação particular de MAX e MIN . O tamanho e a forma da vizinhança, as regiões de pesquisa de MIN e MAX e o algoritmo determinam completamente uma operação de morfologia cinza.

5.2 CONSIDERAÇÕES INICIAIS

5.2.1 Translação e Simetria

Seja um sub-conjunto B de um conjunto $\mathcal{E}$. Este sub-conjunto pode sofrer algumas modificações. Por exemplo, o conjunto obtido de B pela translação do vetor h , denotado $B + h$, é:

$$B + h = \{y \in \mathcal{E} \rightarrow y - h \in B\} \quad (5.1)$$

Este conjunto transladado, chamado B_h , pode ser também definido desta outra maneira:

$$y \in B \implies y + h \in B_h \quad (5.2)$$

com as seguintes propriedades:

$$\begin{aligned} B_o &= B \\ (B_h)_k &= B_{h+k} = (B_k)_h \end{aligned}$$

O conjunto deduzido de B por simetria central pela origem $\{o\}$ do sistema de referência, denotado $\tilde{B}$ e chamado de B transposto é:

$$\begin{aligned} y \in B &\implies -y \in \tilde{B} \\ B_h &= (\tilde{B})_{-h} \end{aligned} \quad (5.3)$$

5.2.2 Operações de Minkowski

Vimos anteriormente como podia ser definida a adição de um vetor com um conjunto (translação). É também possível definir operações entre conjuntos. Duas delas, muito interessantes, desenvolvidas por Minkowski [Min03] e Hadwiger [Had50] [Had57], são conhecidas como a adição e a subtração de Minkowski.

Sejam dois sub-conjuntos A e B de $\mathcal{E}$. A adição de Minkowski de A com B , denotada $A \oplus B$, é o seguinte conjunto:

$$A \oplus B = \{x \in \mathcal{E} : \exists a \in A \text{ e } \exists b \in B : x = a + b\} \quad (5.4)$$

Esta definição pode ser formulada de uma outra maneira [Ser82], a ser utilizada posteriormente:

$$A \oplus B = \cup_{b \in B} A_b \quad (5.5)$$

Da mesma maneira, a subtração de Minkowski do sub-conjunto A com o sub-conjunto B , denotada $A \ominus B$, é:

$$A \ominus B = \{x \in \mathcal{E} : \forall b \in B, \exists a \in A : x = a - b\} \quad (5.6)$$

Uma formulação desta definição [Ser82], a ser usada posteriormente, é:

$$A \ominus B = \cap_{b \in B} A_b \quad (5.7)$$

As propriedades destes operadores são amplamente detalhadas em [BB94], [Min03], [Had50] e [Had57].

5.3 PROPRIEDADES BÁSICAS

Toda operação morfológica compõe-se de uma transformação ψ de um conjunto em outro e de uma medida μ . $\psi(X)$ é um conjunto, $\mu(\psi(X))$ é um número. Esses operadores não são simplesmente matemáticos. Os seus efeitos não podem violar a realidade física que eles representam. ψ e μ devem respeitar as condições de contorno impostas pelas leis da percepção visual:

- a invariância à translação,
- a invariância à homotetia,
- o conhecimento local,
- a continuidade.

5.3.1 Propriedade de Invariância à Translação

Seja X_h o conjunto transladado do conjunto X pelo vetor h . De forma geral, existem dois tipos de tranformações, as que não dependem da posição do referencial e as que dependem desta posição.

Neste caso, pode-se afirmar que a transformação ψ sobre um conjunto X é invariante com a translação h (figura 5.3). Pode-se escrever:

$$\psi(X_h) = [\psi(X)]_h \quad (5.8)$$

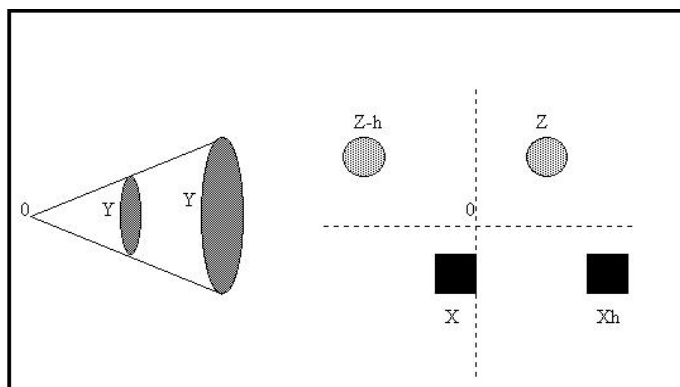


figura 5.3: Translação e homotetia

5.3.2 Propriedade de Invariância à Mudança de Escala

Seja o exemplo de dois pesquisadores trabalhando com a mesma amostra. Pode acontecer que esses pesquisadores trabalhem com uma ampliação diferente desta amostra. Para obter as mesmas conclusões, eles tem que achar um meio de trabalhar independentemente da ampliação usada. Para isto é necessário que a avaliação quantitativa da amostra X seja independente da redução ou da ampliação λ . Isso se traduz por:

$$\psi(\lambda X) = \lambda \psi(X) \quad (5.9)$$

ou ainda

$$\psi(X) = \lambda \psi(X/\lambda) \quad (5.10)$$

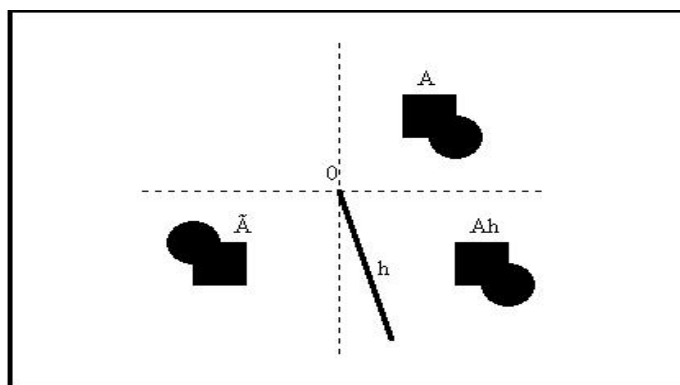


figura 5.4: Translação e simetria

A transformação ψ é então invariante por homotetia. Aplicar ψ a X é então igual a reduzir X , aplicar ψ e depois dilatar o resultado (figura 5.4).

5.3.3 Conhecimento Local

Na realidade, quando estuda-se um conjunto X , raramente tem-se a possibilidade de estudar o conjunto X como um todo, mas provavelmente somente uma parte desse conjunto que denomina-se de Z . Exemplos disto podem ser encontrados em imagens aéreas ou imagens microscópicas. Não se conhece o conjunto X mas em realidade $X \cap Z$ (e claro o conjunto Z). Nestas condições, as transformações ψ

usadas serão aplicadas sobre uma máscara Z' provendo do subconjunto Z onde é conhecido X . A máscara Z' depende somente de Z e $X \cap Z$. Isto representa o terceiro princípio da morfologia matemática:

A transformação ψ satisfaz a propriedade de conhecimento local se para cada conjunto restrito Z' no qual X é conhecido, pode-se encontrar um conjunto restrito Z no qual o conhecimento de X é suficiente para aplicar localmente (através de Z') a transformação ψ :

A $\forall Z'$ restrito, $\exists Z$ restrito tal como:

$$[\psi(X \cap Z)] \cap Z' = \psi(X) \cap Z' \quad (5.11)$$

5.3.4 Continuidade

A noção de borda é complicada, devido às limitações dos sistemas de análise de imagens. Os matemáticos criaram o conceito de borda a partir da definição dos conjuntos complementares X e X^c . Esta definição pode entrar em contradição com a dos físicos que consideram uma borda como o contato entre duas fases. Na realidade, quando estuda-se duas fases, a detecção de bordas realizada de outra maneira. Achar bordas Δ_i consiste em procurar conjuntos que são infinitamente pequenos, mas que não tendem ao conjunto vazio mas em um limite (que pode ter uma espessura nula). Para realizar isto, J. Serra modelou os Δ_i como sendo topologicamente conjuntos fechados que contém as suas bordas. Esta noção faz surgir o quarto princípio da morfologia matemática:

Para qualquer seqüência de conjuntos fechados que tende para um limite Δ e para cada transformação crescente ψ , deve corresponder uma seqüência de conjuntos transformados que tendem para a transformação do limite Δ .

5.4 OPERADORES MORFOLÓGICOS BINÁRIOS BÁSICOS

5.4.1 Aspecto digital

A partir de agora, B representa o elemento estruturante, noção introduzida anteriormente. Nas diferentes transformações, B_x representa o elemento estruturante B centrado no pixel x . Em função do contexto e por necessidade de simplificação, B_x poderá ser simplesmente notado B . Um elemento estruturante é definido pelos pixels que o formam (exemplo 5.1) e que são representados por "." e por "•". Um pixel marcado "." é um pixel inativo ou neutro, quer dizer que ele não interage com o conjunto X . O pixel ".", simplesmente, aparecerá no elemento estruturante para visualizar o seu aspecto geométrico. Um pixel marcado "•" significa um pixel ativo que tem um papel a desenvolver na interação com a imagem X .

Os pixels "•" do elemento estruturante criam um sub-conjunto que vai agir com a imagem X . O resultado dessa interação é colocado numa posição específica, a do *ponto central PC* do elemento estruturante, na imagem no momento da ação. O símbolo "()" representa este ponto central PC no elemento estruturante.

Na forma digital, B é representado entre "{ }" da seguinte forma:

$$B = \left\{ \begin{array}{c} \cdot \\ \cdot \cdot \cdot \\ \cdot \end{array} \right\}$$

Na maioria dos exemplos apresentados, o ponto central do elemento estruturante corresponderá a seu centro físico. Neste caso, num objetivo de simplificação, o símbolo "()" será omitido (exemplo 5.1).

Exemplo 5.1 : *Representação do elemento estruturante*

$$B = \left\{ \begin{array}{c} \bullet \\ \bullet \\ \bullet \end{array} \right\} = \left\{ \begin{array}{c} \bullet \\ \bullet \\ \bullet \end{array} \right\}$$

Muitas vezes será necessário introduzir o elemento estruturante transposto $\tilde{B}$:

$$\text{Se } B = \left\{ \begin{array}{c} \bullet \\ \bullet \\ \bullet \end{array} \right\} \text{ então } \tilde{B} = \left\{ \begin{array}{c} \bullet \\ \bullet \\ \bullet \end{array} \right\}$$

Da mesma maneira, por ser binária, a imagem digital X contém dois tipos de informação, o fundo (representado por ".") e os pixels relevantes (representados por "•"). Na forma digital, a imagem X é representada entre "[]" da seguinte forma:

Exemplo 5.2 : *Representação do conjunto X*

$$X = \left[\begin{array}{c} \bullet \\ \bullet \\ \bullet \\ \bullet \\ \bullet \end{array} \right]$$

5.4.2 Operadores elementares

Um primeiro conceito fundamental em Morfologia Matemática é o de relação de ordem parcial. Seja $\mathcal{E}$ um conjunto não vazio. A relação habitualmente usada no caso de sub-conjuntos de $\mathcal{E}$ é a de inclusão que permite comparar certos sub-conjuntos entre si. Seja $\mathcal{P}(\mathcal{E})$ a coleção de todos os sub-conjuntos de $\mathcal{E}$ associada com a relação de inclusão $\subset$. $\mathcal{P}(\mathcal{E})$ representa um conjunto parcialmente ordenado anotado $(\mathcal{P}(\mathcal{E}), \subset)$. [BB94] mostraram que o conjunto $(\mathcal{P}(\mathcal{E}), \subset)$, provido das operações de união e interseção estendidas às famílias em $\mathcal{P}(\mathcal{E})$, forma um *reticulado completo*. Sejam ψ um operador sobre $\mathcal{P}$ e X uma subcoleção de $\mathcal{P}$. [BB94] demonstraram que qualquer operador pode ser decomposto a partir de quatro classes fundamentais de operadores, chamados de *operadores elementares*, que são a erosão, a anti-erosão, a dilatação e a anti-dilatação.

