

7

REPRESENTAÇÃO E DESCRIÇÃO

Um dos problemas básicos no desenvolvimento de um sistema para análise de imagens é a seleção de um conjunto de características extraídas do objeto de interesse para o propósito de classificação.

Após a imagem ter sido segmentada em regiões ou objetos, os agrupamentos resultantes de pixels são normalmente representados e descritos em formato apropriado para o processamento subsequente.

Um objeto pode ser representado em termos de suas características externas (bordas) ou características internas (pixels que compõem o objeto). A descrição depende da representação adotada e deve permitir a caracterização de sua forma sem ambigüidades e com um número reduzido de características ou medidas extraídas do objeto.

A forma do objeto é uma informação visual importante e tem sido utilizada como uma das características básicas para descrever o conteúdo de imagens. Entretanto, em muitos problemas práticos, a forma dos objetos pode sofrer distorções devido à presença de ruído, oclusão com outros objetos e outros tipos de degradações.

Diversos esquemas de representação e descrição de objetos têm sido propostos. Um aspecto desejável nos descritores de forma é que eles deveriam ser pouco afetados na presença de ruído e por transformações geométricas como mudança de escala, rotação e translação.

As próximas seções apresentam um conjunto de técnicas para representação e descrição de objetos segmentados das imagens. As abordagens são classificadas de acordo com as características extraídas dos contornos ou das próprias regiões que formam os objetos. Outra classificação possível seria considerar o domínio no qual as características são derivadas, tanto no domínio espacial quanto de freqüência.

7.1 Esquemas de Representação

As técnicas de segmentação descritas no capítulo 5 resultam em agrupamentos de pixels que correspondem às regiões homogêneas da imagem ou aos contornos (bordas) dos objetos. Estruturas adequadas de representação facilitam o armazenamento e a manipulação dos objetos segmentados da imagem, além de simplificarem o cálculo de certos descritores de forma. Esta seção discute alguns esquemas de representação de objetos.

7.1.1 Código da Cadeia

Freeman (1961, 1974) propôs um esquema conhecido como *código da cadeia* para representar a borda de um objeto. Ao invés de armazenar as coordenadas absolutas dos pixels, a representação pelo código da cadeia utiliza a posição relativa entre pixels consecutivos da borda. Dado um ponto inicial pertencente à borda, o código é definido por uma sequência formada pelas direções entre cada pixel e seu vizinho, até que todos os pixels da borda sejam considerados. A direção de cada segmento é codificada por um esquema de numeração baseado na vizinhança-4 ou vizinhança-8, como ilustrado na figura 7.1.

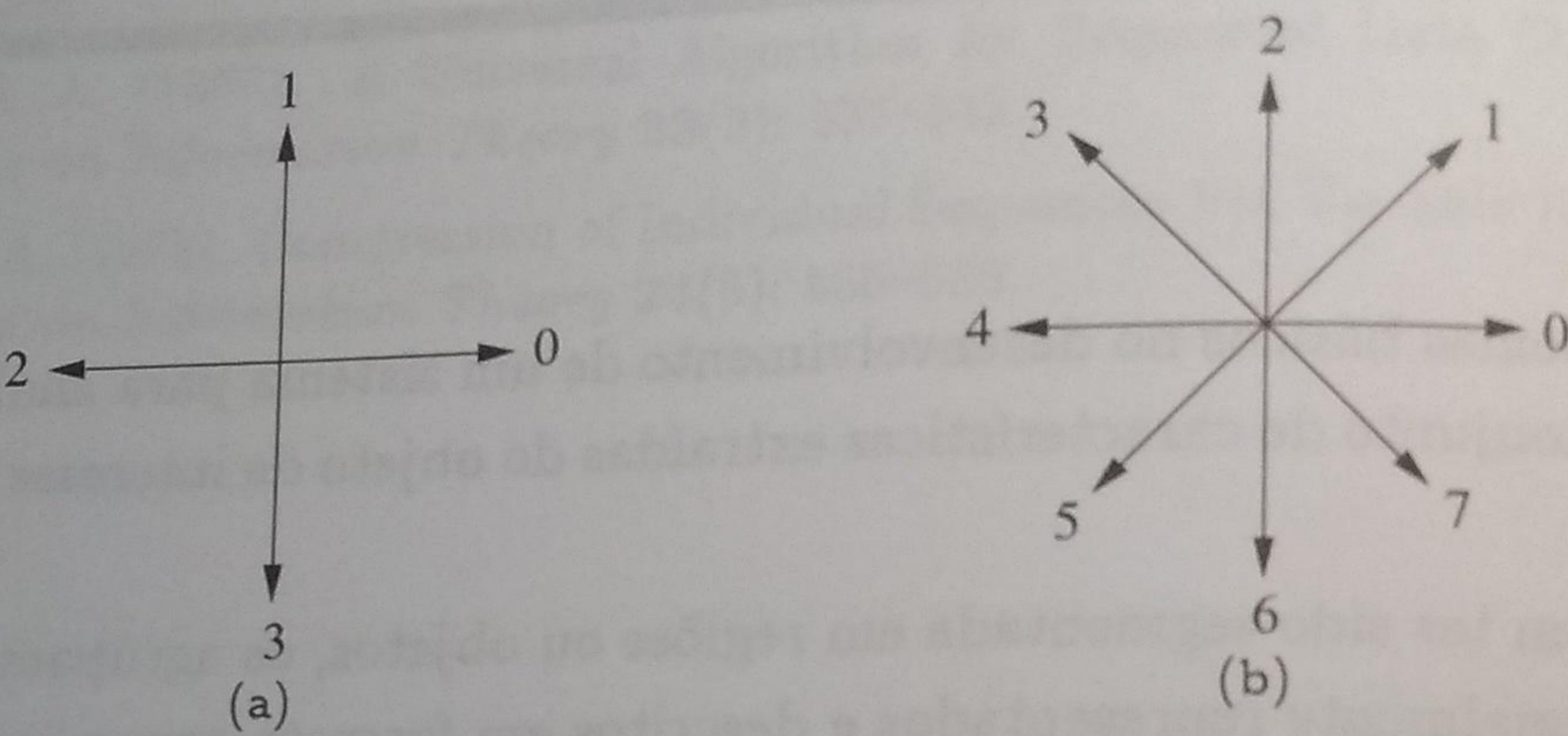


Figura 7.1: Direções do código da cadeia para (a) vizinhança-4 e (b) vizinhança-8.

Essa representação baseia-se inicialmente na amostragem da borda a ser codificada por meio de uma grade cujo espaçamento determina a resolução da codificação. A figura 7.2(a) mostra um contorno e uma grade de amostragem. Os pontos marcados são os vértices da grade que estão mais próximos da curva e que serão usados na codificação. A figura 7.2(b) mostra o resultado da amostragem da borda com a grade.

O código da cadeia é gerado seguindo-se a borda do objeto, por exemplo, no sentido anti-horário e atribuindo-se uma direção aos segmentos que conectam cada par de pixels, de acordo com a vizinhança-4 ou vizinhança-8. Assumindo-se que a posição inicial seja o ponto (3,6), o código da cadeia referente à curva da figura 7.2(b) é dado pela seqüência 1170012344443556670.

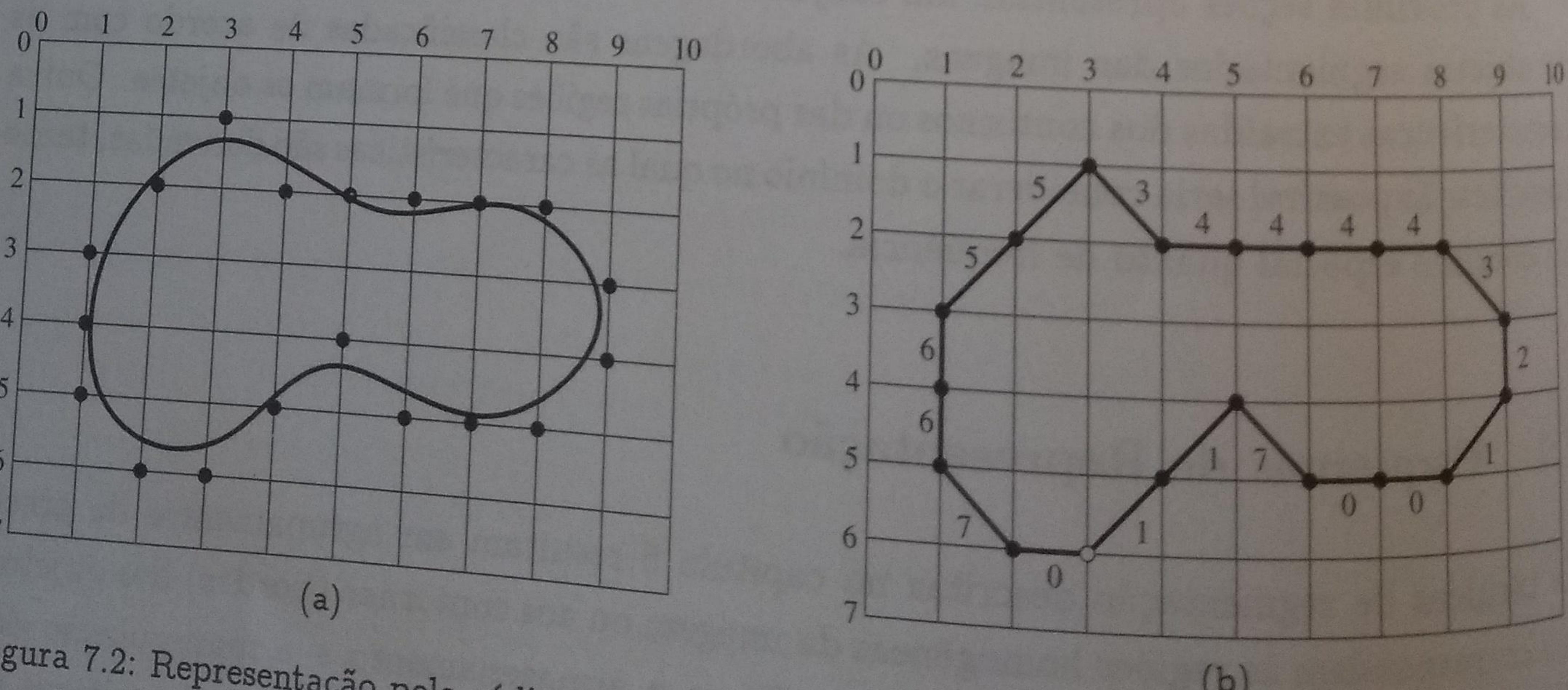


Figura 7.2: Representação pelo código da cadeia. (a) borda original com pontos amostrados sobre a grade; (b) reamostragem da borda, com códigos direcionais.

Na forma descrita, o código da cadeia é dependente do ponto inicial da borda. Para torná-lo invariante com respeito à escolha do primeiro pixel da borda, um procedimento de normalização consiste em encontrar o pixel que resulte no menor número inteiro, considerando a seqüência de dígitos como um número na base 4 ou 8, de acordo com o conceito de vizinhança adotado. Esse pixel é então utilizado como o ponto inicial (Tsai e Yu 1985).

Esse processo é ilustrado na figura 7.3, em que o código da cadeia é obtido a partir do contorno do objeto da figura 7.2(b), tendo como ponto inicial o pixel (3,6). O resultado do primeiro deslocamento circular de dígitos, em que o dígito menos significativo passa a ser o dígito mais significativo, é ilustrado na figura 7.3(b). Isso é equivalente ao código formado utilizando-se o pixel (2,6) como ponto inicial. O resultado do segundo deslocamento, mostrado na figura 7.3(c), é equivalente ao código formado com o pixel (1,5) como ponto inicial. O código cujo número inteiro é mínimo considerando-se todos os deslocamentos é mostrado na figura 7.3(d), o qual é derivado a partir do pixel (7,2).

$$C = 117001234443556670$$

(a)

$$C = 011700123444355667$$

(b)

$$C = 701170012344435566$$

(c)

$$C = 001234443556670117$$

(d)

Figura 7.3: Invariância do código da cadeia com respeito à translação. (a) código da cadeia inicial; (b) resultado após primeiro deslocamento; (c) resultado após segundo deslocamento; (d) código da cadeia formado pelo menor número inteiro.

Para tornar o código da cadeia invariante quanto à mudança de escala, a borda pode ser reamostrada em uma grade de espaçamento conveniente. Com a subdivisão da borda, os comprimentos dos segmentos dos códigos devem-se ajustar às subdivisões.

A normalização pode também ser feita para tornar o código da cadeia invariante com respeito à rotação, utilizando-se a primeira diferença do código da cadeia. A diferença é obtida contando-se, no sentido anti-horário, o número de direções que separam cada elemento adjacente ao código. Seja $a_0 a_1 \dots a_{n-1}$ o código da cadeia que representa a borda de um objeto. A primeira diferença do código da cadeia é dada por $b_0 b_1 \dots b_{n-1}$, em que $b_k = a_k - a_{k-1}$. A diferença entre os elementos adjacentes é realizada de forma circular, ou seja, módulo 4 ou 8, conforme a vizinhança adotada para a adjacência dos pixels.