Definição 5.1

Um operador ψ é uma:

- dilatação se e somente se, para todo $X \subset \mathcal{P}$, $\psi(\sup X) = \sup(\psi(X))$
- erosão se e somente se, para todo $X \subset \mathcal{P}$, $\psi(\inf X) = \inf(\psi(X))$
- anti-dilatação se e somente se, para todo $X \subset \mathcal{P}$, $\psi(\sup X) = \inf(\psi(X))$
- anti-erosão se e somente se, para todo $X \subset \mathcal{P}$, $\psi(\inf X) = \sup(\psi(X))$

5.4.3 Erosão binária

5.4.3.1 Primeira Definição

[Ser82] define a operação de erosão binária **ero** da seguinte maneira:

Definição 5.2

A erosão de um conjunto X pelo elemento estruturante B é:

$$\mathbf{ero}^B(X) = X \mathbf{ero} B = \{x \in \mathcal{E} : B_x \subset X\}$$

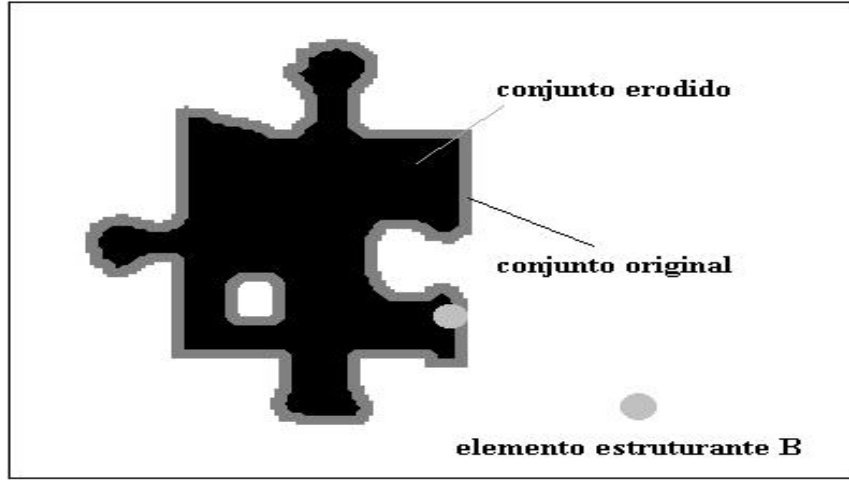


figura 5.5: Princípio da erosão binária

Segundo a definição 5.2, o elemento estruturante B deve deslizar na imagem X (figura 5.5). Por ser binária, X contém dois tipos de informação, o fundo (representado por ".") e os pixels relevantes (representados por "•"). O significado da definição 5.2 é que o elemento estruturante B_x , posicionado e centrado no pixel x de X , tenta aparelhar-se com a vizinhança de x . Entende-se que cada pixel relevante de B_x deve encontrar-se na mesma posição na vizinhança de x . Caso seja verificado, o ponto central na imagem do resultado será um pixel relevante. Caso contrário, ele será marcado como irrelevante. Veja a erosão com X e B definidos no exemplo 5.3.

Exemplo 5.3

$$X = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \text{ e } B = \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\}$$

Por simplificação, o ponto central PC no elemento estruturante B que não estará marcado será o centro físico de B .

Em cada posição x de X , deve ser posicionado B , virando assim B_x . Pode-se ver imediatamente que nos pixels nas bordas de X , B_x está parcialmente fora. A aplicação da definição 5.2 não pode ser rigorosamente seguida. Este fato se ocorrerá sempre porque a borda de uma imagem é por definição descontínua. Portanto, em todos os exemplos citados, as bordas serão ignoradas e as imagens X serão escolhidas de maneira a não fazer aparecer pixels relevantes naquelas bordas.

No primeiro caso onde x é apontado com "[]", B_x não se aparelha com a vizinhança de x . Os pontos relevantes de B não coincidem com os pixels relevantes da vizinhança de x . Então, o x não pertence ao resultado parcial do conjunto erodido, e vira um pixel irrelevante.

$$\left[\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \text{ero} \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} = \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} \cap \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} \right] = \left[\begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right]$$

No segundo caso, x apontado com "[]", B_x se aparelha com a vizinhança de x . Os pontos relevantes de B coincidem com os pixels relevantes da vizinhança de x . Então, o x pertence ao conjunto erodido, e vira um pixel relevante.

$$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \text{ero} \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} = \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \subset \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$$

Como resultado final, tem-se :

$$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \text{ero} \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$$

O papel do ponto central PC do elemento estruturante é de definir a posição do resultado de X com B . O que acontece quando o ponto central muda de lugar? Seja o exemplo 5.4 parecido com o anterior onde PC não é mais o centro de B :

Exemplo 5.4

$$X = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \text{ e } B = \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\}$$

No primeiro caso onde x é apontado com $\begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}$, B_x não se aparelha com a vizinhança de x . Então, o ponto x vira um pixel irrelevante.

$$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \text{ero} \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} = \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \subset \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$$

No segundo caso, x apontado com $\begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}$, B_x se aparelha com a vizinhança de x . Os pontos relevantes de B coincidem com os pixels relevantes da vizinhança de x . Então, o x pertence ao conjunto erodido, e vira um pixel relevante.

$$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \text{ero} \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} = \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \subset \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$$

Seja o resultado final:

$$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \text{ero} \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$$

Pode-se constatar que o conjunto obtido no exemplo 5.4 é idêntico ao obtido no exemplo 5.3. A única diferença ocorre na posição do resultado. **Mudar a posição do ponto central PC do elemento estruturante B gera uma translação do resultado.**

5.4.3.2 Segunda Definição

[Ser82] e [BB94] demonstraram que a primeira definição 5.2 da erosão pode ser relacionada com a subtração de Minkowski (definição 5.7). Seguindo o formalismo de [Ser82], tem-se a seguinte formulação:

Definição 5.3

A erosão de um conjunto X pelo elemento estruturante B é:

$$\text{ero}^B(X) = X \text{ ero } B = X \ominus \tilde{B} = \cap_{b \in \tilde{B}} X_b = \cap_{b \in B} X_{-b}$$

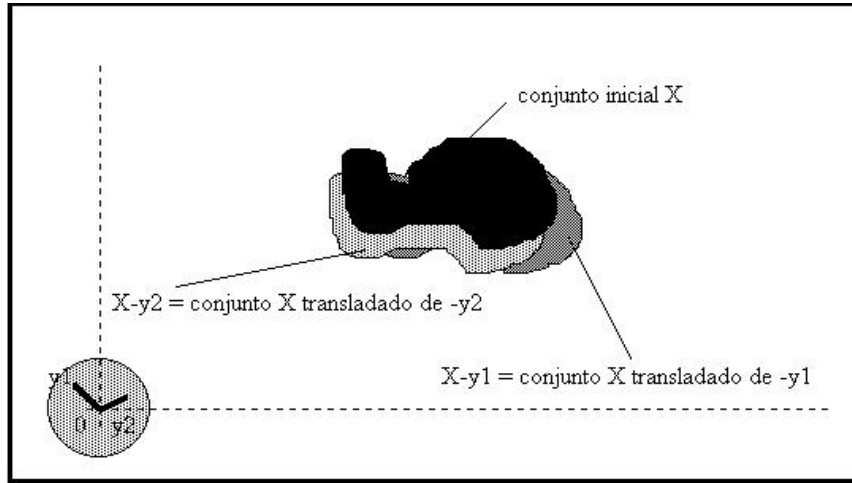


figura 5.6: Interpretação da erosão binária pela subtração de Minkowski

Nesta definição, pode-se constatar que o conjunto a ser erodido, X , é deslocado em função das posições permitidas pelo elemento estruturante B . Entende-se que, na diferença da definição 5.2, tem-se que transladar X e não mais B . Os deslocamentos são realizados em relação ao ponto central de $\tilde{B}$. Veja a diferença de processamento com X e B do exemplo 5.3.

$$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \text{ero} \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \cap \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \cap \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$$

Cada movimento de X é realizado em relação ao ponto central PC de $\tilde{B}$. Aqui B é simétrico, $\tilde{B} = B$. PC representa portanto a origem das translações. No exemplo acima, PC faz parte dos pontos relevantes de $\tilde{B}$. O que significa que $X_{PC} = X_{0,0}$, que representa o conjunto de origem, constitui o primeiro conjunto da interseção. Caso PC não fizesse parte de $\tilde{B}$, $X_{0,0}$ não teria razão de participar da interseção. O elemento estruturante $\tilde{B}$ tem, além de PC , dois outros pontos relevantes, $(PC_x, PC_y - 1)$ e $(PC_x, PC_y + 1)$, onde o $eixo_x$ representa a linha horizontal que passa por PC e o $eixo_y$ a linha vertical que passa por PC e orientada positivamente por cima. Relativamente ao primeiro ponto, tem-se o conjunto $X_{0,-1}$, e para o segundo ponto tem-se $X_{0,1}$. O conjunto resultado é a interseção de $X_{0,0}$, $X_{0,-1}$ e $X_{0,1}$.

A diferença mais importante entre as definições 5.2 e 5.3 situa-se no fato que, enquanto na primeira definição se deve verificar o resultado da interação de B com X (o que consome tempo e cpu), na segunda definição as operações são realizadas "ao pé da letra" sem verificação. Será visto posteriormente a consequência em termos de velocidade e de desempenho dessa diferença.

Agora, veja o exemplo 5.5 onde o ponto central PC de $\tilde{B}$ não faz parte dos pontos relevantes de $\tilde{B}$. $\tilde{B}$ é ainda simétrico, $\tilde{B} = B$. A posição de PC não mudou. A única diferença vem do fato que PC , por não ser relevante, não pode contribuir na criação de $X_{0,0}$. Então tem-se dois X_b que são $X_{0,-1}$ e $X_{0,1}$. Donde o resultado final que, nesse caso particular, é idêntico:

Exemplo 5.5

$$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \text{ero} \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \cap \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$$

5.4.3.3 Efeitos da erosão

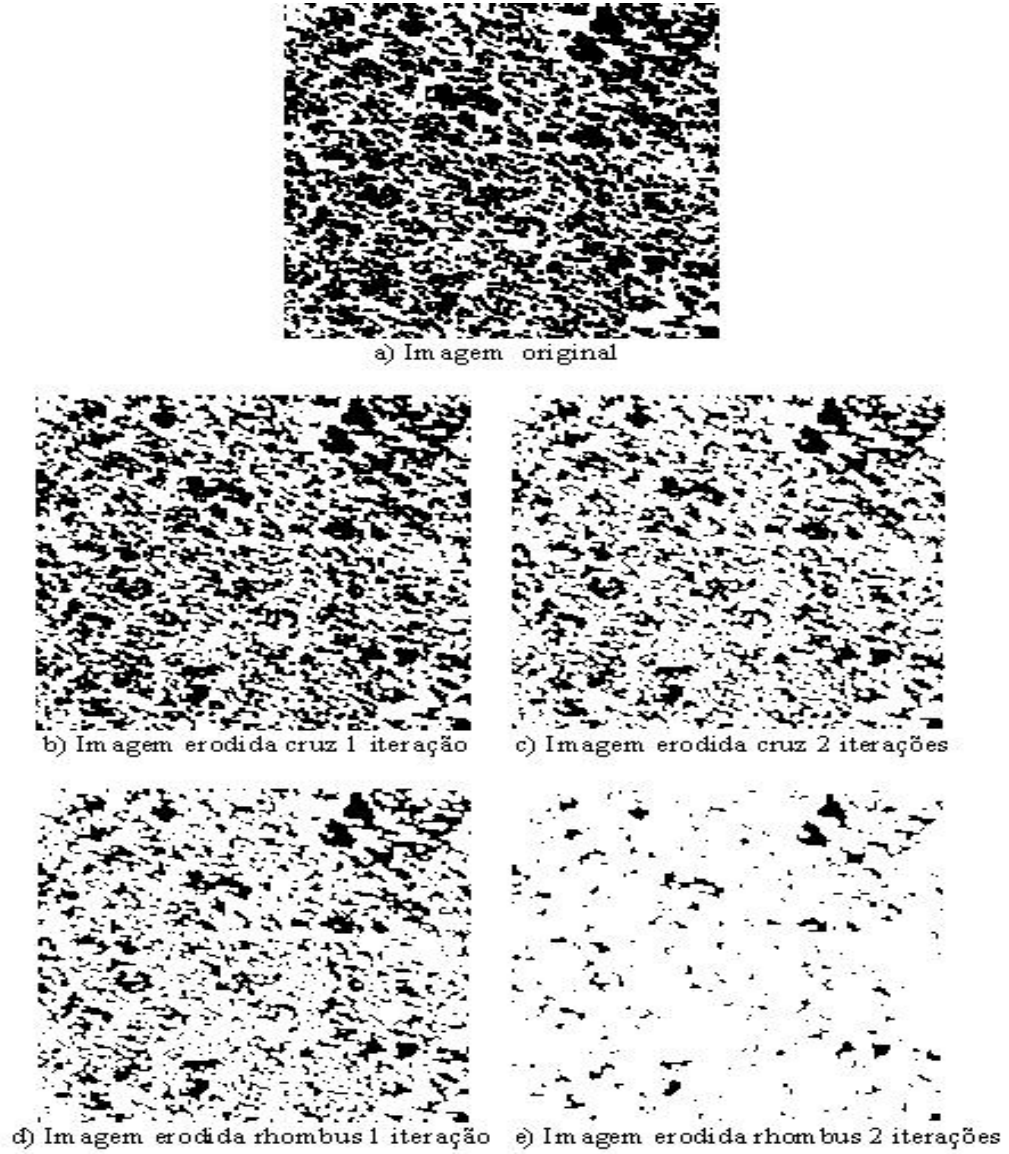


figura 5.7: Exemplos de erosão binária

Pelos exemplos anteriores, pode ser constatado que a erosão modifica o conjunto X . Este fica menor em todos os casos. Através de exemplos, será ilustrado outros efeitos genéricos importantes.

No exemplo 5.6, pode-se constatar que, por erosão, o elemento estruturante elimina os conjuntos inferiores a ele em tamanho.

Exemplo 5.6

$$\begin{bmatrix} \dots \\ \dots \\ \cdot \\ \dots \\ \dots \end{bmatrix} \text{ero} \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} = \begin{bmatrix} \dots \\ \dots \\ \cdot \\ \dots \\ \dots \end{bmatrix} \cap \begin{bmatrix} \dots \\ \cdot \\ \cdot \\ \dots \\ \dots \end{bmatrix} \cap \begin{bmatrix} \dots \\ \dots \\ \cdot \\ \dots \\ \dots \end{bmatrix} = \begin{bmatrix} \dots \\ \dots \\ \cdot \\ \dots \\ \dots \end{bmatrix}$$

No exemplo 5.7, pode-se constatar que a erosão aumenta os furos interiores aos conjuntos.

Exemplo 5.7

$$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{ero} \left\{ \begin{array}{c} \cdot \\ \cdot \\ \cdot \end{array} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \cap \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \cap \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \cap \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \cap \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

Um outro caso interessante de ser estudado é uma imagem apresentando dois conjuntos X_1 e X_2 conectados (exemplo 5.8). Pode-se constatar que, usando um elemento estruturante adequado em tamanho e orientação (nesse caso um elemento estruturante vertical de três pixels para eliminar a ligação horizontal de tamanho 1), a erosão permite separá-los.

Exemplo 5.8

$$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{ero} \left\{ \begin{array}{c} \cdot \\ \cdot \\ \cdot \end{array} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \cap \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \cap \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

Agrupando os resultados apresentados nesses exemplos, pode-se concluir que os efeitos obtidos pela erosão são (figura 5.7):

- diminuir as partículas
- eliminar grãos de tamanho inferior ao tamanho do elemento estruturante
- aumentar os buracos
- permitir a separação de grãos próximos.

Na figura 5.7 são ilustrados os efeitos da erosão sobre uma imagem contendo células, a partir de um elemento estruturante em cruz $\left\{ \begin{array}{c} \cdot \\ \cdot \\ \cdot \end{array} \right\}$ (imagem erodida cruz)

e a partir do elemento estruturante Rhombus $\left\{ \begin{array}{ccc} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{array} \right\}$ (imagem erodida rhombus).

5.4.4 Dilatação binária

5.4.4.1 Primeira Definição

[Ser82] define a operação de dilatação binária **dil** da seguinte maneira:

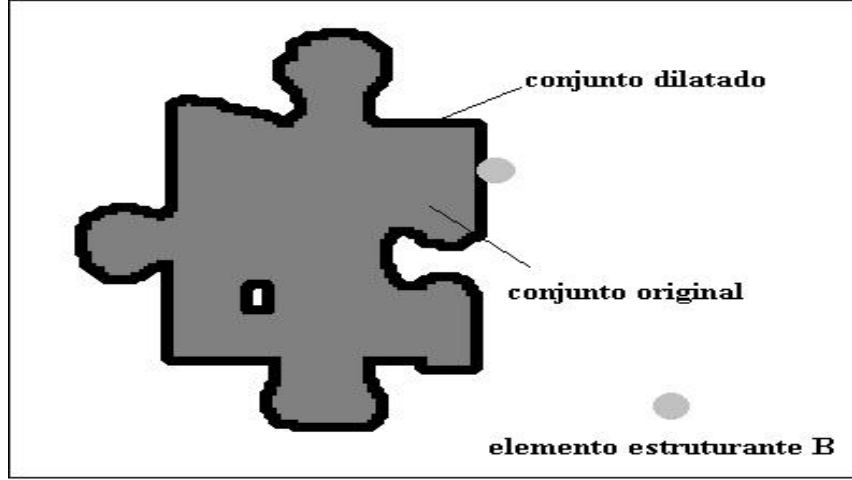


figura 5.8: Princípio da dilatação binária

Definição 5.4 *A dilatação de um conjunto X pelo elemento estruturante B é:*

$$\text{dil}^B(X) = X \text{ dil } B = \{x \in X : B_x \cap X \neq \emptyset\}$$

As duas operações de erosão e de dilatação são portanto duais e a interpretação da dilatação é complementar da interpretação da erosão. O complemento da proposição “ B_x está incluído em X ” é a proposta “a interseção de B_x e X não é vazia” (figura 5.8).

Segundo a definição 5.4, o elemento estruturante B deve deslizar na imagem X (figura 5.8). O significado é que o elemento estruturante B_x , posicionado e centrado em cada pixel x de X , verifica uma possível interseção com a vizinhança de x . Caso seja verdadeiro, o ponto central na imagem resultado será um pixel relevante (•). Caso contrário, ele será marcado como irrelevante (.). Veja a dilatação com X e B definidos no exemplo 5.9:

Exemplo 5.9

$$X = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \text{ e } B = \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\}$$

B é posicionado em cada posição x de X , virando assim B_x . Pode-se ver imediatamente que nos pixels nas bordas de X , os exemplos serão escolhidos de maneira a não fazer aparecer pixels relevantes nas bordas para evitar ambigüidades nelas.

No primeiro caso onde x é apontado com [], os pontos relevantes de B_x não tem interseção com os pixels correspondentes na vizinhança de x . Então, na posição atual do ponto central PC de B_x em X , é colocado um pixel irrelevante.

$$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \text{dil} \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} = \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}$$

No segundo caso, x apontando com [], B_x chega a ter pelo menos interseção com os pixels correspondentes na vizinhança de x . Pelo menos, um ponto relevante de B coincide com um pixel relevante da vizinhança de x . Então, o x pertence ao conjunto dilatado e vira um pixel relevante.

$$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \text{dil} \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} = \left[\begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right] = \left[\begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right]$$

Confira o resultado final:

$$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \text{dil} \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$$

5.4.4.2 Segunda Definição

[Ser82] e [BB94] demonstraram que a primeira definição 5.4 da dilatação pode ser relacionada com a adição de Minkowski (definição 5.5). Seguindo o formalismo de [Ser82], tem-se a seguinte formulação:

Definição 5.5 *A dilatação de um conjunto X pelo elemento estruturante B é:*

$$\text{dil}^B(X) = X \text{ dil } B = X \oplus \tilde{B} = \cup_{b \in \tilde{B}} X_b$$

Na definição 5.5, pode-se constatar que, o conjunto que deve ser dilatado, X , é deslocado em função das posições permitidas pelo elemento estruturante $\tilde{B}$. Entende-se que, na diferença da definição 5.4, não é mais $\tilde{B}$ que translada-se mas X . Os deslocamentos são realizados em relação ao ponto central de $\tilde{B}$. Veja o seguinte exemplo 5.10:

Exemplo 5.10

$$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \text{dil} \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \oplus \widetilde{\left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\}} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \oplus \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \cup \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$$

X é movimentado em relação ao ponto central PC de $\tilde{B}$ que representa portanto a origem das translações. PC faz parte dos pontos relevantes de $\tilde{B}$. O que significa que $X_{PC} = X_{0,0}$, que representa o conjunto de origem, constitui o primeiro conjunto da união. Se PC não fizesse parte de B , $X_{0,0}$ não participaria dessa união. No exemplo, o elemento estruturante não é simétrico, isto faz com que $B \neq \tilde{B}$. O outro ponto relevante, além de PC , que é o simétrico de $(PC_x, PC_y - 1)$ seja $(PC_x, PC_y + 1)$ faz parte de $\tilde{B}$. O $eixo_x$ representa a linha horizontal que passa por PC e o $eixo_y$ é a linha vertical que passa por PC e orientada positivamente para cima. Relativamente a este ponto, tem-se o conjunto transladado $X_{0,1}$. O conjunto resultado é a união de $X_{0,0}$ e $X_{0,1}$.

Pode-se de novo anotar que a diferença mais importante entre as definições 5.4 e 5.5 se situa na ausência de verificação. Enquanto na primeira definição se deve verificar o resultado da interação de B com X (o que consome tempo), na segunda definição as operações são simplesmente realizadas.

Será estudado o exemplo 5.11 onde o ponto central PC de $\tilde{B}$ não faz parte dos pontos relevantes de $\tilde{B}$. A posição de PC não mudou. A única diferença vem do fato que PC , por não ser relevante, não pode contribuir na criação de $X_{0,0}$. Então tem-se somente dois X_b que são $X_{0,-1}$ e $X_{0,1}$. Donde o resultado final:

Exemplo 5.11

$$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \text{dil} \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \oplus \widetilde{\left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\}} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \cup \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$$

5.4.4.3 Efeitos da dilatação

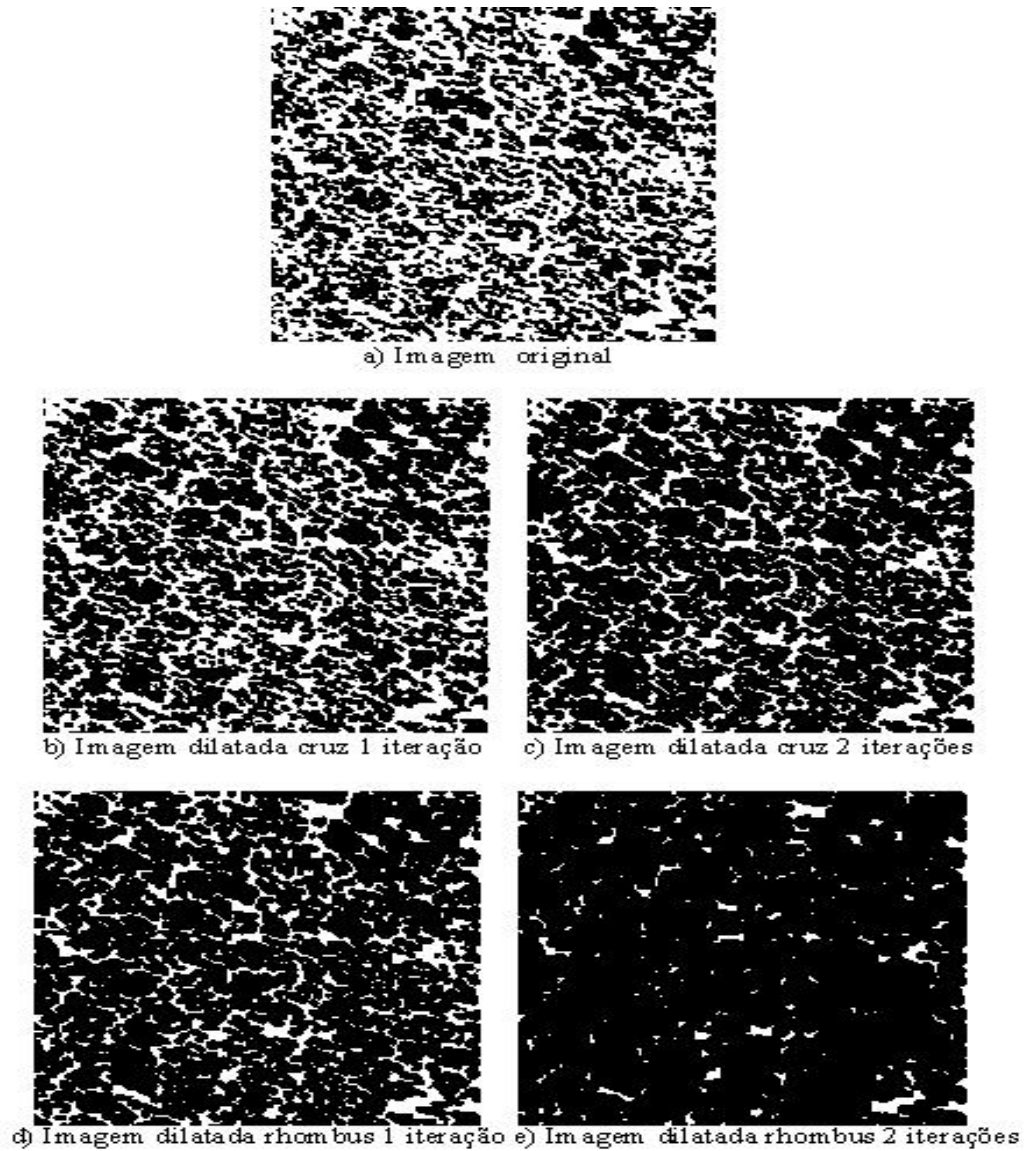


figura 5.9: Exemplos de dilatação binária

Nos exemplos precedentes, vem-se que a dilatação modifica o conjunto X . Este fica maior em todos os casos. Através de outros exemplos, será ilustrado outros efeitos genéricos importantes.

No exemplo 5.12, pode-se observar que, por dilatação, pode-se preencher furos. De forma geral, pode-se afirmar que a dilatação preenche todos os furos que são inferiores em tamanho ao elemento estruturante.

Exemplo 5.12

$$\begin{aligned}
 \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{dil} \left\{ \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \right\} &= \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \oplus \widetilde{\left\{ \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \right\}} \\
 &= \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \cup \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \cup \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \cup \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \\
 &= \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}
 \end{aligned}$$

Uma outra propriedade interessante é apresentada no exemplo 5.13. A dilatação permite conectar conjuntos separados. A maneira mais eficiente é empregar um elemento estruturante maior que o espaço entre conjuntos e que lhe é perpendicular. Aqui o espaço horizontal de tamanho 2 é eliminado com um elemento estruturante horizontal de 3 pixels.

Exemplo 5.13

$$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{dil} \left\{ \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \cup \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \cup \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

Agrupando os resultados apresentados nesses exemplos, pode-se concluir que os efeitos da dilatação são (figura 5.9):

- engordar partículas
- preencher pequenos buracos
- conectar grãos próximos.