Exemplo 7.1 Seja a borda B_1 de um objeto mostrado na figura 7.4(a). Considerando o ponto inicial como (4,5), o código da cadeia de vizinhança-4 correspondente à borda B_1 é $C_1 = 11112232333000$. Seja B_2 a borda do mesmo objeto, porém, após ter sofrido uma rotação de 90° no sentido horário. Considerando o ponto inicial como (5,2), o código da cadeia associado à borda B_2 é $C_2 = 22223303000111$. Uma observação interessante é que a diferença (módulo 4) entre cada par de dígitos correspondentes para os códigos C_1 e C_2 é sempre igual a 1, ou seja

$$C_2 = 22223303000111$$

$$C_1 = 11112232333000$$

$$C_2 - C_1 = \underline{11111111111111}$$

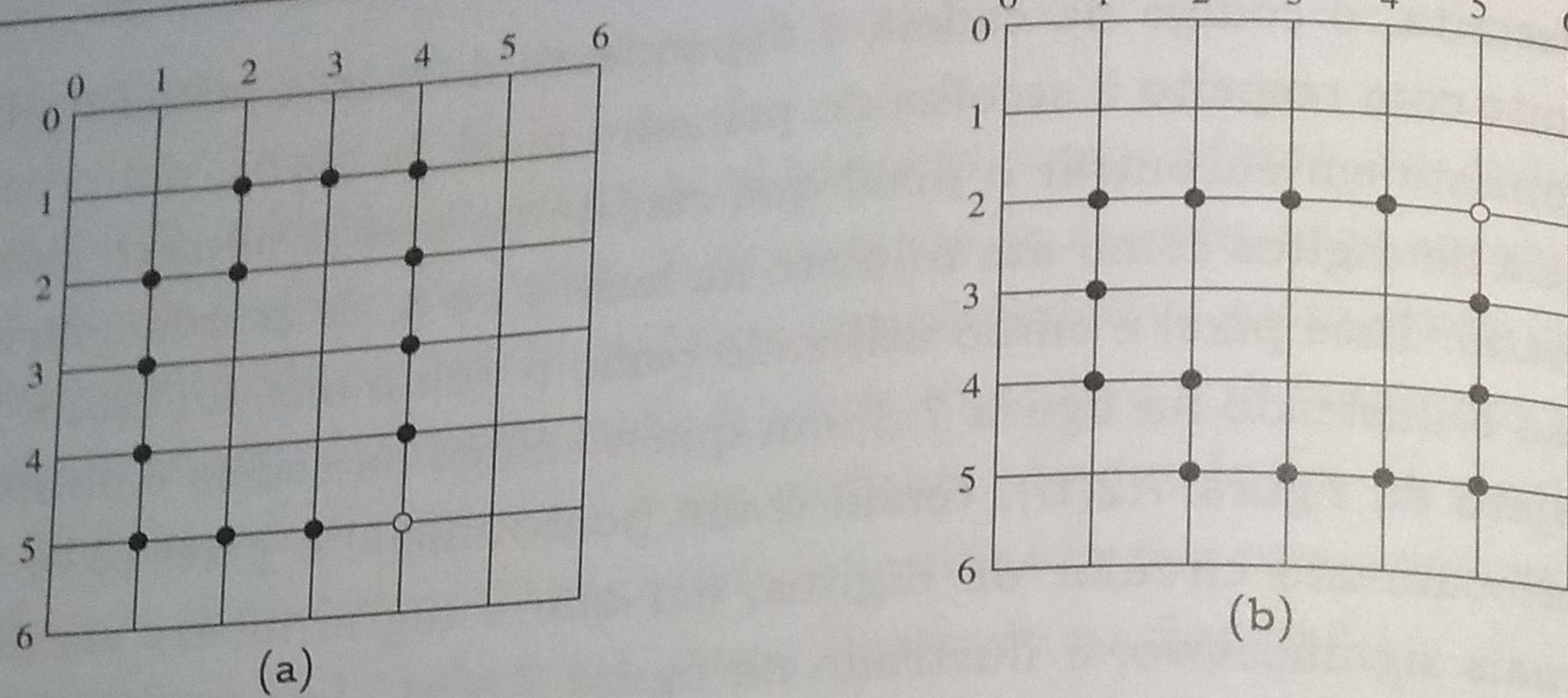


Figura 7.4: Normalização do código da cadeia com respeito à rotação. (a) borda original do objeto; (b) borda após rotação do objeto em 90° .

Após a rotação da borda, cada dígito do código da cadeia é alterado pelo mesmo valor, dependendo do grau de rotação, o que significa que dígitos adjacentes em um código da cadeia sempre terão a mesma diferença, mesmo após a rotação. Portanto, os códigos da cadeia representam a mesma borda de um objeto se pares de dígitos nos códigos possuem a mesma diferença. Assim, para a borda B_2 do objeto após a rotação, cujo código da cadeia é 22223303000111, o código da primeira diferença é 0001013100100. Analogamente, para a borda B_1 do objeto original, cujo código da cadeia é 11112232333000, o código da primeira diferença também é 0001013100100. Portanto, esses dois códigos da cadeia representam a mesma borda, uma vez que seus códigos de primeira diferença são iguais.

Além da redução do espaço de armazenamento, o código da cadeia pode ser aplicado em várias tarefas de processamento e análise de imagens. Freeman (1974) propôs algoritmos para cálculo de perímetro e área baseado no código da cadeia para contornos fechados. A detecção de cantos ou junções em contornos de objetos, descrita na seção 5.1.4, é outra operação que pode ser realizada com o uso do código da cadeia. Kuhl e Giardina (1982) desenvolveram um procedimento para determinar os coeficientes de Fourier a partir do código da cadeia.

Uma desvantagem do código da cadeia é sua sensibilidade a pequenas perturbações ao longo da borda, devido a ruído ou a imperfeições durante o processo de segmentação. Li e Zhiying (1988) propuseram uma versão suavizada do código da cadeia, em que médias das direções são calculadas ao longo do comprimento do caminho.

7.1.2 Aproximações Poligonais

A borda de um objeto pode ser aproximada por um polígono, segundo uma precisão arbitrária. No caso de um contorno fechado, a aproximação é exata quando o número de segmentos do polígono for igual ao número de pontos na borda, de maneira que cada par de pontos adjacentes defina um segmento do polígono.

O principal objetivo da aproximação poligonal é capturar a essência da forma da borda com o menor número possível de segmentos poligonais. Esse problema não é, em geral, trivial, podendo se tornar uma busca iterativa bastante custosa. No entanto, várias técnicas de aproximação poligonal, com complexidade computacional relativamente baixa, podem ser bem adaptadas para aplicações em processamento de imagens.

Representação
Uma solução
se ajustam à
em medidas
poligonal. U
quadrático d
limiar. Quan
o erro é igual
borda até qu
dos segment
Uma das
pondem a i
até que o lin
sendo traç
limiar) apó
Uma té
um segmen
exemplo de
que une se
ponto mai
inicial em
em geral,
figura 7.5.
mais dista
perpendic
a maior d
mostrado
distância

Uma solução possível é realizar a busca de polígonos de perímetro mínimo, os quais se ajustam à geometria estabelecida pela grade da imagem. Técnicas de fusão baseadas em medidas de erro ou outros critérios têm sido aplicadas em problemas de aproximação poligonal. Uma abordagem consiste em unir os pontos ao longo da borda até que o erro quadrático de um segmento de reta ajustado aos pontos unidos exceda um determinado limiar. Quando essa condição ocorrer, os parâmetros do segmento de reta são armazenados, o erro é igualado a zero e o procedimento repetido, juntando-se novos pontos ao longo da borda até que o erro exceda novamente o limiar. No final do procedimento, as intersecções dos segmentos adjacentes formam os vértices do polígono.

Uma das principais dificuldades desse método é que os vértices geralmente não correspondem a inflexões (como cantos) na borda, uma vez que o novo segmento não é iniciado até que o limiar de erro seja excedido. Se, por exemplo, um longo segmento de reta estivesse sendo traçado e ele cruzasse um canto, um determinado número de pontos (dependendo do limiar) após passar pelo canto ainda seria considerado antes que o erro excedesse o limiar.

Uma técnica baseada na divisão de segmentos consiste em subdividir sucessivamente um segmento em duas partes até que um determinado critério de erro seja satisfeita. Um exemplo de critério é que a maior distância perpendicular entre o segmento da borda e a linha que une seus dois pontos extremos não exceda um determinado limiar. Caso isso ocorra, o ponto mais distante no segmento da borda se torna um vértice que subdivide o segmento inicial em dois segmentos. No caso de uma borda fechada, os melhores pontos iniciais são, em geral, os dois pontos mais distantes entre si da borda. Esse processo é ilustrado na figura 7.5. A borda do objeto, mostrada na figura 7.5(a), é subdividida em seus pontos mais distantes, conforme ilustrado na figura 7.5(b). O ponto A possui a maior distância perpendicular entre o segmento superior e a linha BD. De maneira similar, o ponto C possui a maior distância em relação ao segmento inferior. O processo termina com o polígono mostrado na figura 7.5(d), uma vez que nenhum ponto nos segmentos da borda possui uma distância perpendicular ao seu segmento correspondente maior que o limiar definido.

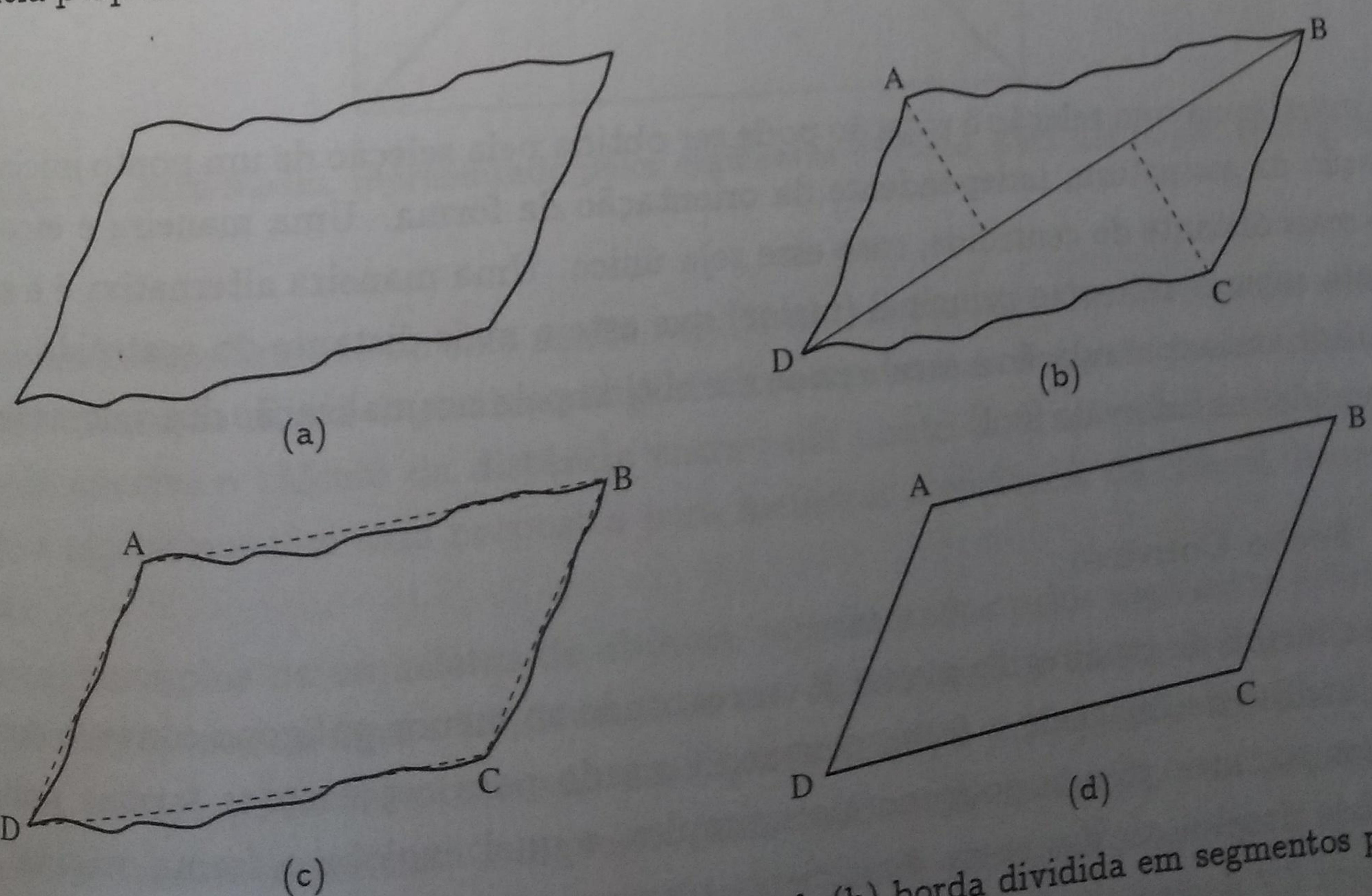


Figura 7.5: Aproximação poligonal. (a) borda original; (b) borda dividida em segmentos por meio de cálculo de distância; (c) união dos vértices; (d) polígono resultante.

7.1.3 Assinatura

Uma assinatura é uma representação unidimensional da borda de um objeto. Uma das formas mais simples é dada pelo gráfico da distância da borda ao centróide em função do ângulo, como ilustrado na figura 7.6. Para essa definição, as assinaturas são invariantes com respeito à translação, mas são dependentes das transformações de rotação e mudanças de escala.

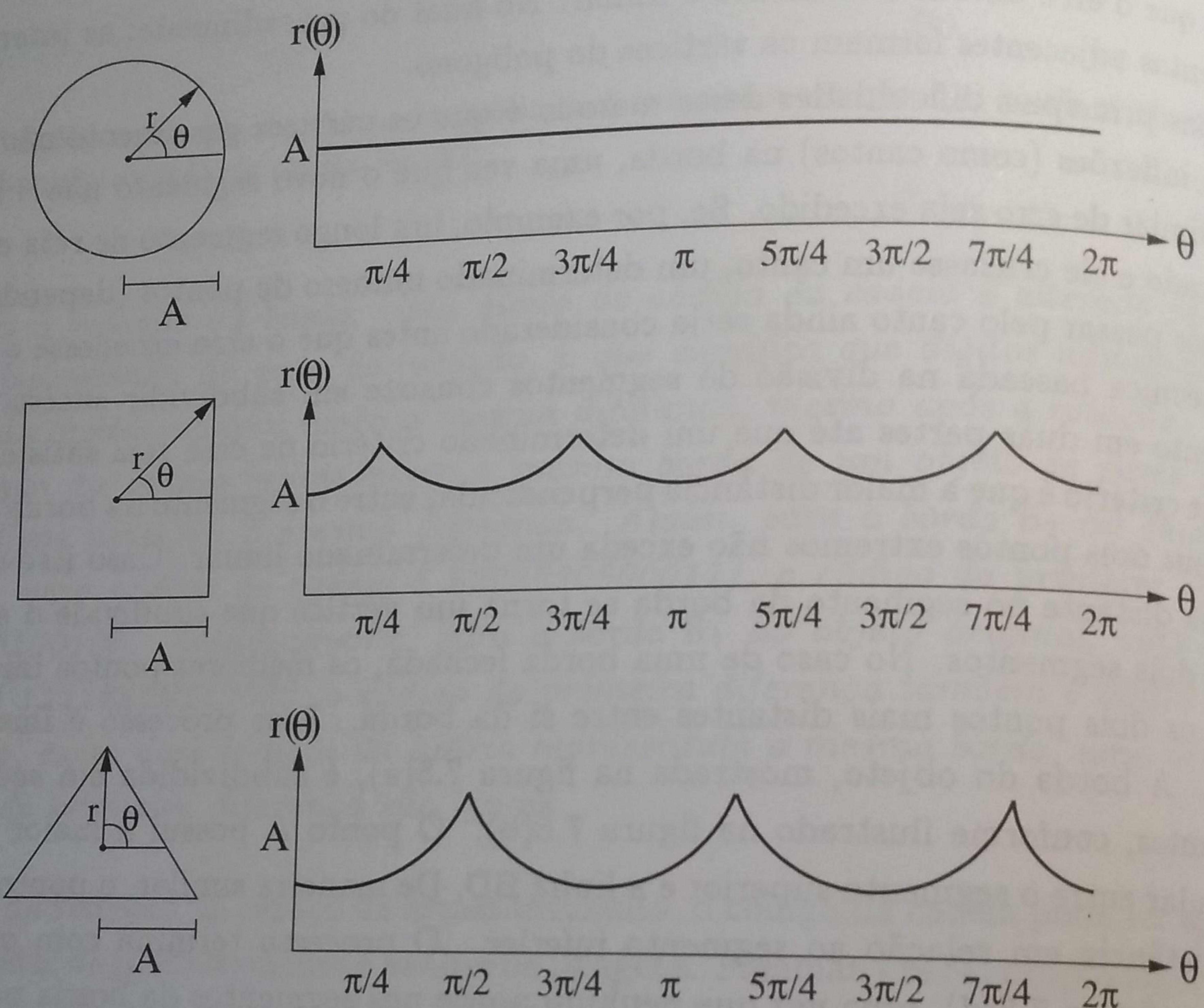


Figura 7.6: Três bordas simples e suas respectivas assinaturas.

A invariância com relação à rotação pode ser obtida pela seleção de um ponto inicial para a geração da assinatura, independente da orientação da forma. Uma maneira é escolher o ponto mais distante do centróide, caso esse seja único. Uma maneira alternativa é a seleção do ponto sobre o auto-eixo principal (maior) que esteja mais distante do centróide.

A invariância com relação à escala pode ser obtida pela normalização dos valores de $r(\theta)$, por exemplo, no intervalo $[0, 1]$.

7.1.4 Fecho Convexo

O fecho convexo de uma região planar R corresponde ao menor polígono convexo contendo R . Em análise de imagens, o fecho convexo é usado para representar formas poligonais complexas por meio de um polígono mais simples, o qual engloba a forma original. Essa forma mais simples, muitas vezes, é suficiente para permitir o reconhecimento do objeto. A figura 7.7 ilustra o fecho convexo, representado pelo polígono em linha tracejada, para um objeto extraído da imagem. Cada um dos ângulos dos vértices internos ao fecho convexo deve, naturalmente, ser menor que 180 graus.

Figura 7.7:

7.1.5 Es...
Uma abor...
estrutura...
por meio...
de Voron...
A tra...
pontos q...
ao esque...
De u...
junto de...
cuja tan...
objeto r...

Fig

Em...
sua im...
menta...
Inúme...
forma...

O...
preto...
os po...
pode...
muit...
local...
A...

Pont

Uma das variantes com mudanças de

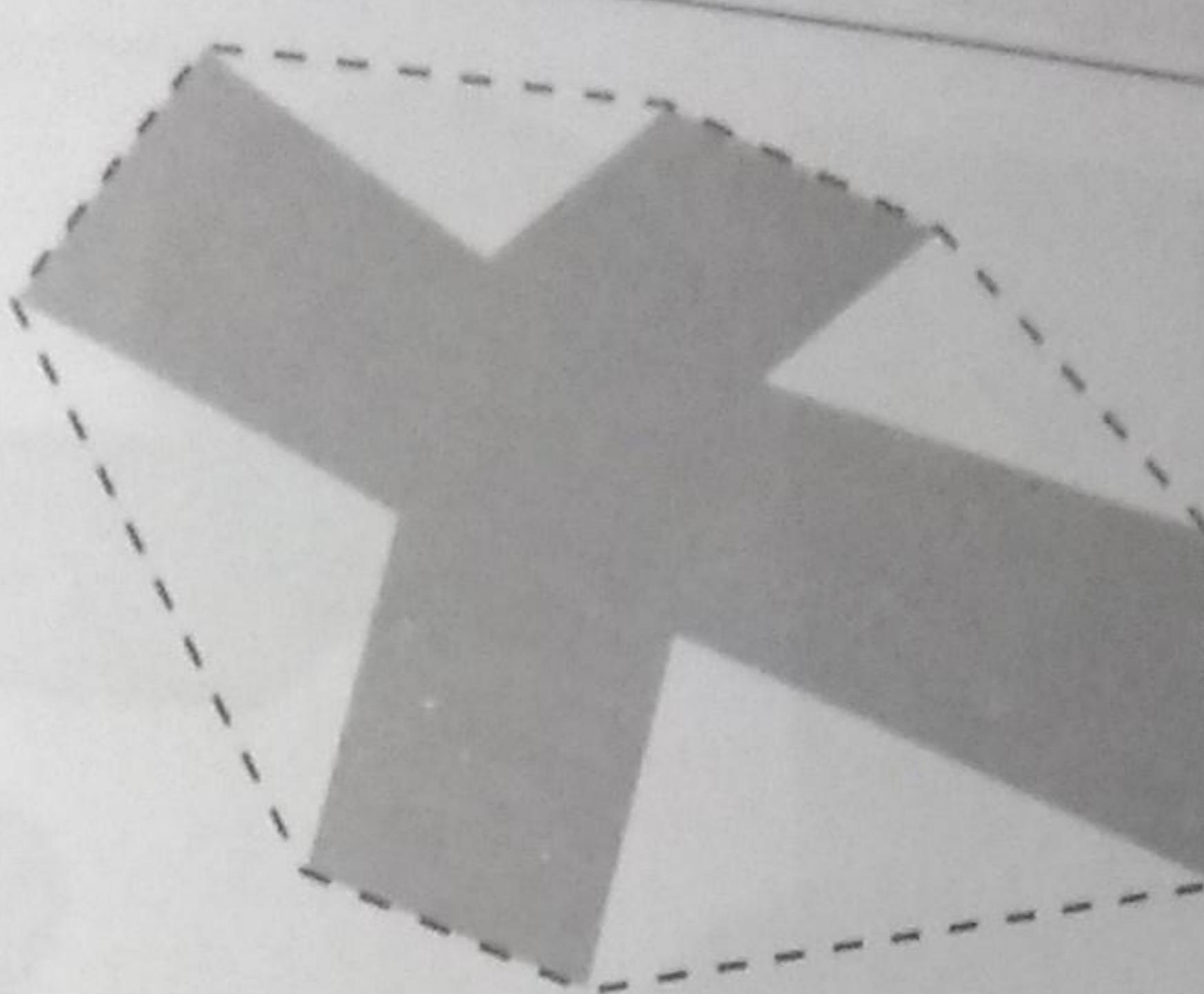


Figura 7.7: Fecho convexo (representado pelo polígono em linhas tracejadas) para uma região planar.

7.1.5 Esqueleto de um Objeto

Uma abordagem importante para a representação da forma de um objeto é reduzi-la a uma estrutura mais simples, chamada de *esqueleto*. O esqueleto de um objeto pode ser obtido por meio de técnicas como transformada do eixo médio, transformada de distância, diagrama de Voronoi e afinamento de objetos.

A *transformada do eixo médio*, proposta por Blum (1967), determina o conjunto de pontos que estão mais próximos da borda do objeto. Um ponto interno ao objeto pertencerá ao esqueleto do objeto se ele tiver, no mínimo, dois pontos mais próximos da borda.

De uma maneira mais formal, o eixo médio de um objeto pode ser definido como o conjunto de pontos formado pelos centros das maiores circunferências que inscrevem o objeto, cuja tangência ocorre em dois ou mais pontos. A figura 7.8 ilustra essa definição para um objeto retangular.

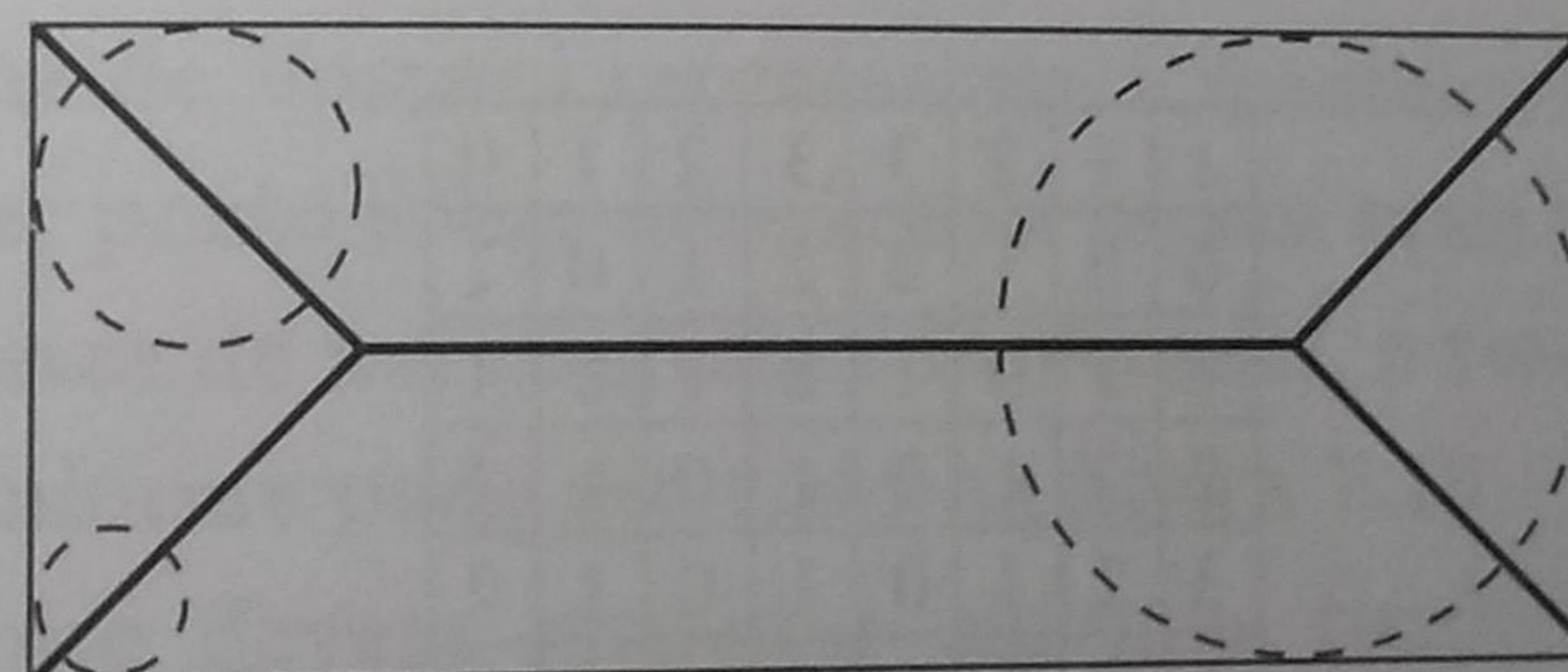


Figura 7.8: Eixo médio, representado pelos segmentos internos, para um objeto retangular.

Embora a transformada do eixo médio resulte em um esqueleto intuitivamente aceitável, sua implementação direta é geralmente proibitiva em termos computacionais. Tal implementação envolve o cálculo da distância entre cada ponto interior e da borda do objeto. Inúmeros algoritmos têm sido propostos para melhorar a eficiência do cálculo dessa transformada.

Outros exemplos de esqueletos de objetos, representados pelos segmentos internos em preto, são mostrados na figura 7.9, levando-se em consideração a distância Euclidiana entre os pontos. Como pode ser observado nas figuras 7.9(a) e 7.9(e), dois objetos diferentes podem ter o mesmo esqueleto. O esqueleto, mostrado na figura 7.9(d), contém um número muito menor de pontos comparado com o objeto original, entretanto, representa as simetrias locais e a estrutura topológica do objeto.

A *transformada de distância* corresponde ao mapa resultante da distância entre cada ponto interior do objeto e de sua borda mais próxima. Essa transformada depende do

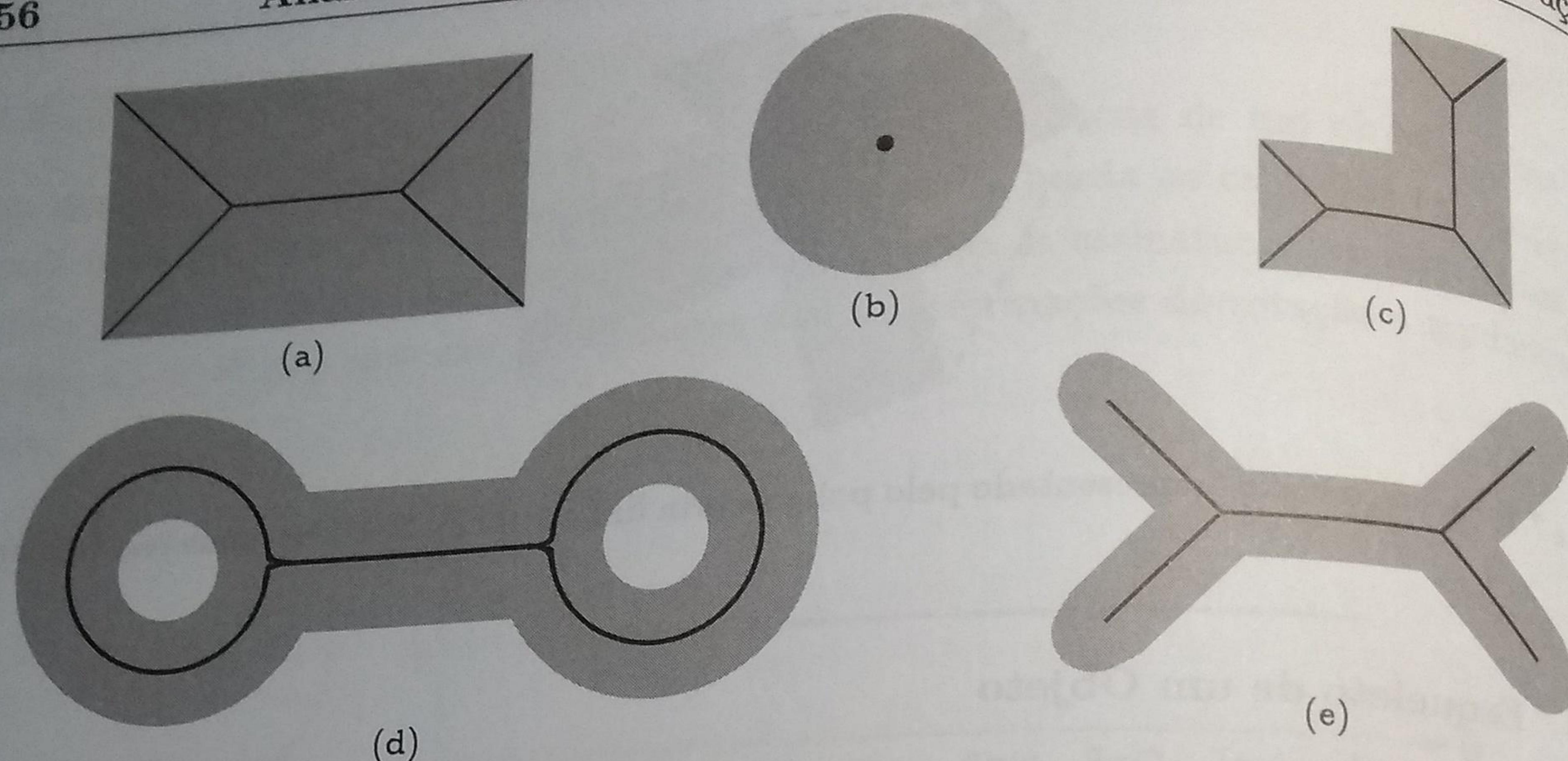


Figura 7.9: Exemplos de objetos e seus respectivos esqueletos.

conceito de distância adotado, conforme discutido na seção 2.11.7. Os valores mais altos formam o esqueleto do objeto.