A figura 5.9 apresenta os efeitos da dilatação, a partir de um elemento estruturante em cruz $\left\{ \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \right\}$ (imagem dilatada cruz) e a partir do elemento estruturante Rhombus $\left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\}$ (imagem dilatada rhombus) sobre uma imagem contendo células.

5.4.5 Propriedades da Erosão e da Dilatação binárias

Propriedade 5.1 Enquanto a erosão é uma transformação não comutativa, a dilatação é comutativa:

$$\mathbf{ero}^B(X) \neq \mathbf{ero}^X(B) \quad (5.12)$$

$$\mathbf{dil}^B(X) = \mathbf{dil}^X(B) \quad (5.13)$$

O exemplo 5.14 ilustra a equação 5.12 dessa propriedade:

Exemplo 5.14 $\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \mathbf{ero} \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \quad mas \quad \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \mathbf{ero} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$

O exemplo 5.15 ilustra a equação 5.13 dessa propriedade:

Exemplo 5.15 $\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \mathbf{dil} \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \mathbf{dil} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$

Propriedade 5.2 *A erosão e a dilatação por um ponto reduzem-se à uma translação. Como consequência, a dilatação e a erosão são invariantes por translação.*

$$\mathbf{ero}^x(X) = \mathbf{dil}^x(X) = X_x \quad (5.14)$$

O exemplo 5.16 ilustra a equação 5.14 desta propriedade:

Exemplo 5.16

$$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \mathbf{ero} \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \mathbf{dil} \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$$

Propriedade 5.3 *A erosão e a dilatação tem comportamentos interessantes em relação à interseção e à união.*

$$\mathbf{dil}^{B \cup B'}(X) = \mathbf{dil}^B(X) \cup \mathbf{dil}^{B'}(X) \quad (5.15)$$

Pelo uso da dualidade:

$$\mathbf{ero}^{B \cup B'}(X) = \mathbf{ero}^B(X) \cap \mathbf{ero}^{B'}(X) \quad (5.16)$$

e

$$\mathbf{ero}^B(X \cap Z) = \mathbf{ero}^B(X) \cap \mathbf{ero}^B(Z) \quad (5.17)$$

As relações mostram que um conjunto X pode ser dilatado ou erodido tomando B por pedaços e combinando os resultados pelas operações de união ou de interseção.

O exemplo 5.17 ilustra a equação 5.15 dessa propriedade:

Exemplo 5.17

$$\begin{aligned} \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \mathbf{dil} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} &= \left(\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \mathbf{dil} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} \right) \cup \left(\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \mathbf{dil} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} \right) \\ &= \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \cup \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \end{aligned}$$

O exemplo 5.18 ilustra a equação 5.17 dessa propriedade:

Exemplo 5.18

$$\begin{aligned} \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{ero} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} &= \left(\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{ero} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} \right) \cap \left(\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{ero} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} \right) \\ &= \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \cap \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \end{aligned}$$

Propriedade 5.4 A erosão e a dilatação têm duas propriedades interessantes relativas à repetição.

A primeira é:

$$\text{ero}^{B'}(\text{ero}^B(X)) = \text{ero}^{\text{dil}^{B'}(\tilde{B})}(X) \quad (5.18)$$

A equação 5.18 é detalhada no exemplo 5.19 para colocar em evidência as vantagens da decomposição do elemento estruturante. Na primeira fase, pela definição de Minkowski, o elemento estruturante inteiro (não decomposto) exige 8 deslocamentos (um para cada ponto diferente do ponto central). No caso da dilatação decomposta em duas etapas, em cada uma delas, usa-se um elemento estruturante de somente 3 pontos, ocorrendo então 2 deslocamentos. Seja no total das duas etapas, 4 translações contra 8 no primeiro caso. O que ocasiona uma economia de 4 operações para cada pixel processado da imagem X :

Exemplo 5.19

$$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{ero} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{ero} \left(\left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} \text{dil} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} \right) = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

Neste exemplo, a imagem X contém 49 pixels. Sem os das bordas, sobram 25 pixels. Usando o elemento estruturante não decomposto $\left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\}$ foram efetuados 8×25 , ou seja 200 deslocamentos, contra somente 100 ($2 \times (2 \times 25)$) decompondo o elemento em $\left(\left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} \text{dil} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} \right)$. O que significa uma economia de 50%.

A segunda propriedade é:

$$\text{dil}^{B'}(\text{dil}^B(X)) = \text{dil}^{\text{dil}^{B'}(\tilde{B})}(X) \quad (5.19)$$

Ilustrando a equação 5.19 dessa propriedade no exemplo 5.20, tem-se :

Exemplo 5.20

$$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{dil} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{dil} \left(\left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} \text{dil} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} \right) = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

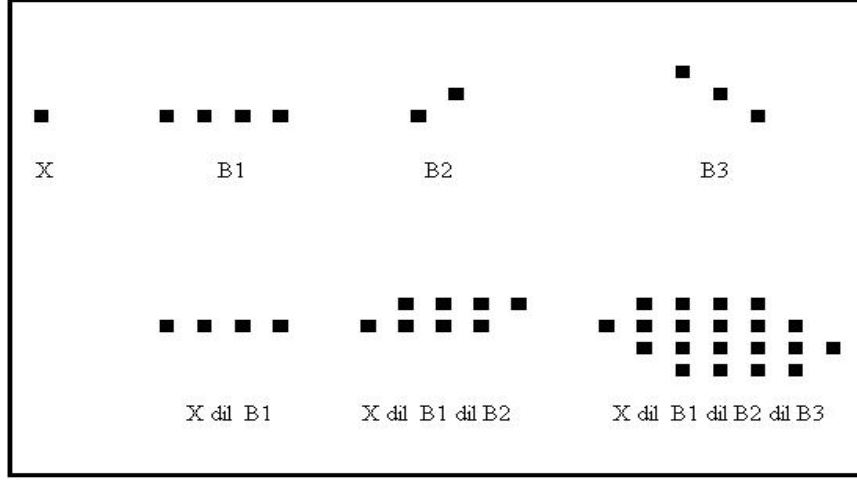


figura 5.10: Exemplo de composição do elemento estruturante por dilatação

Essas propriedades são fundamentais porque mostram que a erosão e a dilatação por um elemento estruturante podem ser decompostas em erosões e dilatações por elementos estruturantes elementares. A aplicação dessas propriedades reside no fato que a maioria dos elementos estruturantes podem ser então compostos pela adição de Minkowski a partir de elementos mais simples (figura 5.10).

Essa propriedade é particularmente interessante para elementos estruturantes convexas. Erosões e dilatações bidimensionais podem ser assim realizadas por erosões e dilatações unidimensionais.

Propriedade 5.5 *A erosão e a dilatação são operações crescentes.*

$$X \subset X' \implies \text{ero}^B(X) \subset \text{ero}^B(X') \quad (5.20)$$

$$\implies \text{dil}^B(X) \subset \text{dil}^B(X') \quad (5.21)$$

Ilustrando no exemplo 5.21 a equação 5.21:

Exemplo 5.21

$$\begin{aligned} \begin{bmatrix} \dots & \dots & \dots \\ \dots & \bullet & \dots \\ \dots & \bullet & \dots \\ \dots & \bullet & \dots \\ \dots & \bullet & \dots \\ \dots & \dots & \dots \end{bmatrix} \subset \begin{bmatrix} \dots & \dots & \dots \\ \dots & \bullet & \dots \\ \dots & \bullet & \dots \\ \dots & \bullet & \dots \\ \dots & \bullet & \dots \\ \dots & \dots & \dots \end{bmatrix} \implies \begin{bmatrix} \dots & \dots & \dots \\ \dots & \bullet & \dots \\ \dots & \bullet & \dots \\ \dots & \bullet & \dots \\ \dots & \bullet & \dots \\ \dots & \dots & \dots \end{bmatrix} \text{ero} \left\{ \begin{bmatrix} \dots \\ \dots \\ \dots \end{bmatrix} \right\} \subset \begin{bmatrix} \dots & \dots & \dots \\ \dots & \bullet & \dots \\ \dots & \bullet & \dots \\ \dots & \bullet & \dots \\ \dots & \bullet & \dots \\ \dots & \dots & \dots \end{bmatrix} \text{ero} \left\{ \begin{bmatrix} \dots \\ \dots \\ \dots \end{bmatrix} \right\} \\ \implies \begin{bmatrix} \dots & \dots & \dots \\ \dots & \bullet & \dots \\ \dots & \bullet & \dots \\ \dots & \bullet & \dots \\ \dots & \bullet & \dots \\ \dots & \dots & \dots \end{bmatrix} \subset \begin{bmatrix} \dots & \dots & \dots \\ \dots & \bullet & \dots \\ \dots & \bullet & \dots \\ \dots & \bullet & \dots \\ \dots & \bullet & \dots \\ \dots & \dots & \dots \end{bmatrix} \end{aligned}$$

E por dualidade:

$$B \subset B' \implies \text{ero}^{B'}(X) \subset \text{ero}^B(X) \quad (5.22)$$

$$\implies \text{dil}^B(X) \subset \text{dil}^{B'}(X) \quad (5.23)$$

O exemplo 5.22 ilustra a equação 5.23:

Exemplo 5.22

$$\begin{aligned} \left\{ \begin{array}{c} \cdot \\ \cdot \\ \cdot \end{array} \right\} \subset \left\{ \begin{array}{cc} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{array} \right\} &\Rightarrow \left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right] \text{ero} \left\{ \begin{array}{cc} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{array} \right\} \subset \left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right] \text{ero} \left\{ \begin{array}{c} \cdot \\ \cdot \\ \cdot \end{array} \right\} \\ &\Rightarrow \left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right] \subset \left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right] \end{aligned}$$

Propriedade 5.6 *A dilatação é um operação extensiva.*

$$X \subset \text{dil}^B(X) \quad (5.24)$$

Ilustrando a equação 5.24 da propriedade de extensividade no exemplo 5.23:

Exemplo 5.23

$$\left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right] \subset \left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right] \text{dil} \left\{ \begin{array}{c} \cdot \\ \cdot \\ \cdot \end{array} \right\} = \left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right]$$

Propriedade 5.7 *A erosão é uma operação anti-extensiva.*

$$\text{ero}^B(X) \subset X \quad (5.25)$$

Ilustrando a equação 5.25 da propriedade de anti-extensividade no exemplo 5.24:

Exemplo 5.24

$$\left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right] \text{ero} \left\{ \begin{array}{ccc} \cdot & \cdot & \cdot \end{array} \right\} \left(= \left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right] \right) \subset \left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right]$$

Propriedade 5.8 *A dilatação e a erosão são transformações contínuas.*

5.4.6 Principais classes de elementos estruturantes

Sabe-se que o princípio básico da morfologia matemática reside numa operação de conjuntos entre o conjunto em teste e o elemento estruturante conhecido. Através desse, é possível obter uma informação relativa à geometria e à topologia desse conjunto. Toda a dificuldade é conseguir escolher o elemento estruturante adequado para o resultado procurado (figura 5.11).

Para facilitar a interpretação, os elementos estruturantes devem ser os mais simples possíveis. Na maioria dos casos, os elementos estruturantes são escolhidos em função das propriedades de convexidade, não-convexidade, isotropia e anisotropia. A erosão pelo segmento de reta leva a anisotropia através da noção de orientação. As erosões pelo círculo e segmento de reta são semelhantes pelo fato de que esses elementos são convexos. O círculo serve mais particularmente na determinação do tamanho dos objetos. O segmento de reta, também convexo mas anisotrópico, serve para estudar as distribuições de tamanho. A erosão pelo par de pontos permite caracterizar o estado de dispersão de uma estrutura.

Do ponto de vista digital, os elementos estruturantes o segmento de reta e o par de pontos são definidos sem ambigüidade, como por exemplo: $B_{\text{segmento}} = \left\{ \begin{array}{ccc} \cdot & \cdot & \cdot \end{array} \right\}$ e $B_{\text{par pontos}} = \left\{ \begin{array}{cc} \cdot & \cdot \end{array} \right\}$

elemento estruturante	convexo	não convexo
isótopo		
não isótopo		

figura 5.11: Exemplos de elementos estruturantes

Já o elemento estruturante o círculo é mais difícil de ser representado. Pode ser usado um elemento estruturante aproximado grosseiro como uma *cruz*, um *quadrado*. Uma versão mais correta pode ser o elemento estruturante *rhombus*:

$$B_{cruz} = \left\{ \begin{array}{c} \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{array} \right\} \quad B_{quadrado} = \left\{ \begin{array}{ccc} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{array} \right\} \quad B_{rhombus} = \left\{ \begin{array}{ccccc} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right\}$$

O resultado da aplicação de um elemento estruturante depende também da posição de seu ponto central. De forma geral pode-se dizer que a mudança da posição do ponto central num mesmo elemento estruturante resulta em uma translação do resultado final (figura 5.12 página 84).