Um exemplo da aplicação da transformada de distância em uma região da imagem é ilustrado na figura 7.10, sendo adotados dois conceitos para a adjacência entre pixels na imagem, a vizinhança-4 e vizinhança-8. A figura 7.11 mostra um objeto representado por uma imagem binária e o resultado da transformada da distância Euclidiana. O esqueleto do objeto, formado pelos valores mais altos de distância, corresponde às regiões mais claras mostradas na figura 7.11(b).

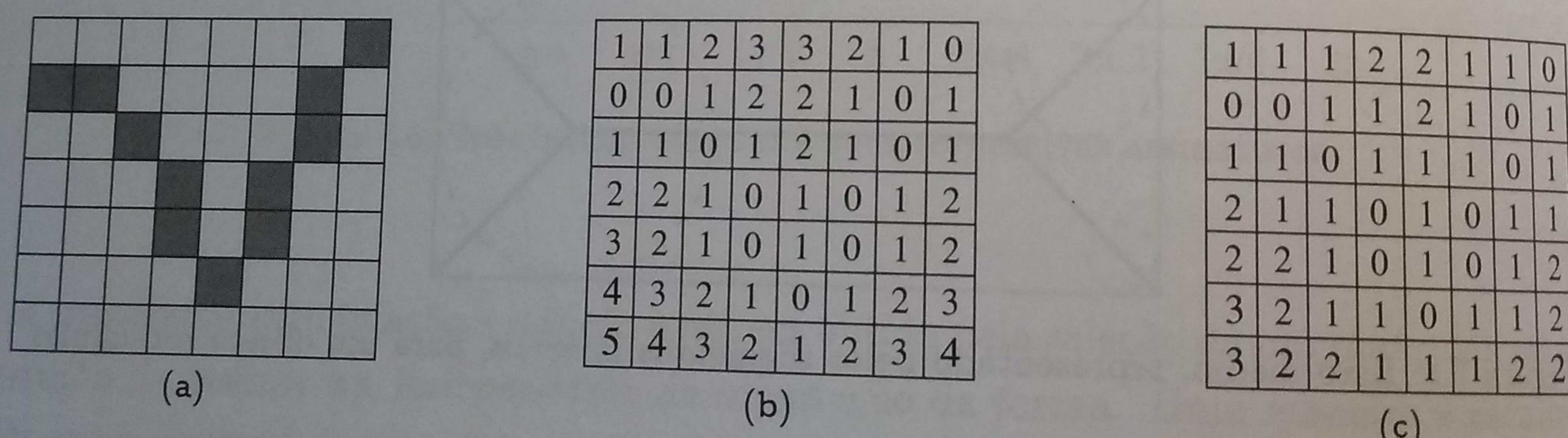


Figura 7.10: Transformada de distância. (a) imagem binária; (b) distâncias para vizinhança-4; (c) distâncias para vizinhança-8.

A transformada de distância pode ser executada em tempo linear, com tempo proporcional ao número de pixels da imagem. Embora essa técnica satisfaça requisitos geométricos, caso o cálculo utilize a distância Euclidiana, por exemplo, não há garantia de que a topologia do objeto original seja preservada.

O esqueleto de uma região também pode ser obtido por meio do *diagrama de Voronoi*. Seja S um conjunto de n pontos no plano. O diagrama de Voronoi do conjunto S é uma partição do plano em n regiões convexas $V(p)$ ao redor de cada ponto p , tal que cada ponto no interior da região $V(p)$ está mais próximo de p do que de qualquer outro ponto em S . Um exemplo de diagrama de Voronoi para um determinado conjunto de pontos é mostrado na figura 7.12.

Figura 7.11: Transformada de distância Euclidiana. (a) Imagem binária de um triângulo preto sobre fundo branco. (b) Resultado da transformada de distância para vizinhança-4. (c) Resultado da transformada de distância para vizinhança-8. (d) Esqueleto do objeto, formado pelos valores mais altos de distância.

As áreas
diagrama,
Assim, to
leto interr
Técni
remoção
per a co
realizar
sejam ac

Para
definida
enquant

O m
imagem
de dois

(a) 2
(b) 3
(c) 4
(d) 5

$$(d) p_4 p_6 p_8 = 0$$

$$(c) p_2 p_4 p_6 = 0$$

$$(b) S(p) = 1$$

$$(a) 2 \leq N(p) \leq 6$$

O método de afinamento de Zhang e Sun (1984) verifica se um determinado pixel da enquadramento que os pixels pertencentes ao fundo possuem valor 0 (branco). O método pode ser eliminado pela análise de seus oito pixels vizinhos. O método é composto de dois passos. No passo 1, um pixel p é eliminado se satisfizer todas as seguintes condições:

Para a discussão dos algoritmos de afinamento a seguir, a vizinhança de um pixel é definida conforme a figura 7.14. Assume-se que os objetos possuem pixels de valor 1 (preto), enquanto que os pixels pertencentes ao fundo possuem valor 0 (branco).

Sejam aquelas que pertencem ao esqueleto do objeto.

realizar o afinamento é por meio de sucessivas remoções, até que os pixels remanescentes per a conectividade e nem causar erosão excessiva do objeto. Uma maneira comum de remoção de pixels. O algoritmo não deve remover pontos extremos, não deve interromper a conectividade do objeto, conforme visto na figura 7.13.

Técnicas de afinamento devem respeitar algumas condições durante o processo de leitura interno do objeto, conforme visto na figura 7.13.

Assim, tomam-se as arestas de um diagrama de Voronoi, o resultado é chamado de esqueleto do diagrama, tal que qualquer ponto sobre uma aresta possui mais de um ponto mais próximo. As arestas do diagrama de Voronoi representam a fronteira entre regiões adjacentes no

Figura 7.12: Diagrama de Voronoi para um dado conjunto de pontos.

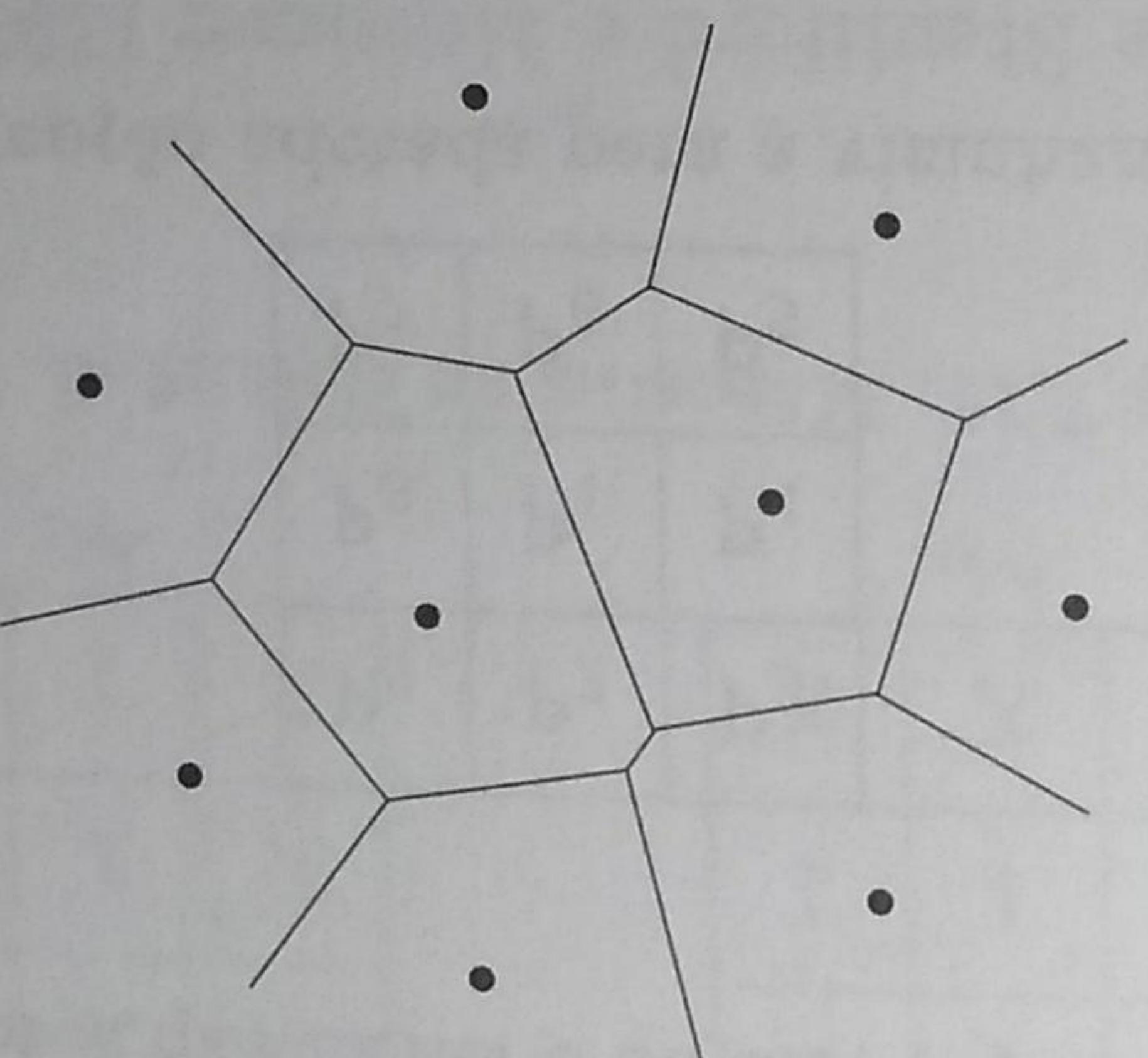
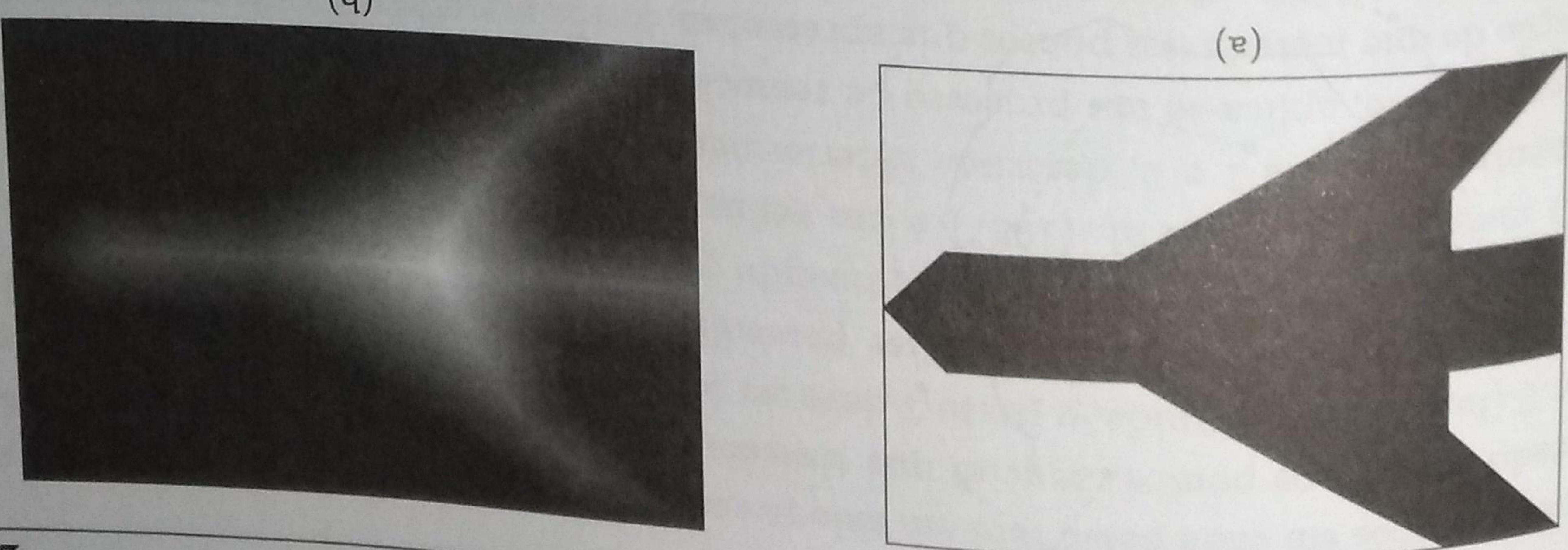


Figura 7.11: Transformada de distância. (a) imagem binária; (b) resultado da transformada da distância Euclidiana.



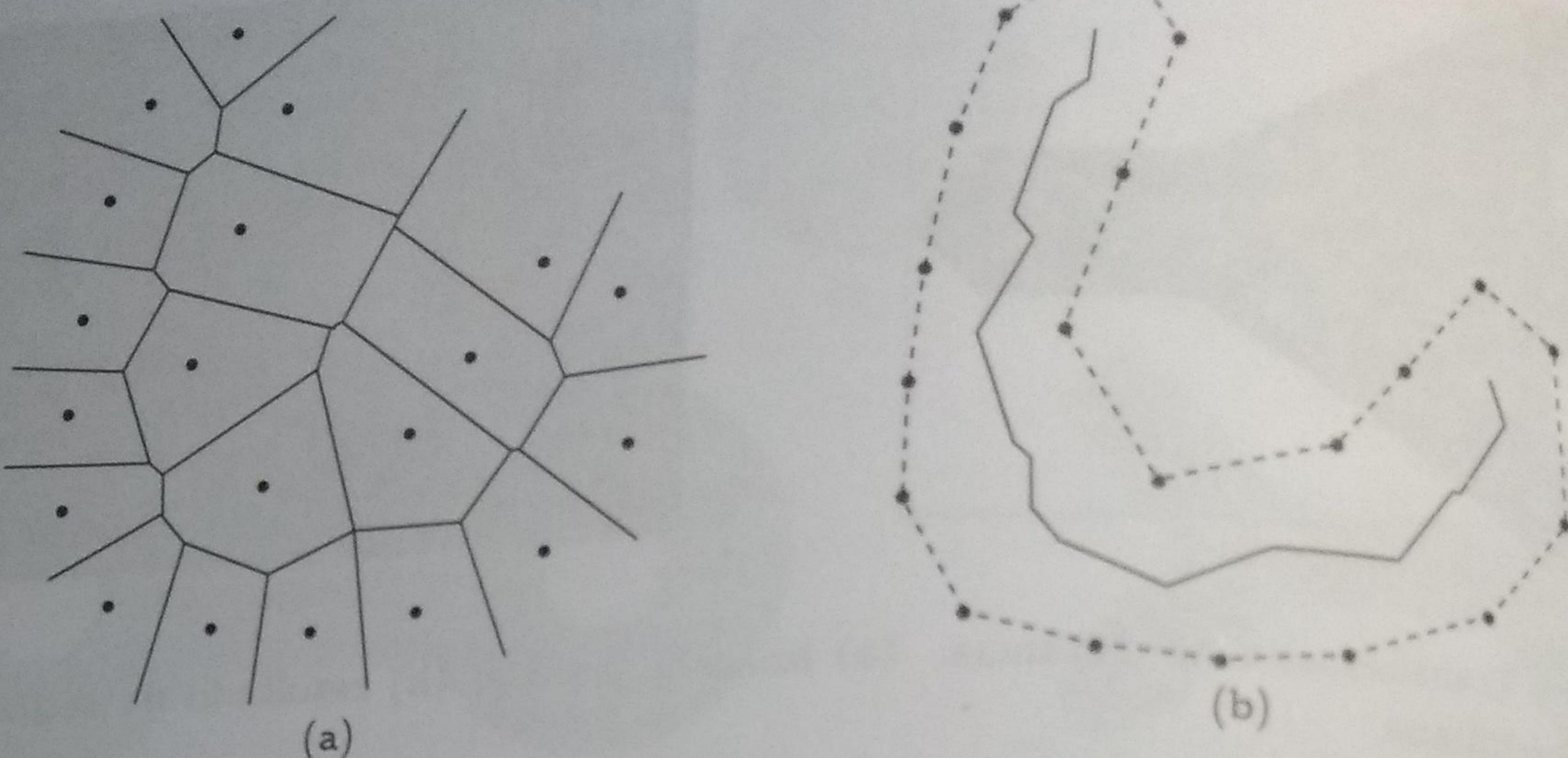


Figura 7.13: Esqueleto formado pelas arestas internas do diagrama de Voronoi. (a) diagrama de Voronoi para um conjunto de pontos que forma o objeto; (b) arestas internas ao objeto.

P ₉	P ₂	P ₃
P ₈	P ₁	P ₄
P ₇	P ₆	P ₅

Figura 7.14: Convenção adotada para a vizinhança-8 de um pixel.

em que $N(p)$ é o número de vizinhos não nulos de p , ou seja

$$N(p) = p_2 + p_3 + p_4 + p_5 + p_6 + p_7 + p_8 + p_9 \quad (7.1)$$

e $S(p)$ é o número de transições de branco para preto (do valor 0 para o valor 1) na seqüência ordenada $p_2, p_3, p_4, \dots, p_9, p_2$, ou seja, nos pixels que circundam o pixel central.

A condição (a) verifica se existem ao menos dois pixels pretos que sejam vizinhos ao pixel p , e não mais do que seis. Essa condição procura remover sucessivamente pixels da borda do objeto, ao invés de suas partes internas. A condição (b) verifica se o número de conectividade do pixel p é igual a 1. Essa condição assegura que um pixel a ser removido pertence a apenas um único objeto. A condição (c) verifica se ao menos um dos pixels vizinhos p_2, p_4 ou p_6 é fundo (branco) da imagem. A condição (d) verifica se ao menos um dos pixels p_4, p_6 ou p_8 é fundo da imagem.

No passo 2, um pixel p é eliminado se todas as condições abaixo forem satisfeitas:

- (a) $2 \leq N(p) \leq 6$
- (b) $S(p) = 1$
- (c) $p_2 p_4 p_8 = 0$
- (d) $p_2 p_6 p_8 = 0$

Se um ponto satisfizer todas as condições (a) a (d), ele deve ser, inicialmente, marcado para ser removido. Em outras palavras, o ponto não deve ser efetivamente eliminado até que

todos os pontos em cada passo tenham sido processados. Uma vez que o passo 1 tenha sido aplicado a todos os pontos, aqueles que tiverem sido marcados para remoção receberão o valor 0 (fundo). Em seguida, o passo 2 deve ser aplicado aos pontos resultantes exatamente da mesma maneira que o passo 1. Esse procedimento deve ser repetido até que não haja mais nenhum ponto a ser apagado, produzindo o esqueleto do objeto.

O método de Holt et al. (1987) é um refinamento do método de Zhang e Suen. Inicialmente, os passos 1 e 2, descritos anteriormente, são aplicados aos objetos da imagem. Posteriormente, aplica-se um processo de remoção de serrilhamento, o qual consiste na observação de que metade dos pontos que apresentam forma semelhante a uma escada pode ser removida sem afetar o formato ou a conectividade do objeto. O pixel central das máscaras mostradas na figura 7.15 pode ser removido se um dos valores X for 0 (branco).

<table border="1"> <tr><td>0</td><td>1</td><td>X</td></tr> <tr><td>1</td><td>1</td><td>X</td></tr> <tr><td>X</td><td>X</td><td>0</td></tr> </table> M_1	0	1	X	1	1	X	X	X	0	<table border="1"> <tr><td>X</td><td>1</td><td>0</td></tr> <tr><td>X</td><td>1</td><td>1</td></tr> <tr><td>0</td><td>X</td><td>X</td></tr> </table> M_2	X	1	0	X	1	1	0	X	X	<table border="1"> <tr><td>0</td><td>X</td><td>X</td></tr> <tr><td>X</td><td>1</td><td>1</td></tr> <tr><td>X</td><td>1</td><td>0</td></tr> </table> M_3	0	X	X	X	1	1	X	1	0	<table border="1"> <tr><td>X</td><td>X</td><td>0</td></tr> <tr><td>1</td><td>1</td><td>X</td></tr> <tr><td>0</td><td>1</td><td>X</td></tr> </table> M_4	X	X	0	1	1	X	0	1	X
0	1	X																																					
1	1	X																																					
X	X	0																																					
X	1	0																																					
X	1	1																																					
0	X	X																																					
0	X	X																																					
X	1	1																																					
X	1	0																																					
X	X	0																																					
1	1	X																																					
0	1	X																																					

Figura 7.15: Máscaras de Holt et al. (1987).

O método de afinamento de Stentiford e Mortimer (1983) utiliza quatro máscaras M_1 , M_2 , M_3 e M_4 , mostradas na figura 7.16, em que o valor 0 representa um pixel branco (fundo), valor 1 representa um pixel preto (objeto) e X denota que o pixel pode ter valor 0 ou 1.

<table border="1"> <tr><td>X</td><td>0</td><td>X</td></tr> <tr><td>X</td><td>1</td><td>X</td></tr> <tr><td>X</td><td>1</td><td>X</td></tr> </table> M_1	X	0	X	X	1	X	X	1	X	<table border="1"> <tr><td>X</td><td>X</td><td>X</td></tr> <tr><td>0</td><td>1</td><td>1</td></tr> <tr><td>X</td><td>X</td><td>X</td></tr> </table> M_2	X	X	X	0	1	1	X	X	X	<table border="1"> <tr><td>X</td><td>1</td><td>X</td></tr> <tr><td>X</td><td>1</td><td>X</td></tr> <tr><td>X</td><td>0</td><td>X</td></tr> </table> M_3	X	1	X	X	1	X	X	0	X	<table border="1"> <tr><td>X</td><td>X</td><td>X</td></tr> <tr><td>1</td><td>1</td><td>0</td></tr> <tr><td>X</td><td>X</td><td>X</td></tr> </table> M_4	X	X	X	1	1	0	X	X	X
X	0	X																																					
X	1	X																																					
X	1	X																																					
X	X	X																																					
0	1	1																																					
X	X	X																																					
X	1	X																																					
X	1	X																																					
X	0	X																																					
X	X	X																																					
1	1	0																																					
X	X	X																																					

Figura 7.16: Máscaras de Stentiford e Mortimer (1983).

Um *ponto final* é definido como um pixel preto que possui apenas um outro pixel preto como seu vizinho. Esses pixels não podem ser removidos da imagem. Os passos do método de Stentiford e Mortimer, bem com a ordem na qual as quatro máscaras devem percorrer a imagem, são mostrados no algoritmo 7.1.

Exemplos de aplicação dos algoritmos de afinamento propostos por Zhang e Suen (1984), Holt et al. (1987) e Stentiford e Mortimer (1983) são ilustrados na figura 7.17.

7.2 Descritores de Bordas

Esta seção apresenta um conjunto de descritores de bordas baseados em propriedades geométricas. Devido ao aspecto discreto das imagens digitais, os descritores de borda são sensíveis à resolução da imagem.

Afinamento de Stentiford e Mortimer

- 1 Cada máscara deve percorrer a imagem em uma determinada ordem:
- 2 M_1 : da esquerda para a direita e de cima para baixo
- 3 M_2 : de baixo para cima e da esquerda para a direita
- 4 M_3 : da direita para a esquerda e de baixo para cima
- 5 M_4 : de cima para baixo e da direita para a esquerda
- 6 Percorrer a imagem até encontrar um pixel que coincida com a máscara M_1 .
- 7 Se este pixel não for um ponto final e se o seu número de transições S_p for igual a 1, marcar este ponto para que seja removido posteriormente.
- 8 Repetir os passos (1) e (2) para todos os pixels que coincidam com a máscara M_1 .
- 9 Repetir os passos (1), (2) e (3) para cada uma das máscaras M_2 , M_3 e M_4 , nesta ordem.
- 10 Se algum ponto estiver marcado para ser removido, ele deve ser eliminado alterando seu valor para 0.
- 11 Se algum ponto foi removido no passo (5), repetir todos os passos a partir do passo (1). Caso contrário, o processo termina.

Algoritmo 7.1: Afinamento de Stentiford e Mortimer.

A	B	C	D	E	F	G	H	I	J	K	L	M	N
O	P	Q	R	S	T	U	V	W	X	Y	Z		
0	1	2	3	4	5	6	7	8	9				

A	B	C	D	E	F	G	H	I	J	K	L	M	N
O	P	Q	R	S	T	U	V	W	X	Y	Z		
0	1	2	3	4	5	6	0	7	8	9			

(a)

A	B	C	D	E	F	G	H	I	J	K	L	M	N
O	P	Q	R	S	T	U	V	W	X	Y	Z		
0	1	2	3	4	5	6	0	7	8	9			

A	B	C	D	E	F	G	H	I	J	K	L	M	N
O	P	Q	R	S	T	U	V	W	X	Y	Z		
0	1	2	3	4	5	6	6	7	8	9			

(b)

A	B	C	D	E	F	G	H	I	J	K	L	M	N
O	P	Q	R	S	T	U	V	W	X	Y	Z		
0	1	2	3	4	5	6	0	7	8	9			

A	B	C	D	E	F	G	H	I	J	K	L	M	N
O	P	Q	R	S	T	U	V	W	X	Y	Z		
0	1	2	3	4	5	6	6	7	8	9			

(c)

A	B	C	D	E	F	G	H	I	J	K	L	M	N
O	P	Q	R	S	T	U	V	W	X	Y	Z		
0	1	2	3	4	5	6	0	7	8	9			

A	B	C	D	E	F	G	H	I	J	K	L	M	N
O	P	Q	R	S	T	U	V	W	X	Y	Z		
0	1	2	3	4	5	6	6	7	8	9			

(d)

Figura 7.17: Exemplos de aplicação de algoritmos de afinamento. (a) imagem original; (b) método de Zhang e Suen; (c) método de Holt et al.; (d) método de Stentiford e Mortimer.

7.2.1 Diâmetro

O diâmetro de uma borda B é definido como a maior distância entre dois pontos p_i e p_j pertencentes à borda, ou seja

$$\text{Diâmetro}(B) = \max_{i,j} [D(p_i, p_j)] \quad (7.2)$$

sendo que D é uma medida de distância entre os pontos. O conceito de diâmetro de uma borda é ilustrado na figura 7.18. O diâmetro é invariante quanto à rotação e translação do objeto.