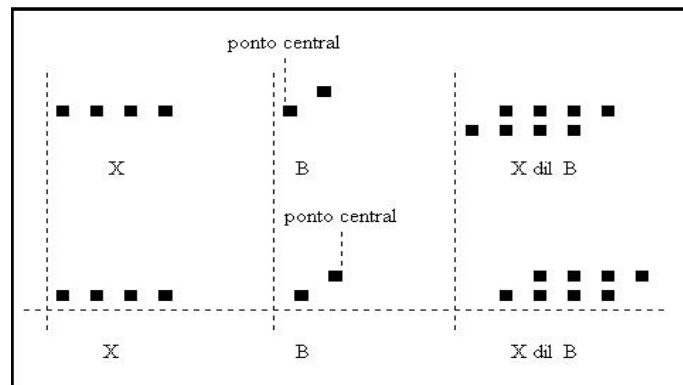


figura 5.12: Exemplos de dilatação a partir do mesmo elemento estruturante com um posicionamento diferente do ponto central

5.4.7 Detecção de bordas

O contorno é um dos conceitos mais importantes da área de Processamento de Imagens. Segmentar uma imagem em termos de segmentos de retas, arcos de círculos, enfim partes de curvas, é uma das tarefas mais comuns. No caso da manipulação de

$$= \begin{bmatrix} \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{bmatrix} - \begin{bmatrix} \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{bmatrix} = \begin{bmatrix} \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{bmatrix}$$

Esse resultado é satisfatório porém não ótimo. O seguinte exemplo é constituído do precedente resultado onde foram marcados com "[]" alguns pontos. Suprimindo esses pontos, o novo contorno não apresenta falhas. Isto significa que os pontos eliminados não modificam a conectividade do contorno, ou seja ainda que os pontos marcados são supérfluos. Isto mostra que o elemento estruturante $B1_{borda}$ permite mas não otimiza a obtenção do contorno, qualquer que seja a imagem.

$$\begin{bmatrix} \dots & \dots & \dots & \dots \\ \dots & \dots & [\bullet] & \dots \\ \dots & [\bullet] & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & [\bullet] & \dots & \dots \\ \dots & \dots & [\bullet] & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix} = \begin{bmatrix} \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}$$

A detecção de bordas por erosão da mesma imagem X do exemplo 5.25 a partir do elemento estruturante $B2_{borda}$ é:

$$\begin{aligned} X - \text{ero}^{B2_{borda}}(X) &= \begin{bmatrix} \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{bmatrix} - \left(\begin{bmatrix} \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{bmatrix} \text{ero} \left\{ \begin{bmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{bmatrix} \right\} \right) \\ &= \begin{bmatrix} \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{bmatrix} - \begin{bmatrix} \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{bmatrix} = \begin{bmatrix} \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{bmatrix} \end{aligned}$$

$B2_{borda}$ representa uma outra alternativa de detecção de bordas. Reduzindo o número de pontos sem prejudicar a conectividade do contorno, $B2_{borda}$ permite otimizar a detecção.

A determinação de elementos estruturantes eficientes para a detecção de contornos fechados em imagens binárias foi realizada empregando a abordagem do operador de erosão. As mesmas conclusões podem ser formuladas usando o operador de dilatação. A única diferença aparece na localização da borda. O contorno obtido por erosão situa-se exatamente na borda do conjunto processado enquanto o contorno obtido por dilatação se localiza um pixel "fora" desta borda. É possível também usar os dois operadores de erosão e dilatação para obter uma borda mais espessa (de dois pixels). A seguir é apresentado um exemplo de detecção de bordas de uma imagem binária pelo elemento estruturante $B2_{borda}$ a partir das operações de erosão e de dilatação.(figura 5.13).

Pode-se concluir que com o elemento estruturante $B1_{borda}$, obtém-se um contorno de boa qualidade mas não sempre da espessura de um pixel. Por outro lado,

$B1_{borda}$ pode ser decomposto em 4 elementos estruturantes elementares, $\begin{Bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{Bmatrix}$, $\begin{Bmatrix} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{Bmatrix}$, $\begin{Bmatrix} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{Bmatrix}$, e $\begin{Bmatrix} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{Bmatrix}$. O segundo elemento estruturante $B2_{borda}$ permite obter um contorno otimizado da espessura de um pixel. Porém, $B2_{borda}$ não pode ser decomposto, o que deixa o processo um pouco mais lento.

5.4.8 Abertura binária

Vimos anteriormente que a erosão e a dilatação podem corrigir defeitos numa imagem como ruídos, furos, conexões, enfim filtrar. Porém, nenhum conjunto retocado por essas operações mantém o mesmo tamanho. A erosão reduz e a dilatação engorda. A partir da propriedade de iteratividade, introduzir-se-á novas operações cujo objetivo consiste em eliminar as partículas indesejáveis sem modificar o tamanho das outras entidades. Intuitivamente, pode-se prever que esta operação consiste em erodir e depois dilatar o resultado da erosão. Definimos assim uma nova operação morfológica chamada de “*abertura binária*” e o novo conjunto processado pelo elemento estruturante B chama-se de *conjunto aberto* por B (figura 5.14).

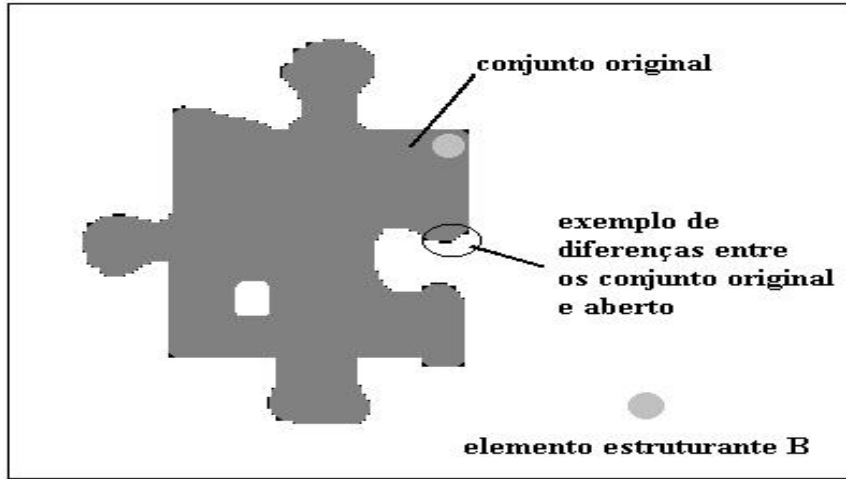


figura 5.14: Princípio da abertura binária

5.4.8.1 Definição da abertura binária

Definida por [Ser82] a partir das definições básicas de Minkowski, a abertura binária **abe** escreve-se como sendo:

Definição 5.6 A abertura de um conjunto X pelo elemento estruturante B é:

$$\mathbf{abe}^B(X) = X \mathbf{abe} B = (X \ominus \tilde{B}) \oplus B$$

Levando em consideração as definições 5.3 e 5.5 da erosão e dilatação, pode-se definir a abertura como sendo:

$$\mathbf{abe}^B(X) = \mathbf{dil}^{\tilde{B}}(\mathbf{ero}^B(X)) \quad (5.27)$$

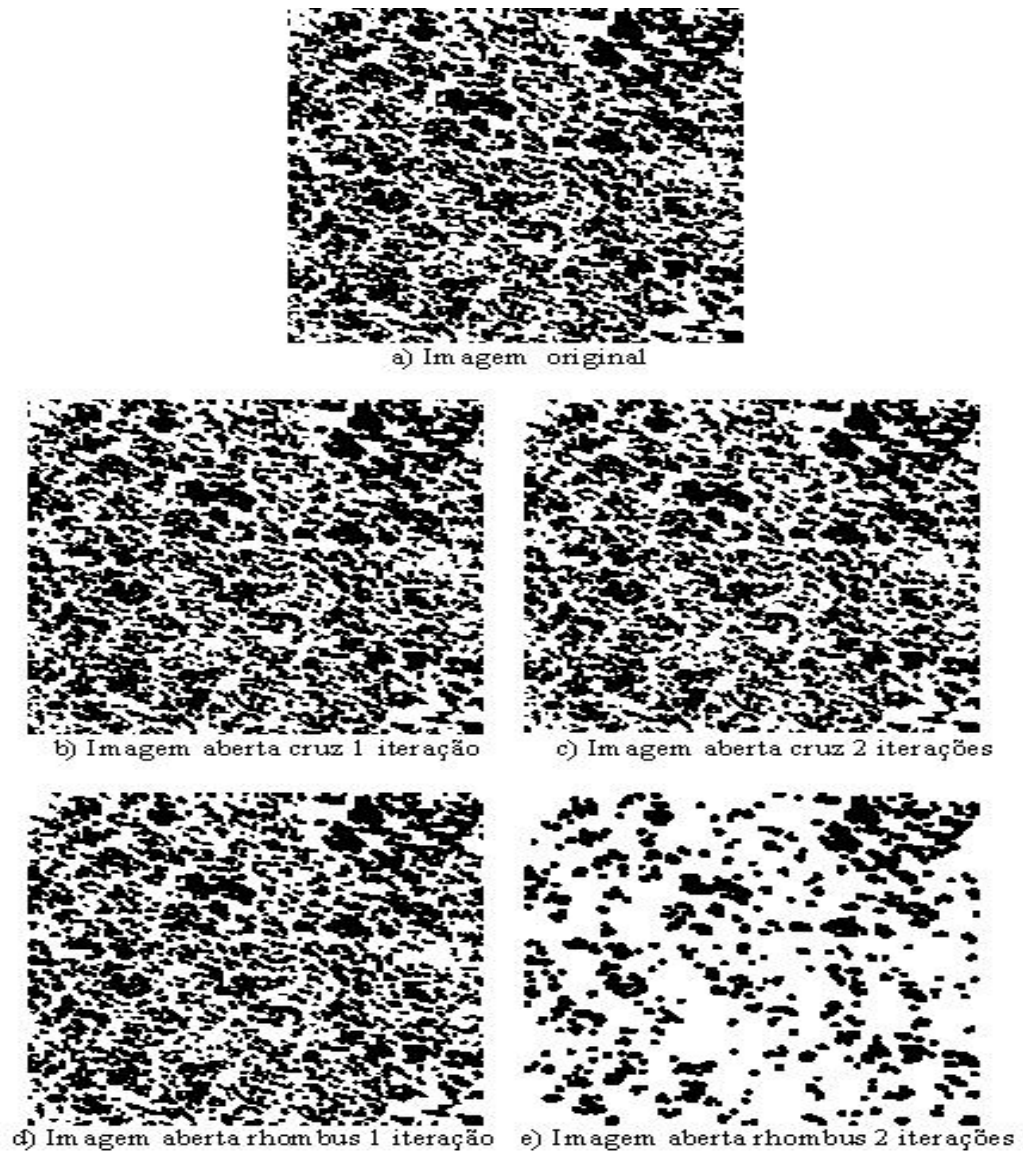


figura 5.15: Exemplos de abertura binária

Ilustra-se a equação 5.6 a partir do exemplo 5.26 onde a imagem X apresenta dois conjuntos conectados:

Exemplo 5.26 $\left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right] \mathbf{abe} \left\{ \begin{array}{c} \cdot \\ \cdot \\ \cdot \end{array} \right\}$

O objetivo é destruir a conexão sem modificar o tamanho e a forma dos conjuntos. Isto pode ser realizado usando um elemento estruturante perpendicular à conexão e de maior tamanho.

$$\left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right] \mathbf{abe} \left\{ \begin{array}{c} \cdot \\ \cdot \\ \cdot \end{array} \right\} = \left(\left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right] \ominus \left\{ \begin{array}{c} \cdot \\ \cdot \\ \cdot \end{array} \right\} \right) \oplus \left\{ \begin{array}{c} \cdot \\ \cdot \\ \cdot \end{array} \right\}$$

O primeiro passo consiste em efetuar a subtração de Minkowski com $\tilde{B}$. Essa subtração consiste em pegar a interseção de X com X deslocado segundo as duas posições permitidas por $\tilde{B}$.

$$\left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right] \ominus \left\{ \begin{array}{c} \cdot \\ \cdot \\ \cdot \end{array} \right\} = \left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right] \cap \left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right] \cap \left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right] = \left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right]$$

O segundo passo consiste em processar o resultado anterior por adição de Minkowski com B . Essa adição consiste em pegar a união de X com X deslocado segundo as posições permitidas por B .

$$\left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right] \oplus \left\{ \begin{array}{c} \cdot \\ \cdot \\ \cdot \end{array} \right\} = \left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right] \cup \left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right] \cup \left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right] = \left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right]$$

O resultado final mostra que os dois conjuntos da imagem aberta $\mathbf{abe}^B(X)$ são desconectados e tem o mesmo tamanho e a mesma forma que os dois conjuntos da imagem original X .

Veja agora o processo de abertura usando da definição da equação 5.27 a partir do mesmo exemplo 5.26. O objetivo é o mesmo, destruir a conexão sem modificar o tamanho e a forma dos conjuntos. O primeiro passo consiste em efetuar a erosão de X com B . Essa erosão consiste em verificar para cada pixel x se $B_x \subset X$ e guardar os pixels x onde essa definição é verificada. O resultado é então dilatado por $\tilde{B}$. Essa dilatação consiste em guardar os pixels x onde $\tilde{B}_x \cap X_{erodido} \neq \emptyset$.

Novamente os dois conjuntos da imagem aberta $\mathbf{abe}^B(X)$ são desconectados e tem o mesmo tamanho e a mesma forma que os dois conjuntos da imagem original X .

$$\begin{aligned}
\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{abe} \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} &= \left(\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{ero} \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} \right) \text{dil} \left\{ \widetilde{\begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}} \right\} \\
&= \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{dil} \left\{ \widetilde{\begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}
\end{aligned}$$

5.4.8.2 Efeitos da abertura binária

Nos exemplos precedentes, vimos alguns resultados da abertura. Principalmente vimos que a abertura separa os conjuntos. Através de outros exemplos, ilustra-se outros efeitos genéricos importantes.