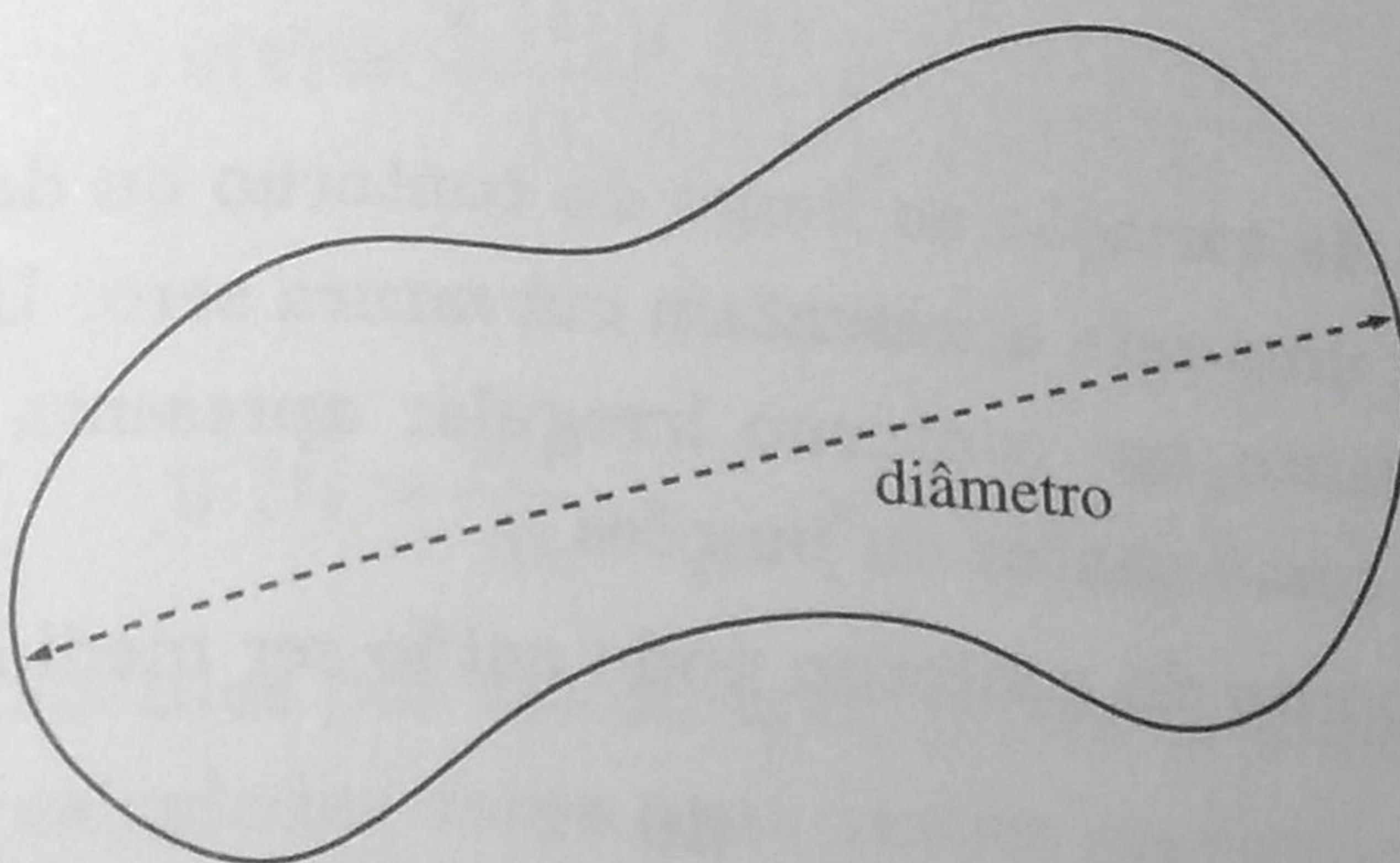


Figura 7.18: Diâmetro de uma borda.

7.2.2 Perímetro

O perímetro representa o comprimento da borda de um objeto. Uma aproximação para este descriptor pode ser obtida pela contagem dos pixels ao longo da borda do objeto na imagem. O perímetro é invariante quanto às operações de translação e rotação.

Uma curva fechada que representa a borda de um objeto, como ilustrada na figura 7.19, pode ser representada pela equação paramétrica complexa $z(t) = x(t) + i y(t)$, tal que $x(t)$ e $y(t)$ possuem período igual a 2π .

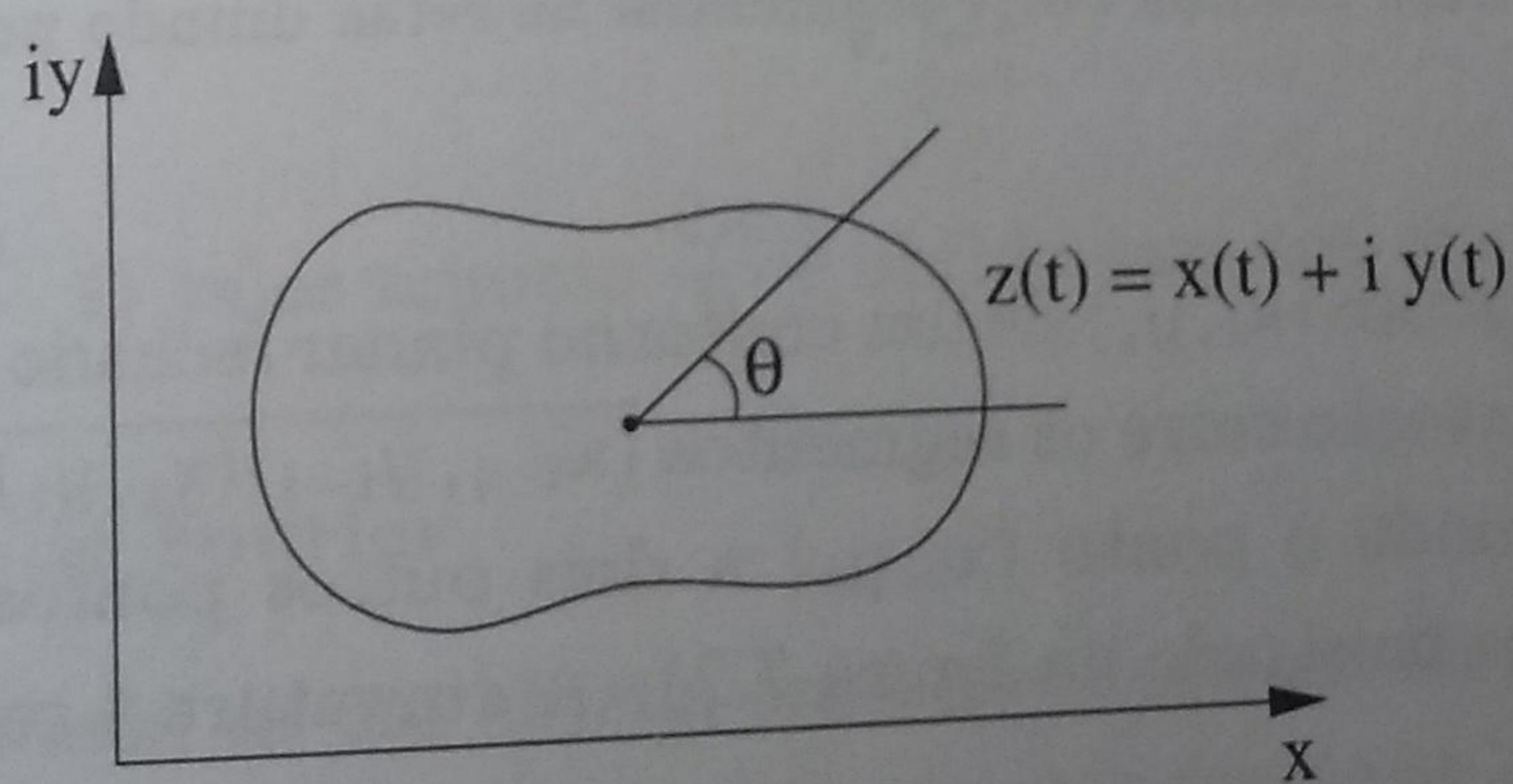


Figura 7.19: Representação paramétrica de uma curva fechada.

O perímetro relativo à borda do objeto é dado por

$$\text{Perímetro} = \int \sqrt{x^2(t) + y^2(t)} dt \quad (7.3)$$

No caso de uma borda representada pelo código da cadeia baseada na vizinhança-8, ilustrado na figura 7.1(b), o perímetro pode ser calculado pelo número de códigos nas direções horizontal e vertical, adicionado pelo número de códigos nas direções diagonais, levando-se em conta o fator $\sqrt{2}$, ou seja

$$\text{Perímetro} = N_p + \sqrt{2} N_i \quad (7.4)$$

em que N_p e N_i representam o número de códigos direcionais pares (componentes horizontal e vertical) e ímpares (componentes diagonais), respectivamente.

7.2.3 Curvatura

A curvatura mede a taxa de variação ao longo do contorno ou da borda do objeto. Dessa forma, pontos situados em uma reta apresentam curvatura zero. Uma circunferência possui curvatura constante, enquanto um contorno irregular apresenta curvatura mais alta em regiões de irregularidades, como cantos ou junções.

A curvatura κ em um ponto do contorno pode então ser medida como

$$\kappa = \frac{d\theta}{ds} \quad (7.5)$$

em que θ é a direção tangente do contorno e s é o perímetro do contorno. O valor de ds pode ser calculado como $ds = \sqrt{dx^2 + dy^2}$ (Schalkoff 1989).

A obtenção de medidas de curvatura confiáveis em uma borda digital não é simples. Um descritor comum calcula a diferença entre as inclinações de segmentos de retas que unem pontos consecutivos da borda, conforme mostrado na figura 7.20.

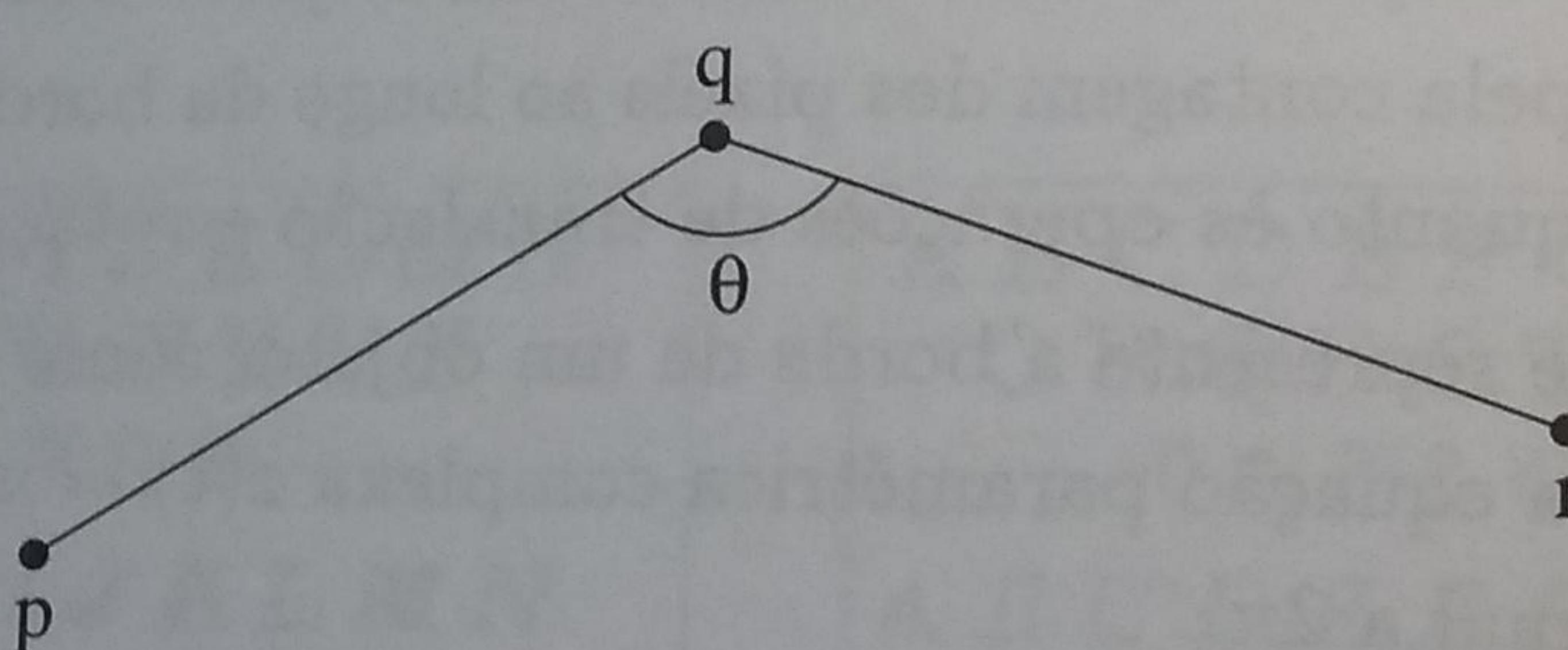


Figura 7.20: Curvatura medida entre segmentos de retas unindo pontos da borda.

A curvatura κ em um ponto (x_i, y_i) de um contorno planar fechado pode ser determinada por uma mudança na orientação entre os segmentos $\overrightarrow{(x_{i-j}, y_{i-j})(x_i, y_i)}$ e $\overrightarrow{(x_i, y_i)(x_{i+j}, y_{i+j})}$, ou seja, os segmentos unindo o ponto (x_i, y_i) a dois outros pontos localizados em uma vizinhança $j \geq 1$, conforme mostrado na figura 7.21. A curvatura é calculada como

$$\kappa(x_i, y_i) = \arctan\left(\frac{y_{i+j} - y_i}{x_{i+j} - x_i}\right) - \arctan\left(\frac{y_i - y_{i-j}}{x_i - x_{i-j}}\right) \quad (7.6)$$

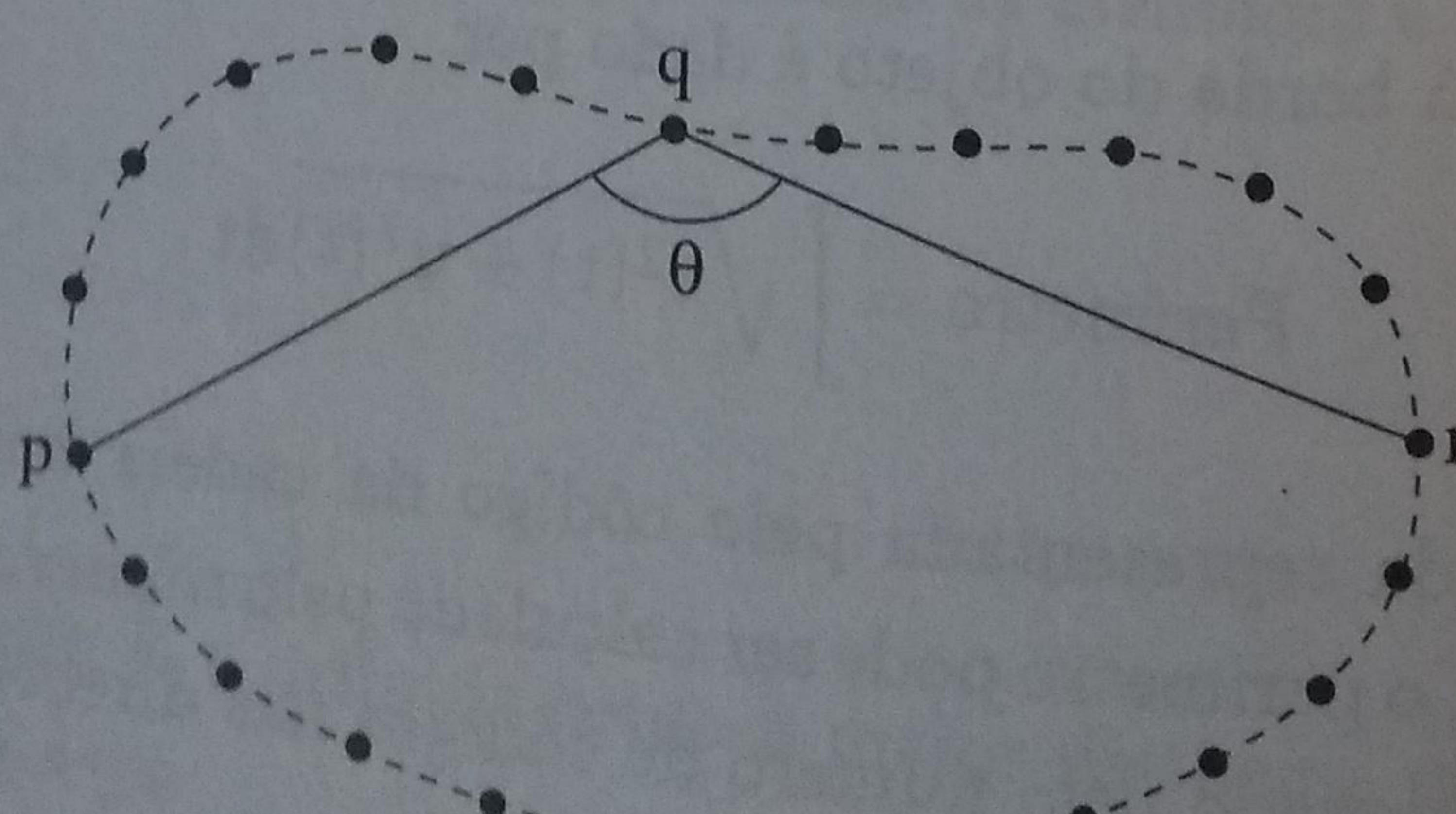


Figura 7.21: Curvatura em um ponto da borda com relação a dois pontos vizinhos.

Uma outra forma comum de medir a curvatura em um contorno utiliza a equação paramétrica da curva que forma o contorno. As coordenadas espaciais são dadas como uma função de um parâmetro $t \in [0, 1]$, expressas como $x(t)$ e $y(t)$. A curvatura pode ser dada pela equação

$$\kappa(t) = \frac{x'(t)y''(t) - x''(t)y'(t)}{(x'(t)^2 + y'(t)^2)^{3/2}} \quad (7.7)$$

tal que

$$x'(t) = \frac{dx}{dt} \quad y'(t) = \frac{dy}{dt} \quad x''(t) = \frac{d^2x}{dt^2} \quad y''(t) = \frac{d^2y}{dt^2} \quad (7.8)$$

Para reduzir efeitos causados por amostragem pobre de pontos de contorno na imagem, as derivadas podem ser calculadas sobre uma versão suavizada da imagem, por exemplo, obtida pela aplicação de um filtro Gaussiano.

Assim, a curvatura absoluta total ao longo de um contorno com perímetro L é dada por

$$\kappa_{\text{total}} = \frac{1}{L} \sum_{\rho=1}^L |\kappa(\rho)| = \oint |\kappa(\rho)| d\rho \quad (7.9)$$

7.2.4 Energia de Deformação

A energia de deformação pode ser entendida como a energia necessária para ajustar um elástico ou arame à forma desejada da borda. A energia de deformação pode ser calculada como a soma dos quadrados da curvatura $\kappa(\rho)$ da borda sobre o perímetro L da borda, ou seja

$$\text{Energia} = \frac{1}{L} \sum_{\rho=1}^L \kappa^2(\rho) \quad (7.10)$$

em que $\text{Energia} \geq \frac{2\pi}{R}$. O valor mínimo $\frac{2\pi}{R}$ é obtido para um círculo de raio R .

7.2.5 Descritores de Fourier

Seja C o contorno fechado de um objeto, conforme ilustrado na figura 7.22(a). O percurso no sentido anti-horário ao longo desse contorno, mantendo-se uma velocidade constante, fornece uma função complexa $z(t)$, em que t é uma variável que representa o tempo. A velocidade deve ser escolhida de tal forma que o percurso sobre o contorno leve um tempo igual a 2π . Assim, uma função periódica 2π é obtida após múltiplos percursos ao longo do contorno. Isso permite uma representação de Fourier¹ de $z(t)$ como

$$z(t) = \int T_n e^{int} \quad (7.11)$$

Os coeficientes T_n da série são chamados de descritores de Fourier do contorno C . Seja L o comprimento do contorno e s o comprimento do arco ao longo do contorno, então

$$t = \frac{2\pi s}{L} \quad (7.12)$$

¹A transformada de Fourier é apresentada no capítulo 3.

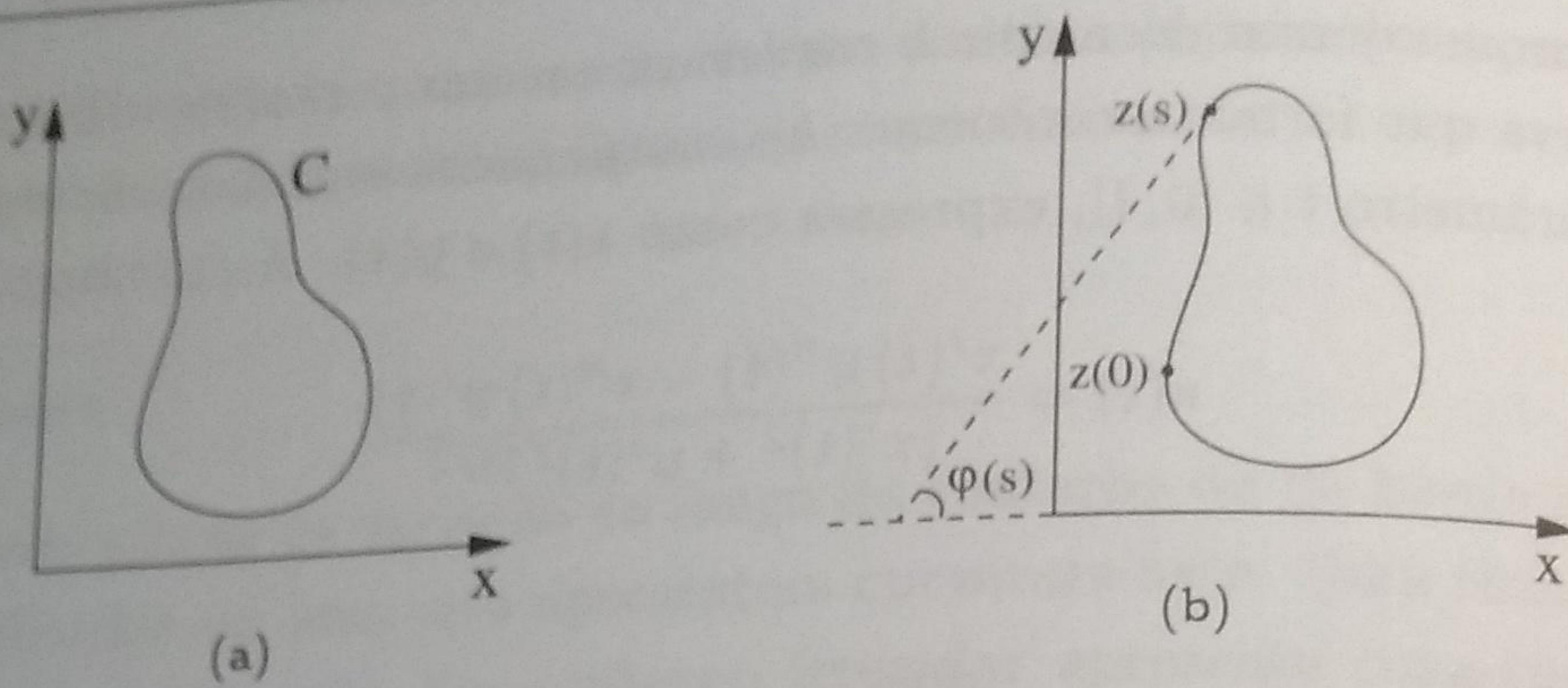


Figura 7.22: Descritores de Fourier. (a) descritores T_n ; (b) descritores S_n .

Os descritores de Fourier, T_n , são dados por

$$T_n = \frac{1}{L} \int_0^L z(t) e^{-i2\pi n t/L} dt \quad (7.13)$$

Os descritores são influenciados pela forma do contorno e pelo seu ponto inicial. Em imagens digitais, as coordenadas do contorno são discretas e a função $z(t)$ não é contínua. Assumindo que $z(k)$ é uma versão discreta de $z(t)$, na qual a vizinhança-4 é considerada entre os pixels do contorno para obter um intervalo de amostragem constante, os descritores T_n podem ser calculados a partir da transformada discreta de Fourier, discutida na seção 3.2, como

$$T_n = \frac{1}{N} \sum_{k=0}^{N-1} z(k) e^{-i2\pi n k/N} \quad (7.14)$$

em que (x_k, y_k) , $k = 0, 1, \dots, N - 1$, representam as coordenadas da borda do objeto, sendo N o comprimento da borda.

Os descritores de Fourier podem ser invariantes quanto à translação e rotação se o sistema de coordenadas for adequadamente escolhido (Pavlidis 1977, Persoon e Fu 1977, Wallace e Wintz 1980, Crimmins 1982, Lin e Chellappa 1987). Para um contorno representado pelas coordenadas (x_k, y_k) , adjacentes por vizinhança-4, em que $(x_0, y_0) = (x_N, y_N)$, então

$$\begin{aligned} a_n &= \frac{1}{N} \sum_{k=0}^{N-1} x_k e^{-i(2\pi/N)nk} \\ b_n &= \frac{1}{N} \sum_{k=0}^{N-1} y_k e^{-i(2\pi/N)nk} \end{aligned} \quad (7.15)$$

Os coeficientes a_n e b_n não são invariantes com relação às transformações de escala, translação e rotação. Invariância quanto à translação e rotação pode ser obtida com os descritores r_n definidos como

$$r_n = \sqrt{|a_n|^2 + |b_n|^2} \quad (7.16)$$

Para se obter a invariância com respeito à escala, os descritores w_n podem ser utilizados, tal que

$$w_n = \frac{r_n}{r_1} \quad (7.17)$$

Para aplicação em reconhecimento óptico de caracteres, Shridhar e Badreldin (1984) mostram que os primeiros 10 a 15 descritores w_n são eficientes para descrever caracteres.

Um contorno fechado pode também ser representado como uma função $\varphi(s)$ que mede a direção angular da reta tangente em relação ao comprimento do arco. A figura 7.22(b) ilustra a direção angular em um ponto da curva. Seja φ_k o ângulo medido no k -ésimo ponto do contorno e seja l_k a distância entre o ponto inicial do contorno e o k -ésimo ponto. A função periódica pode ser definida como

$$a(l_k) = \varphi_k + \mu_k \quad (7.18)$$

$$\mu_k = \frac{2\pi l_k}{L} \quad (7.19)$$

O descritor é então definido como

$$S_n = \frac{1}{2\pi} \int_0^{2\pi} a(u) e^{-inu} du \quad (7.20)$$

(7.13)

cial. Em
contínua.
siderada
scritores
eção 3.2,

(7.14)

, sendo
sistema
llace e
o pelas
o

(7.15)

m os

7.16)

17)

7.2.6 Modelos Deformáveis

Os modelos deformáveis ou modelos de contornos ativos são uma abordagem bastante útil para representação da forma de objetos e têm sido aplicados a um grande número de problemas em análise de imagens, tais como segmentação de regiões e rastreamento de movimento.

Os primeiros modelos deformáveis, conhecidos como *snakes*, foram propostos por Kass et al. (1988). Uma curva inicial é gradativamente deformada sob a ação de forças internas e externas, com o objetivo de minimizar uma função de energia. A curva atinge uma posição de equilíbrio quando ela se aproximar da borda do objeto na imagem.

Os modelos deformáveis permitem a incorporação de informações conhecidas *a priori* sobre o contorno dos objetos, tais como suavidade ou topologia, além de prover mecanismos de interação com o usuário, caso seja necessária sua intervenção para influenciar os resultados da aproximação. Extensões desses modelos para aproximação de objetos tridimensionais também têm sido propostas (Cohen e Cohen 1992, 1993).

Uma *snake* no plano pode ser definida como uma curva paramétrica fechada $c(s) = (x(s), y(s))$, em que $x(s)$ e $y(s)$ são as coordenadas x e y ao longo da curva e $s \in [0, 1]$ é o comprimento do arco normalizado ao longo do contorno.

O modelo original de *snakes* (Kass et al. 1988) é dado por

$$E_{\text{total}}(c(s)) = \int_0^1 [E_{\text{interna}}(c(s)) + E_{\text{externa}}(c(s))] ds \quad (7.21)$$

em que E_{interna} representa a energia interna da curva que determina a suavidade da curva e E_{externa} representa a energia externa que é responsável por atrair a curva à borda do objeto.

A energia interna pode ser expressa como

$$E_{\text{interna}}(c(s)) = \frac{\alpha(s)|c'(s)|^2 + \beta(s)|c''(s)|^2}{2} \quad (7.22)$$

$$E_{\text{interna}}(c(s)) = \frac{\alpha(s)|c'(s)|^2 + \beta(s)|c''(s)|^2}{2}$$

em que $c'(s)$ e $c''(s)$ são as derivadas de primeira e segunda ordens, respectivamente, de $c(s)$ com respeito a s . O parâmetro $\alpha(s)$ controla a elasticidade da curva e $\beta(s)$ controla a rigidez da curva.

A energia externa pode ser expressa como

$$E_{\text{externa}}(c(s)) = E_{\text{imagem}}(c(s)) + E_{\text{restrição}}(c(s)) \quad (7.23)$$

tal que o termo $E_{\text{imagem}}(c(s))$ está diretamente ligado às características da borda na imagem. O termo $E_{\text{restrição}}(c(s))$ está relacionado com outras restrições externas impostas à borda pelo usuário. Duas expressões comuns para $E_{\text{imagem}}(c(s))$, dadas pelas equações 7.24 e 7.25, são

$$E_{\text{imagem}}(c(s)) = -|\nabla f(x, y)|^2 \quad (7.24)$$

$$E_{\text{imagem}}(c(s)) = -|\nabla[G_\sigma(x, y) * f(x, y)]|^2 \quad (7.25)$$

em que ∇ é o operador gradiente e $G_\sigma(x, y)$ denota um filtro de suavização Gaussiano bidimensional com desvio padrão σ aplicado à imagem $f(x, y)$. O termo $E_{\text{imagem}}(c(s))$ assume valores baixos quando o gradiente é alto, por exemplo, nas bordas do objeto.