No exemplo 5.27, pode-se verificar que, quando o elemento estruturante B é maior que X , o resultado da abertura é o conjunto vazio:

Exemplo 5.27

$$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \text{abe} \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} = \left(\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \text{ero} \left\{ \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \right\} \right) \text{dil} \left\{ \widetilde{\begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \text{dil} \left\{ \widetilde{\begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$$

O usuário pode pensar que efetuar uma erosão seguida de uma dilatação devolve o conjunto inicial. De forma geral, isto não acontece. Porque o conjunto aberto é de forma geral mais regular e menos rico em detalhes. Seja o conjunto X , do exemplo 5.28, com contornos irregulares a ser aberto. Pode-se constatar no resultado que o conjunto final é mais regular e contém menos detalhes.

Exemplo 5.28

$$\begin{aligned}
\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{abe} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} &= \left(\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{ero} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} \right) \text{dil} \left\{ \widetilde{\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}} \right\} \\
&= \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{dil} \left\{ \widetilde{\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}
\end{aligned}$$

Pode-se concluir que (figura 5.15):

- a abertura não devolve, de forma geral, o conjunto inicial
- a abertura nivela os contornos pelo interior
- a abertura separa as partículas
- a abertura elimina as pequenas partículas inferiores em tamanho ao em relação ao elemento estruturante
- as entidades restantes após abertura ficam quase idênticas às originais
- o conjunto aberto é mais regular que o conjunto inicial
- o conjunto aberto é menos rico em detalhes que o conjunto inicial

5.4.9 Fechamento binário

Seguindo o mesmo raciocínio usado na abertura, pode-se realizar a operação dual, ou seja dilatar e depois erodir o resultado da dilatação.

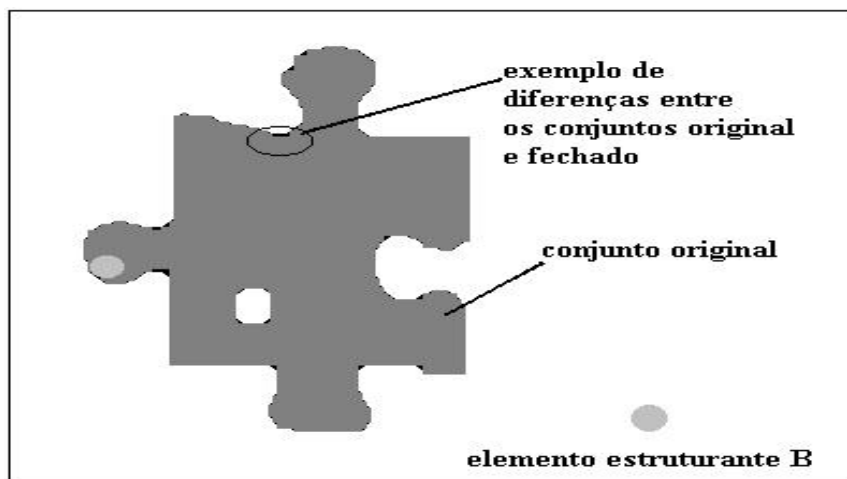


figura 5.16: Princípio do fechamento binário

Esta operação morfológica chama-se de “*fechamento binário*” e o novo conjunto processado pelo elemento estruturante B chama-se de *conjunto fechado* por B (figura 5.16). O fechamento é a operação dual da abertura binária.

5.4.9.1 Definição do fechamento binário

O fechamento binário é por dualidade:

$$(\text{abe}^B(X))^c = X \text{ abe } B = X^c \text{ fec } B = \text{fec}^B(X^c)$$

Ou seja:

Definição 5.7
Definição 5.7
 O fechamento de um conjunto X pelo elemento estruturante B é:

O fechamento de um conjunto X pelo elemento estruturante B é:

$$\text{fec}^B(X) = (X \oplus \tilde{B}) \ominus B$$

Exemplo 5.29

$$\left[\begin{array}{cccccccc} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right] \text{fec} \left\{ \begin{array}{cc} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{array} \right\}$$

O resultado final mostra os dois conjuntos de X fechado conectados e com tamanho e forma idênticos aos de X inicial.

$$\left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right] \text{fec} \left\{ \begin{array}{cc} \cdot & \cdot \\ \cdot & \cdot \end{array} \right\} = \left(\left[\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right] \oplus \widetilde{\left\{ \begin{array}{cc} \cdot & \cdot \\ \cdot & \cdot \end{array} \right\}} \right) \ominus \left\{ \begin{array}{cc} \cdot & \cdot \\ \cdot & \cdot \end{array} \right\}$$

$$\left[\begin{array}{cccc} \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \end{array} \right] \oplus \left\{ \begin{array}{cc} \bullet & \bullet \\ \bullet & \bullet \end{array} \right\} = \left[\begin{array}{cccc} \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \end{array} \right] \cup \left[\begin{array}{cccc} \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \end{array} \right] \cup \left[\begin{array}{cccc} \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \end{array} \right] \cup \left[\begin{array}{cccc} \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \end{array} \right] = \left[\begin{array}{cccc} \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \end{array} \right]$$
[illegible]

Relacionando a operação de fechamento 5.7 com as operações de dilatação e de erosão, pode-se escrever a seguinte relação:

$$\mathbf{fec}^B(X) = (X \oplus \tilde{B}) \ominus B = (X \oplus \tilde{B}) \ominus \tilde{\tilde{B}}$$

$$\mathbf{fec}^B(X) = X \mathbf{fec} B = (X \mathbf{dil} B) \mathbf{ero} \tilde{B} = \mathbf{ero}^{\tilde{B}}(\mathbf{dil}^B(X)) \quad (5.28)$$

Veja o processamento do mesmo exemplo 5.29 a partir da equação 5.28. No primeiro passo é efetuada a dilatação com B que consiste em guardar os pixels x onde $B_x \cap X \neq \emptyset$. O resultado é processado por erosão com $\tilde{B}$. Essa erosão consiste em conservar os pixels x onde $\tilde{B}_x \subset X_{dilataado}$.

$$\begin{aligned} \begin{bmatrix} \dots & \dots & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix} \text{fec} \left\{ \begin{bmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{bmatrix} \right\} &= \left(\begin{bmatrix} \dots & \dots & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix} \text{dil} \left\{ \begin{bmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{bmatrix} \right\} \right) \text{ero} \left\{ \widetilde{\begin{bmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{bmatrix}} \right\} \\ &= \begin{bmatrix} \dots & \dots & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix} \text{ero} \left\{ \begin{bmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{bmatrix} \right\} = \begin{bmatrix} \dots & \dots & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix} \end{aligned}$$

O resultado final mostra como os dois conjuntos de X fechado estão conectados e mostra também que o tamanho e forma são conforme tamanho e a forma dos dois conjuntos de X inicial.

5.4.9.2 Efeitos do fechamento binário

Nos precedentes exemplos foram vistos alguns resultados do fechamento. Vimos principalmente que este operador junta os conjuntos. Através de outros exemplos, ilustra-se outros efeitos genéricos importantes.

No exemplo 5.30, pode-se observar que o fechamento permite, com um elemento estruturante B adequado, preencher furos que lhe são menores sem modificar o tamanho desse conjunto inicial. O que não seria possível realizar usando somente a dilatação.

Veja o seguinte exemplo:

Exemplo 5.30

$$\begin{bmatrix} \dots & \dots & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix} \text{fec} \left\{ \begin{bmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{bmatrix} \right\}$$

Neste exemplo, estuda-se o comportamento do fechamento em relação ao preenchimento de furos:

$$\begin{bmatrix} \dots & \dots & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix} \text{fec} \left\{ \begin{bmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{bmatrix} \right\} = \left(\begin{bmatrix} \dots & \dots & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \bullet & \bullet & \bullet & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix} \oplus \left\{ \widetilde{\begin{bmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{bmatrix}} \right\} \right) \ominus \left\{ \begin{bmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{bmatrix} \right\}$$

$$\begin{aligned}
&= \left(\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \cup \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \cup \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \cup \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \right) \ominus \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} \\
&= \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \ominus \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} \\
&= \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \cap \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \cap \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \cap \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \\
&= \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}
\end{aligned}$$

No exemplo 5.31, mostra-se que o fechamento não restitui, de forma geral, a imagem original. Isto porque o conjunto fechado é mais regular e menos rico que o conjunto inicial. Escolhe-se para isto um elemento estruturante B simples:

Exemplo 5.31

$$\begin{aligned}
\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{fec} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} &= \left(\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{dil} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} \right) \text{ero} \left\{ \widetilde{\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}} \right\} \\
&= \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{ero} \left\{ \widetilde{\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}
\end{aligned}$$

Em conclusão (figura 5.17):

- o fechamento suaviza as fronteiras pelo exterior
- o fechamento preenche os buracos no interior das partículas inferior em tamanho em relação ao elemento estruturante
- o fechamento emenda partículas próximas
- as entidades restantes após fechamento ficam quase idênticas
- o conjunto fechado é mais regular que o conjunto inicial
- o conjunto fechado é menos rico em detalhes que o conjunto inicial

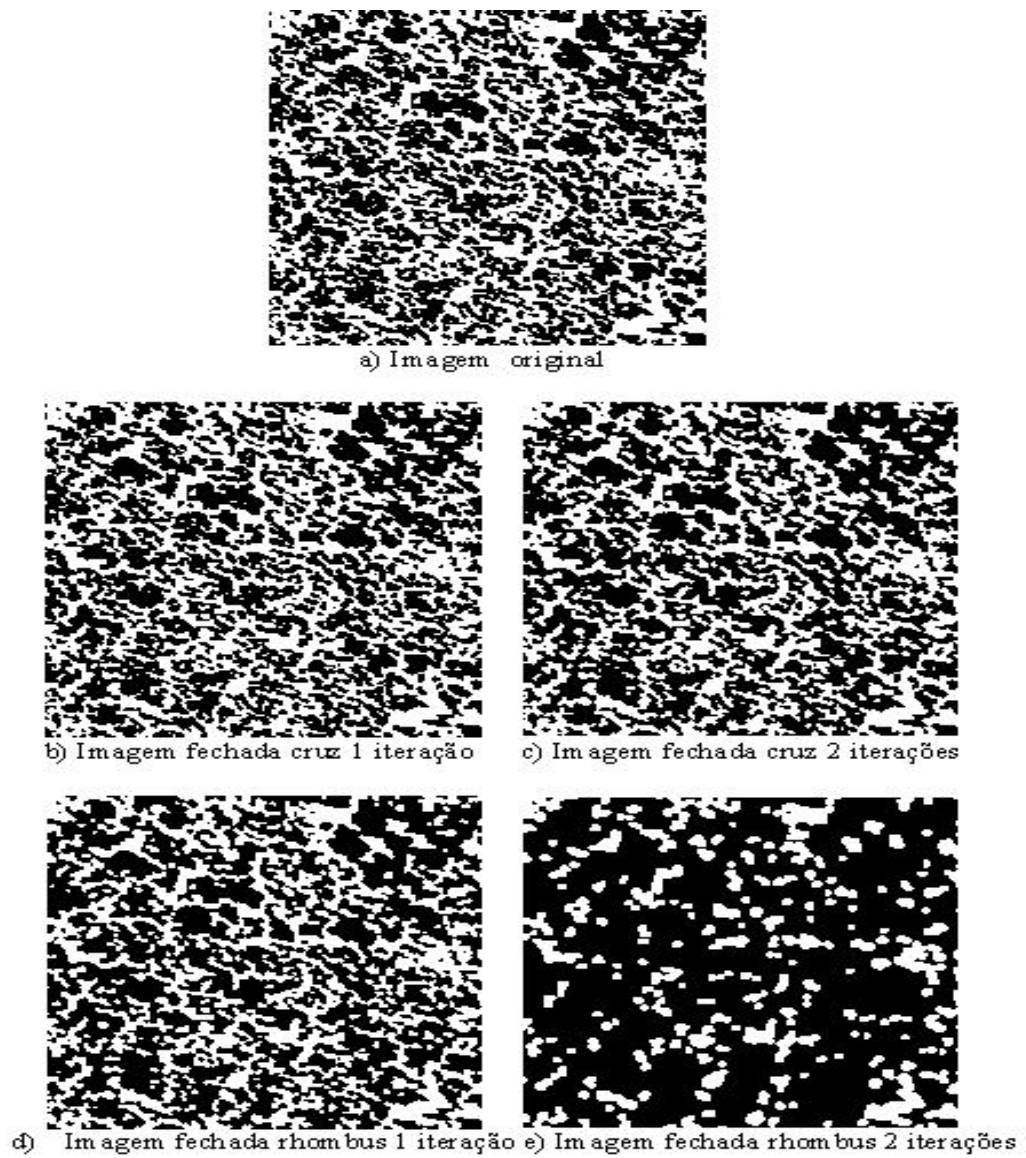


figura 5.17: Exemplos de fechamento binário

5.4.10 Propriedades da abertura e do fechamento

Propriedade 5.9 *A abertura e o fechamento são transformações crescentes.*

$$\begin{aligned} Y \subset X &\implies \mathbf{abe}^B(Y) \subset \mathbf{abe}^B(X) \\ Y \subset X &\implies \mathbf{fec}^B(Y) \subset \mathbf{fec}^B(X) \end{aligned} \quad (5.29)$$

Exemplificando essas propriedades, tem-se :

$$\begin{aligned} \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \subset \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \implies \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \mathbf{abe} \left\{ \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \right\} \subset \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \mathbf{abe} \left\{ \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \right\} \implies \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \subset \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \\ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \subset \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \implies \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \mathbf{fec} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} \subset \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \mathbf{fec} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} \implies \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \subset \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \end{aligned}$$

Propriedade 5.10 *Enquanto a abertura é uma transformação anti-extensiva, o fechamento é uma transformação extensiva.*

$$\begin{aligned} \mathbf{abe}^B(X) &\subset X \\ X &\subset \mathbf{fec}^B(X) \end{aligned} \quad (5.30)$$

Exemplificando essas propriedades, tem-se :

$$\begin{aligned} \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \mathbf{abe} \left\{ \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \right\} &= \left(\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \mathbf{ero} \left\{ \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \right\} \right) \mathbf{dil} \left\{ \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \subset \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \\ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \subset \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \mathbf{fec} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} &\implies \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \subset \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \end{aligned}$$

Propriedade 5.11 *A abertura e o fechamento são transformações idempotentes.*

$$\begin{aligned} \mathbf{abe}^B(\mathbf{abe}^B(X)) &= \mathbf{abe}^B(X) \\ \mathbf{fec}^B(\mathbf{fec}^B(X)) &= \mathbf{fec}^B(X) \end{aligned} \quad (5.31)$$

O seguinte exemplo ilustra essa propriedade muito importante. A imagem apresentada tem dois conjuntos conectados. O objetivo consiste em desconectar esses dois conjuntos sem modificar a forma e o tamanho. A conexão entre eles é de 3 pixels de largura e 4 pixels de comprimento. Para separá-los, sabe-se que tem-se que usar um elemento estruturante perpendicular e maior que o padrão a ser eliminado. Para mostrar o efeito da idempotência, deliberadamente escolhe-se um elemento estruturante simétrico B menor que a conexão, seja ele aqui no formato cruz 3×3 B_{cruz} . Por abertura, o detalhe da etapa de erosão segunda a definição 5.6 mostra

que a conexão é parcialmente destruída mas não o suficiente para desaparecer por dilatação, o que tem como consequência a presença da conexão no resultado final.

$$\begin{aligned}
 \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{abe} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} &= \left(\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \ominus \widetilde{\left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\}} \right) \oplus \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} \\
 &= \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \oplus \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}
 \end{aligned}$$

Para destruir a conexão que foi quase adequadamente eliminada, pode-se ter em mente a idéia de reiterar o processo de abertura com o mesmo elemento estruturante. O que é ilustrado no seguinte exemplo com a definição 5.27.

$$\begin{aligned}
 \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{abe} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} &= \left(\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{ero} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} \right) \text{dil} \left\{ \widetilde{\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}} \right\} \\
 &= \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{dil} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}
 \end{aligned}$$

O resultado mostrado é rigorosamente idêntico ao anterior. Esse exemplo ilustra bem a propriedade muito importante que é a idempotência. Efetuar uma nova abertura sobre o precedente resultado da abertura com o mesmo elemento estruturante não altera o resultado. A única maneira de alterar o resultado consiste em escolher o elemento estruturante como mostrado no exemplo a seguir.

A imagem original é a mesma e o elemento estruturante é agora um retângulo 4×2 $\left\{ \begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} \right\}$.

O processo é portanto:

$$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{abe} \left\{ \begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{dil} \left\{ \widetilde{\begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}} \right\} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

No caso do processo de fechamento, ilustra-se essa propriedade no seguinte exemplo: a imagem tem dois conjuntos que possuem dois buracos internos. O objetivo consiste em preencher os buracos sem emendar os dois conjuntos. O maior tamanho dos buracos é de 3 pixels enquanto o espaço entre eles é de 5 pixels. Para não juntá-los, tem-se que usar um elemento estruturante menor que este espaço. Escolhe-se

Obtem-se portanto um resultado parcialmente satisfatório:

[illegible]

[illegible]
$$\left[\begin{array}{ccccccc} & & & & & & \\ & & & & & & \\ & & & & & & \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ & & & & & & \\ & & & & & & \\ & & & & & & \end{array} \right] \text{fec} \left\{ \begin{array}{cc} . & . \\ . & . \\ . & . \end{array} \right\} = \left[\begin{array}{ccccccc} . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ & & & & & & \\ & & & & & & \\ & & & & & & \end{array} \right] \text{ero} \left\{ \begin{array}{cc} . & . \\ . & . \\ . & . \end{array} \right\} = \left[\begin{array}{ccccccc} & & & & & & \\ & & & & & & \\ & & & & & & \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ & & & & & & \\ & & & & & & \\ & & & & & & \end{array} \right]$$

A abertura e o fechamento diferem da erosão e da dilatação pela propriedade da idempotência. Esta propriedade é crucial nos processos de

5.5 CONDICIONALIDADE E RECONSTRUÇÃO

Os operadores apresentados anteriormente consideravam as imagens como sendo conjuntos indivisíveis. Porém, pode surgir a necessidade de restringir os processos em uma região específica de uma imagem, como por exemplo processar algumas falhas em imagens metalográficas, estudar uma célula específica numa imagem contendo vários tipos delas, etc.... Mostrar-se-á que as transformações morfológicas podem ser modificadas de maneira a trabalhar somente em um subconjunto da imagem. Serão apresentados operadores morfológicos condicionais e a noção de reconstrução de conjuntos.

Uma primeira possibilidade de processar parcialmente uma imagem consiste em definir um subconjunto da imagem onde as operações são válidas, por exemplo, tratar um subconjunto Z do conjunto inicial X . Operadores *erosão e dilatação* ditos *condicionais* permitem realizar esse tipo de processamento.

$$\mathbf{ero}_{c_Y}^B(Z) = (\mathbf{ero}^B(Z \cup X^c)) \cap X$$
$$\mathbf{dil}_X^B(Z) = \mathbf{dil}^B(Z) \cap X$$

Exemplo 5.32

[illegible]

A partir do subconjunto Z foi possível, usando o elemento estruturante $3B$, pelo processo de dilatação condicional, recuperar o subconjunto Y da imagem X . Porém no exemplo, o resultado não foi totalmente satisfatório porque uma parte da letra A apareceu na imagem. Isto aconteceu porque o sub-conjunto A está "muito perto" do sub-conjunto Y . Essa noção de proximidade está não somente ligada às posições relativas dos sub-conjuntos mas também ao aspecto de forma e de tamanho dos sub-conjuntos envolvidos e principalmente à forma e ao tamanho do elemento estruturante.

Veja como é possível reverter a situação usando o mesmo elemento estruturante nB porém reformulando um pouco a dilatação condicional. Sabe-se que a dilatação condicional por nB pode ser decomposta em n dilatações condicionais por B . Portanto processando a partir do resultado da dilatação condicional por $1B$ por B , tem-se :

$$\begin{aligned}
 \text{dil}_{cX}^{2B}(Z) &= \text{dil}_{cX}^{1B}(\text{dil}_{cX}^{1B}(Z)) = \text{dil}^B(\text{dil}_{cX}^{1B}(Z)) \cap X \\
 &= \left(\left[\begin{array}{ccccccc} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right] \text{dil} \left\{ \begin{array}{ccc} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{array} \right\} \right) \cap \left[\begin{array}{ccccccc} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right] \\
 &= \left[\begin{array}{ccccccc} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right] \cap \left[\begin{array}{ccccccc} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right] = \left[\begin{array}{ccccccc} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right]
 \end{aligned}$$

Dilatando condicionalmente o resultado anterior por B , tem-se :

$$\begin{aligned}
 \text{dil}_{cX}^{3B}(Z) &= \text{dil}_{cX}^{1B}(\text{dil}_{cX}^{2B}(Z)) = \text{dil}^B(\text{dil}_{cX}^{2B}(Z)) \cap X \\
 &= \left(\left[\begin{array}{ccccccc} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right] \text{dil} \left\{ \begin{array}{ccc} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{array} \right\} \right) \cap \left[\begin{array}{ccccccc} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right] = \left[\begin{array}{ccccccc} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right]
 \end{aligned}$$

Por decomposição da dilatação condicional por nB em n dilatações condicionais por B , obteve-se o resultado desejado. Pelas propriedades já encontradas anteriormente, sabe-se que esse tipo de decomposição é muito interessante em termos de rapidez e de flexibilidade. Através desse exemplo, foi ilustrado uma outra vantagem, a de reconstruir um subconjunto sem perigo de aparição de partes indesejáveis de outros sub-conjuntos.

5.5.5 Reconstrução binária: teoria

Tendo um subconjunto S da imagem binária X , imagine-se ter um "pedacinho" Z de S (para mais informações, consultar a parte sobre a Última Erosão no parágrafo ??, página ??). Pode-se dizer que S é marcado por Z . Z chama-se portanto **marcador** de S . E S de **máscara**. Imagine-se agora recuperar o conjunto S a partir desse marcador Z . Esse processo chama-se *reconstrução*. Uma primeira maneira de definir a reconstrução de S a partir de Z é dizer que [Vin93]:



a) Imagem original

b) Imagem de marcadores

c) Reconstrução binária 10 iterações d) Reconstrução binária 20 iterações

e) Reconstrução binária 30 iterações f) Reconstrução binária 50 iterações

figura 5.18: Exemplo de reconstrução binária por dilatação condicional

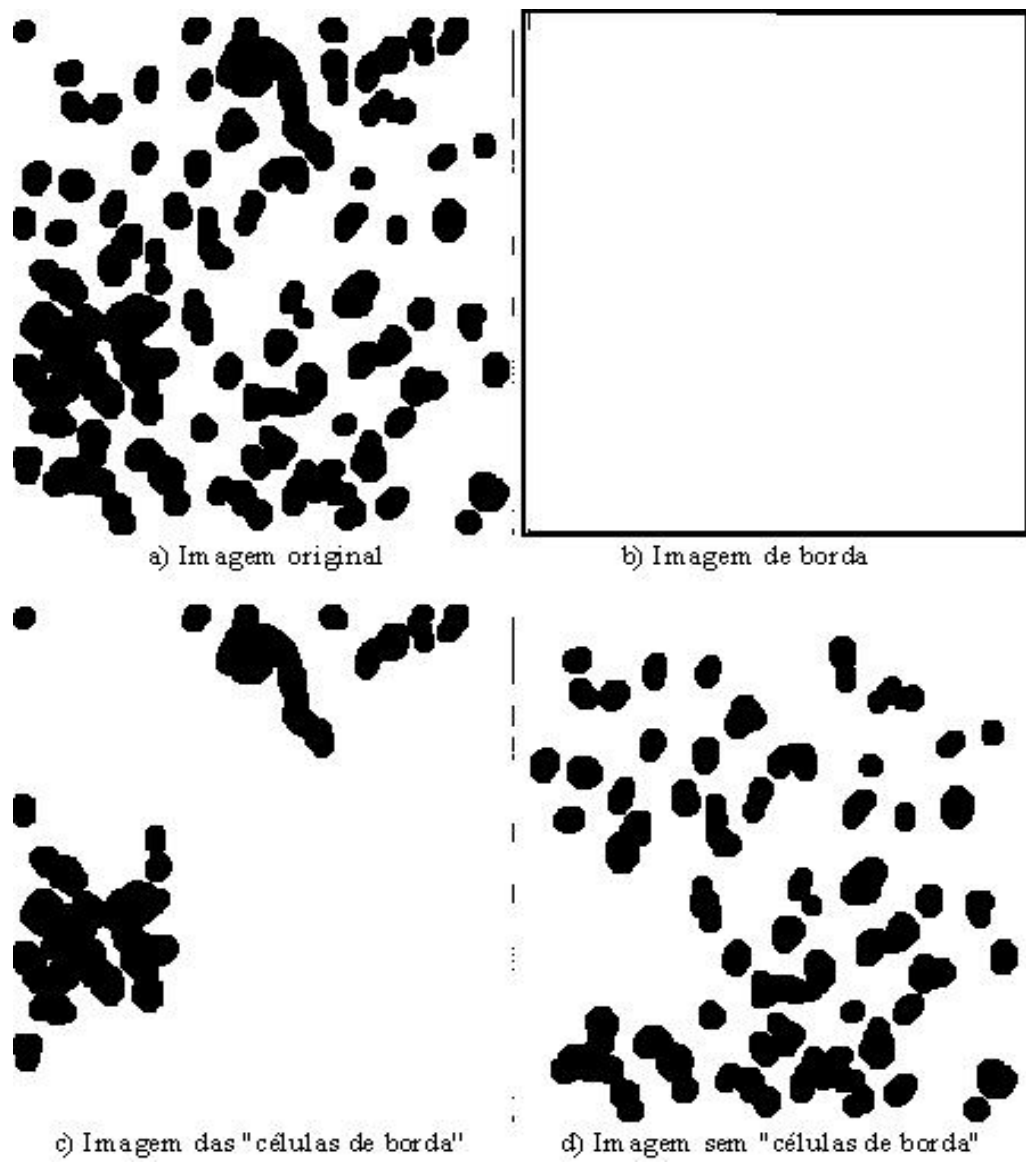


figura 5.19: Eliminação de conjuntos por reconstrução binária

[illegible][illegible]

Na figura 5.19, pode-se verificar os resultados do processo de eliminação de conjuntos tocando as bordas por reconstrução binária. A imagem 5.19-a) representa a imagem original, a imagem 5.19-b) a imagem de borda, a imagem 5.19-c) a reconstrução dos conjuntos tocando as bordas. Pode-se verificar que os conjuntos que realmente tocam as bordas são reconstruídos enquanto que os que estão próximos às bordas não aparecem. A imagem 5.19-d) representa a imagem final sem esses conjuntos.

2- Efetuar a erosão de X por $B1$:

$$X = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \quad e \quad B1 = \left\{ \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \right\}$$

Efetuar a erosão de X por $B2$:

$$X = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \quad e \quad B2 = \left\{ \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \right\}$$

O que pôde-se concluir sobre os resultados finais de 1- e 2-?

Exercício 5.9 *Propor um elemento estruturante B que, por erosão, transforme o conjunto X no conjunto Y :*

$$X = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \quad e \quad Y = \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}$$

Exercício 5.10 *Efetuar a dilatação de X por B usando a 1a definição:*

$$X = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \quad e \quad B = \left\{ \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \right\}$$

Exercício 5.11 *Efetuar a dilatação de X por B usando a 2a definição:*

$$X = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \quad e \quad B = \left\{ \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \right\}$$

Exercício 5.12 *Efetuar a dilatação de X por B usando a 1a definição:*

$$X = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \quad e \quad B = \left\{ \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \right\}$$

Exercício 5.13 *Efetuar a dilatação de X por B usando a 2a definição:*

$$X = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \quad e \quad B = \left\{ \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \right\}$$

Exercício 5.14 *Propor um elemento estruturante B que preencha o furo no conjunto X :*

$$X = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$$

Exercício 5.15 *Decompor o elemento estruturante Rhombus B em elementos estruturantes menores:*

$$B = \left\{ \begin{array}{c} \cdot \cdot \cdot \cdot \cdot \\ \cdot \cdot \cdot \cdot \cdot \\ \cdot \cdot \cdot \cdot \cdot \\ \cdot \cdot \cdot \cdot \cdot \\ \cdot \cdot \cdot \cdot \cdot \end{array} \right\}$$

Bibliografia

- [Abu89] A.S. Abutaleb. Automatic thresholding of gray-level pictures using two-dimensional entropy. *Computer Graphics and Image Processing*, 47(1):22–32, july 1989.
- [Bal81] D. H. Ballard. Generalizing the hough transform to detect arbitrary shapes. *Pattern Recognition*, 13(2):111–122, 1981.
- [BB82] D.H. Ballard and C.M. Brown. *Computer Vision*. Prentice-Hall, 1982.
- [BB94] G.J.F Banon and J. Barrera. *Bases da Morfologia Matemática para a Análise de Imagens Binárias*. IX Escola de Computação, Recife, 1994.
- [Ber86] J. Bernsen. Dynamic thresholding of gray-level images. *Proc. Eighth Int'l Conf. Conf. on Pattern Recognition, Paris, France*, pages 1251–1255, 1986.
- [CA84] C.H. Chien and J.K. Aggarwal. A normalized quadtree representation. *Computer Vision, Graphics and Image Processing*, 26:331–346, 1984.
- [CDG92] B. Costa, L. Darsa, and J.M. Gomes. Image metamorphosis. *V Simpósio Brasileiro de Computação Gráfica e Processamento de Imagens*, pages 19–27, 1992.
- [Cia87] S. Ciarcia. Build a gray-scale video digitizer. *Byte*, May/June 1987.
- [CMVM86] F. Cheevasuvit, H. Maitre, and D. Vidal-Madjar. A robust method for picture segmentation based on a split and merge procedure. *Computer Vision, Graphics and Image Processing*, 34:268–281, 1986.
- [CW94] E. Casey and S. Wong. Adaptative thresholding methodology for images based on intensity gradient information. *Computer Vision, Graphics and Image Processing*, 51:238–247, 1994.
- [DNBC87] L. A. Vieira Dias, G. Camara Neto., J. Barrera, and M. R. Cartaxo. Imagens e computadores - o olho que tudo vê. *Ciência Hoje*, 7(37):34–44, Novembro 1987.
- [Dou81] J.G. Douglas. The role of human visual models in image processing. *Proceedings of the IEEE*, 69(5):552–561, May 1981.
- [Duf86] M. J. B. Duff. *Intermediate-level Image Processing*. Academic Press, 1986.
- [FF87] M.A. Fischler and O. Firschein. *Intelligence, the eye, the brain and the computer*. Addison Wesley, 1987.
- [Fis84] R. Fish. Electronics in medical imaging. *Radio Electronics*, Oct./Nov. 1984.

- [Fre61] H. Freeman. On the encoding of arbitrary geometric configurations'. *IRE Trans. on Electronic Computer*, 1961.
- [GW87] R.C. Gonzales and P. Winz. *Digital Image Processing. 2.a Edição*. Addison-Wesley Publishing Company, 1987.
- [Had50] H. Hadwiger. Minkowskische addition und subtraktion beliebiger punktmengen und die theoreme von erhard schmidt. *Math.Zeitschrift*, 53(3):210–218, 1950.
- [Had57] H. Hadwiger. *Vorlesungen uber inhalt, oberfläche und isoperimetrie*. Springer Verlag, Berlin, 1957.
- [HD88] Q.Z Haralick and P.E Danielson. On minimum error thresholding and its implementations. *Pattern Recognition Letters*, 7(4):201–206, april 1988.
- [HEY82] G. HEYGSTER. Rank filters in digital image processing. *Computer Graphics and Image Processing*, 19:148–164, 1982.
- [JB82] G. Johannsen and J. Bille. A threshold selection method using information measures. *Proceedings, 6th Int. Conf. Pattern Recognition, Munich, Germany*, pages 140–143, 1982.
- [Kat65] Y.H. Katz. Pattern recognition od meteorological satellite cloud photography. *Proc. 3rd Symp. Remote Sensing Environment*, pages 173–214, 1965.
- [KZ93] G.H. Kamel and J. Zhao. *Analyzing Oriented Patterns*. IEEE Computer Society Press, 1993.
- [LM84] C. Lantuéjoul and F. Maisonneuve. Geodesic methods in quantitative image analysis. *Pattern Recognition*, 17:177–187, 1984.
- [Mar82] D. Marr. *Vision*. Freeman, 1982.
- [Mas75] D. Mason. Measurement of c-bands in human chromosomes. *Computers Biology and Medicine*, 5:179–201, 1975.
- [MG89] C.H. Mitchell and L. Gillies. *Fundamentals of Digital Image Processing*. Prentice-Hall Information and System Sciencies Series, Englewood Cliffs, NJ, 1989.
- [Min03] H. Minkowski. Volumem und oberfläche. *Math. Annalen*, 57:447–495, 1903.
- [Mor87a] T. Morand. *Repartir: une méthode pour l'interprétation d'images robotiques en vision 2D*. Université de Technologie de Compiègne, França, 1987.
- [Mor87b] H.P. Moravec. Towards automatic visual obstacle avoidance. *Proc. 5th Int. Joint. Conf. Artificial Intelligence, Cambridge, MA*, page 584, august 1987.
- [MV89] N.D.A. Mascarenhas and F.D.R. Velasco Velasco. *Processamento Digital de Imagens. 2a edição*. IV Escola Brasileiro-Argentina de Informática, Termas de Rio Hondo, Argentina, 1989.
- [NM79] M. Nagao and T. Matsuyama. Edge preserving smoothing. *Computer Graphics and Image Processing*, 13:394–407, 1979.

- [NR78] Y. NAKAGAWA and A. ROSENFELD. A note on the use of local min and max operations in digital picture processing. *IEEE Transactions on Systems, Man and Cybernetics*, 8, 1978.
- [Pug83] A. Pugh. *Robot Vision*. IFS Publication, UK, Springer-Verlag, 1983.
- [Pun81] T. Pun. Entropic thresholding, the new approach. *Computer Graphics and Image Processing*, 16:210–239, 1981.
- [Ros88] A. Rosenfeld. Computer vision: Basic principles. *Proceedings of the IEEE*, 76(8):863–868, August 1988.
- [RRK84] S.S. Reddi, S.F. Rudin, and H.R. Keshavan. An optimal multiple threshold scheme for image segmentation. *IEEE Trans. Systems, Man and Cybernetics*, CMC-14(4):661–665, july-august 1984.
- [RT83] A. Rosenfeld and P. De La Torre. Histogram concavity analysis as an aid in threshold selection. *IEEE Trans. Systems, Man and Cybernetics*, CMC-13:231–235, 1983.
- [Ser82] J. Serra. *Image Analysis and Mathematical Morphology*. Academic Press, London, 1982.
- [Shi87] Y. Shirai. *Three Dimensional Computer Vision*. Springer-Verlag, 1987.
- [SSW88] P.K. Sahoo, S. Soltani, and A.K.C. Wong. A survey of thresholding techniques. *Computer Graphics and Image Processing*, 41:233–260, 1988.
- [Vin93] L. Vincent. Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms. *IEEE Transactions on Image Processing*, 2(2):176–201, April 1993.
- [Wat74] S. Watanabe. A automated apparatus for cancer processing cybest. *Computer Graphics and Image Processing*, 3(4):350–358, december 1974.
- [WNR74] J.S. Weszka, R.N. Nagel, and A. Rosenfeld. A threshold selection technique. *IEEE Trans. Computer*, C-23:1322–1326, 1974.
- [Won78] K. Wong. *Image Segmentation by texture*. IEEE Computer Society Press, 1978.
- [WR83] J.M. White and G.D. Rohrer. Image thresholding for optical character recognition and other applications requiring character image extraction. *IBM J. Research and Development*, 27(4):400–411, july 1983.

Índice

- 4-vizinhança, 12
- 8-vizinhança, 12
- abertura
 - binária, 89
- anti-dilatação, 72
- anti-erosão, 72
- aquisição, 7
- background, 46
- binarização, 46
- borda
 - detecção, 87
- conectividade, 12
- contorno, 45
- detecção
 - de borda, 57
 - de círculo, 61
 - de contorno, 58
 - de reta, 60
- diferenciação, 39
- digitalização, 4, 7, 10
- dilatação, 72
 - binária, 77
 - condicional, 101
- divisão e fusão, 46, 51
- elemento estruturante, 67, 71
 - cruz, 86
 - ponto central, 86
 - quadrado, 86
 - rhombus, 86
 - transposto, 72
- entropia, 55
- equalização
 - de histograma, 22
- erosão, 72
 - binária, 72
 - condicional, 101
- especificação
 - de histograma, 25
- espectro de Fourier, 56
- fechamento
 - binário, 93
- filtragem, 28
- filtro
 - da média, 33
 - da mediana, 35
 - de Nagao-Matsuyama, 38
 - homomórfico, 38
 - não linear, 34
 - passa-alta, 33
 - passa-baixa, 32
 - por convolução, 30
- foreground, 46
- Fourier
 - transformada de, 29
- Freeman, 59
- gradiente, 40, 57
 - argumento, 40
 - módulo, 40
- grafo, 64
- histograma, 21, 46, 48
 - equalização de, 22
 - especificação de, 25
 - modificação de, 20, 22
- Hough
 - transformada de, 60
- iluminância, 9, 38
- iluminação, 48
- imagem, 9
 - digital, 10
 - monocromática, 9
- inspeção visual, 5
- interpretação, 8
- Laplaciano, 42, 49, 57
- limiarização, 46, 58
 - adaptativa, 48
 - global, 48
 - local, 48, 50
 - local adaptativa, 50
- máximo de probabilidade, 55
- média de múltiplas imagens, 27

- matriz de co-ocorrência, 54
- Minkowski
 - adição, 69
 - subtração, 69
- modificação
 - de contraste, 20, 26
 - de histograma, 20, 22
- nível de cinza, 10
- OCR, 5
- operador
 - de Moravec, 57
 - de Prewitt, 41
 - de Sobel, 42
- perseguição de contornos, 59
- pixel, 10
- ponto central, 86
- pré-processamento, 8
- primitiva, 45
- quadtree, 51
- realce, 19, 39
- reconhecimento
 - de caracteres, 5
 - de assinaturas, 6
 - de impressões digitais, 6
 - de padrões, 5
- reconstrução
 - binária, 104
 - tridimensional, 6
- refletância, 9, 38, 46
- região, 45
- representação, 8
- restauração, 19
- robô, 5
- rotação, 14
- segmentação, 8, 45
 - por contorno, 56
 - por região, 45
 - por textura, 52
- segmento, 45
- sensor, 4
- Shannon, 10
- simetria, 68
- sistema
 - visual humano, 1
- split and merge, 46
- suavização, 20
- threshold, 46
- transformação
 - de escala, 14
 - anti-extensiva, 98
 - crescente, 98
 - extensiva, 98
 - idempotente, 98
 - perspectiva, 15
- transformada
 - de Fourier, 29
 - de Hough, 60
- translação, 13, 68
- transmitância, 9
- uniformidade, 55