Uma dificuldade dos modelos deformáveis baseados em *snakes* é determinar a melhor localização para a curva inicial. Se a curva não estiver suficientemente próxima ao objeto de interesse, ela pode não ser atraída para a borda do objeto. Uma forma de resolver esse problema é a geração manual da curva inicial pelo usuário da aplicação. Outras dificuldades incluem a seleção correta dos parâmetros para o modelo, a convergência em mínimos locais de energia e a instabilidade do modelo devido à falta de nitidez das bordas (Gunn e Nixon 1997).

Um método de busca baseado em programação dinâmica para encontrar um mínimo global foi proposto por Amini et al. (1990), entretanto, apresenta alto custo computacional. Um algoritmo guloso proposto por Williams e Shah (1992), também baseado em programação dinâmica, utiliza uma busca local para reduzir o custo computacional. Uma otimização desse algoritmo foi proposta por Lam e Yan (1994). Uma abordagem estatística que utiliza a técnica de resfriamento simulado (do inglês, *simulated annealing*) (Geman e Geman 1984) foi proposta por Storvik (1994), para que a curva possa se mover mais dinamicamente em direção à configuração de energia mínima global.

Uma força adicional externa, chamada *força balão*, pode ser incorporada ao modelo para melhorar seu comportamento (Cohen e Cohen 1993). A força balão, denotada $E_{\text{balão}}(c(s))$, é um vetor unitário perpendicular à curva em um ponto. O objetivo dessa força é mover pontos na direção normal da curva, tal que a curva possa se expandir ou se contrair em direção à borda do objeto.

Embora a força balão torne o modelo de *snakes* menos sensível à localização da curva inicial, concavidades no objeto de interesse podem não ser corretamente identificadas.

Outra energia externa, conhecida como *fluxo do vetor de gradiente* (GVF, do inglês, *gradient vector flow*), foi proposta por Xu e Prince (1998). Nesse modelo, a energia é calculada como a difusão dos vetores de gradiente do mapa de bordas derivado da imagem. Em comparação com os modelos deformáveis originais, o modelo GVF apresenta menor sensibilidade às condições iniciais da curva e melhor capacidade para reconhecer concavidades na borda do objeto.

Figura 7.23: Modelo

O modelo de los (1995), combi geométrica e top tar elevado custo objetos complexo

O modelo du tornos interligad que se expande torna o método modelo requer a para representar

A utilização posta por Mene controla sua tem uso de *B-snake*

Um modelo força balão é c escolha de parâ se contraia de a

7.3 Descri

Nos descritore considerados r da região.

7.3.1 Área

A área pode de interesse. dimensões d

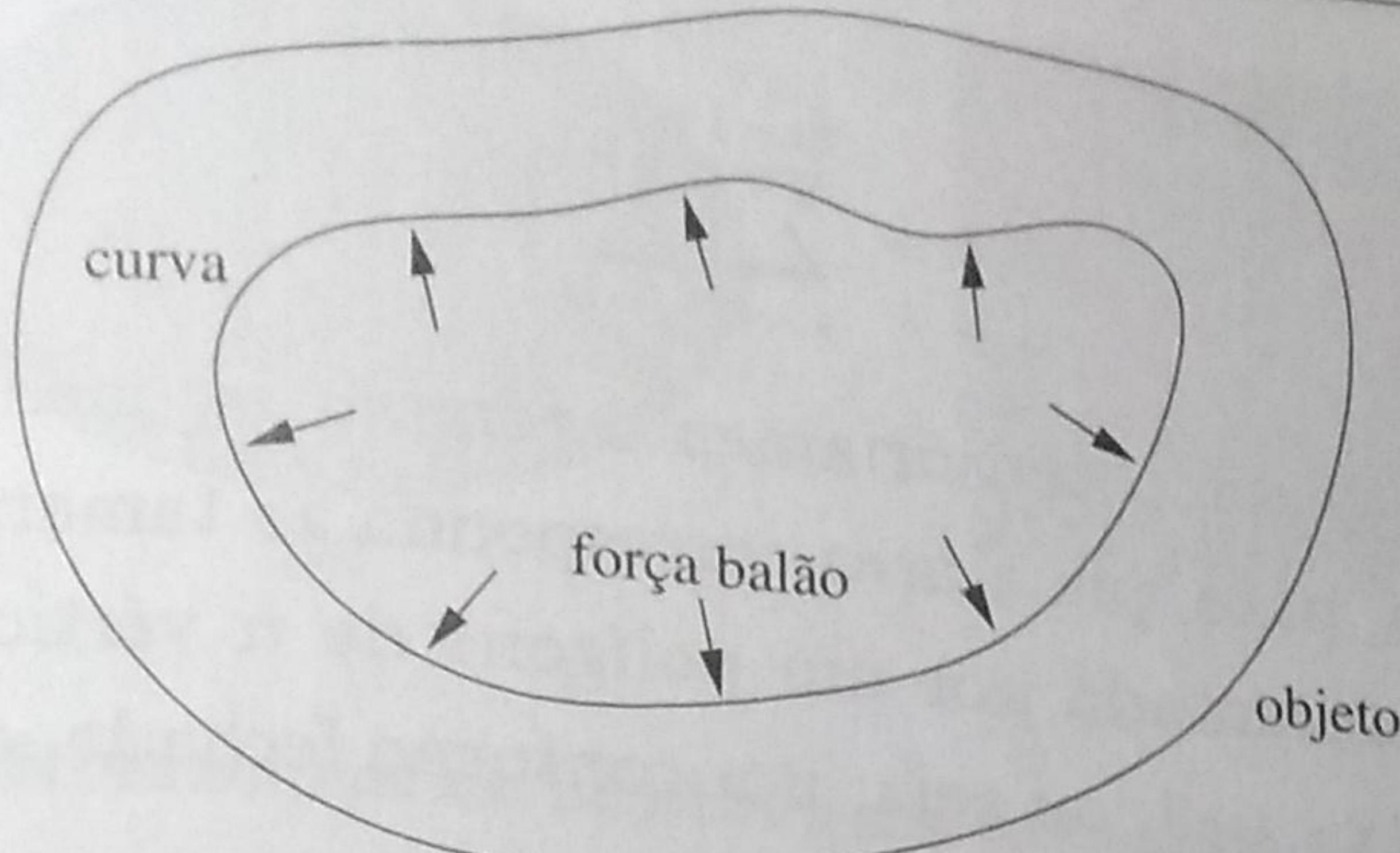


Figura 7.23: Modelo deformável com utilização de força balão para aproximação de borda de objeto.

O modelo de *topologia adaptável* (*T-snakes*), proposto por McInerney e Terzopoulos (1995), combina a representação paramétrica do modelo original com a flexibilidade geométrica e topológica dos modelos implícitos (Malladi et al. 1995). Apesar de apresentar elevado custo computacional, o modelo de topologia adaptável é capaz de representar objetos complexos que contenham saliências, bifurcações ou ramificações.

O modelo *dual*, proposto por Gunn e Nixon (1997), combina a adaptação de dois contornos interligados, um que se contraí a partir do exterior do objeto de interesse e outro que se expande a partir do interior do objeto de interesse. A utilização de duas curvas torna o método mais robusto, evitando convergência para mínimos locais. Entretanto, o modelo requer a determinação dos parâmetros das curvas iniciais e apresenta dificuldades para representar objetos com formas complexas.

A utilização de *B-splines* em conjunto com *snakes*, conhecida como *B-snakes*, foi proposta por Menet et al. (1990). O contorno é aproximado por uma curva *B-spline* que controla sua tensão e rigidez, substituindo a componente interna da equação de energia. O uso de *B-snakes* em múltiplas escalas foi descrito por Brigger e Unser (1998).

Um modelo estatístico para *snakes* foi proposto por Ivins e Porrill (1994), tal que a força balão é controlada em função de características estatísticas dos dados da imagem. A escolha de parâmetros apropriados tem o objetivo de fazer com que a curva se expanda ou se contraia de acordo com uma região inicial definida pelo usuário.

7.3 Descritores de Região

Nos descritores baseados em região, os pixels localizados no interior da região ou objeto são considerados no cálculo do descritor, em vez de utilizar apenas os pixels que formam a borda da região.

7.3.1 Área

A área pode ser expressa como o número de pixels que compreende o objeto ou a região de interesse. Em uma imagem binária, a área de um objeto contido em um retângulo com dimensões de $m \times n$ pixels, em que os pixels $f(x, y)$ do objeto são representados pelo valor

1, é definida como

$$\text{Área} = \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} f(x, y) \quad (7.26)$$

Em algumas aplicações, como sensoriamento remoto ou medicina, o tamanho real do pixel deve ser considerado, para que a área corresponda ao tamanho efetivo da região.

Caso o objeto seja representado por um polígono de n vértices, $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$, com $(x_n, y_n) = (x_0, y_0)$, ou seja, um contorno fechado, a área é dada por

$$\text{Área} = \frac{1}{2} \left| \sum_{k=0}^{n-1} (x_k y_{k+1} - x_{k+1} y_k) \right| \quad (7.27)$$

Embora a área possa ser utilizada como descritor de região, ela é geralmente aplicada a situações em que a escala dos objetos não varia. A área é invariante quanto à translação e rotação do objeto.

7.3.2 Compacidade ou Circularidade

A *compacidade* ou *circularidade* de uma região é definida como

$$C = \frac{P^2}{A} \quad (7.28)$$

em que P e A são o perímetro e a área da região, respectivamente, medidos em unidades de pixels. De acordo com esse descritor, um círculo possui o menor valor de compacidade, igual a 4π .

A compacidade é invariante quanto às transformações de translação, escala e rotação dos objetos. A figura 7.24 ilustra o conceito de compacidade em dois objetos com formas diferentes.

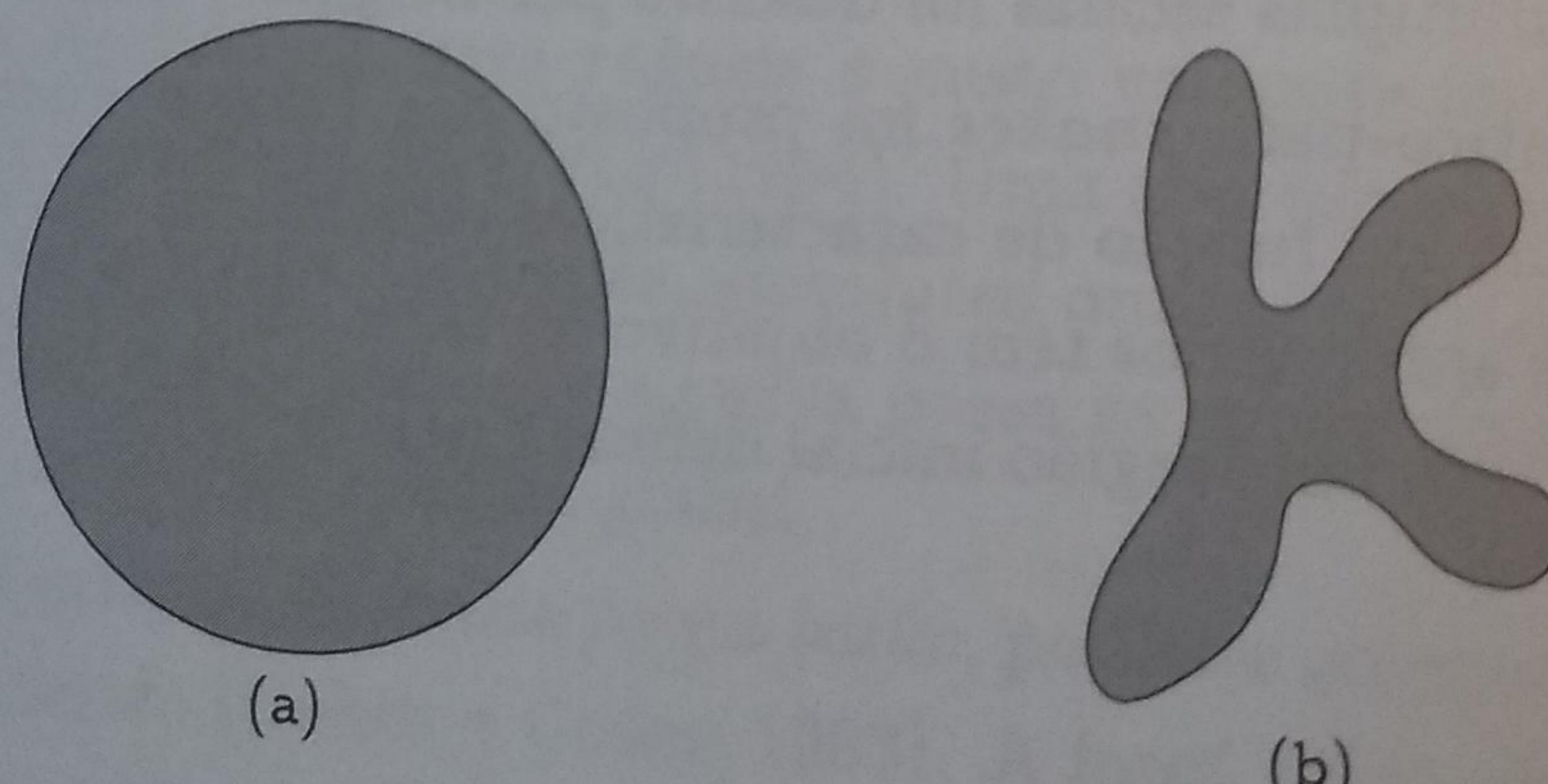


Figura 7.24: Compacidade. (a) objeto com baixa compacidade; (b) objeto com alta compacidade.

Exemplo 7.2 Cálculo da compacidade para os objetos mostrados na figura 7.25.
 (a) círculo

(b) elipse

$$C = \frac{(2\pi r)^2}{\pi r^2} = 4\pi$$

$$C = \frac{\left(2\pi\sqrt{\frac{a^2 + b^2}{2}}\right)^2}{\pi ab} = \frac{2\pi(a^2 + b^2)}{ab}$$

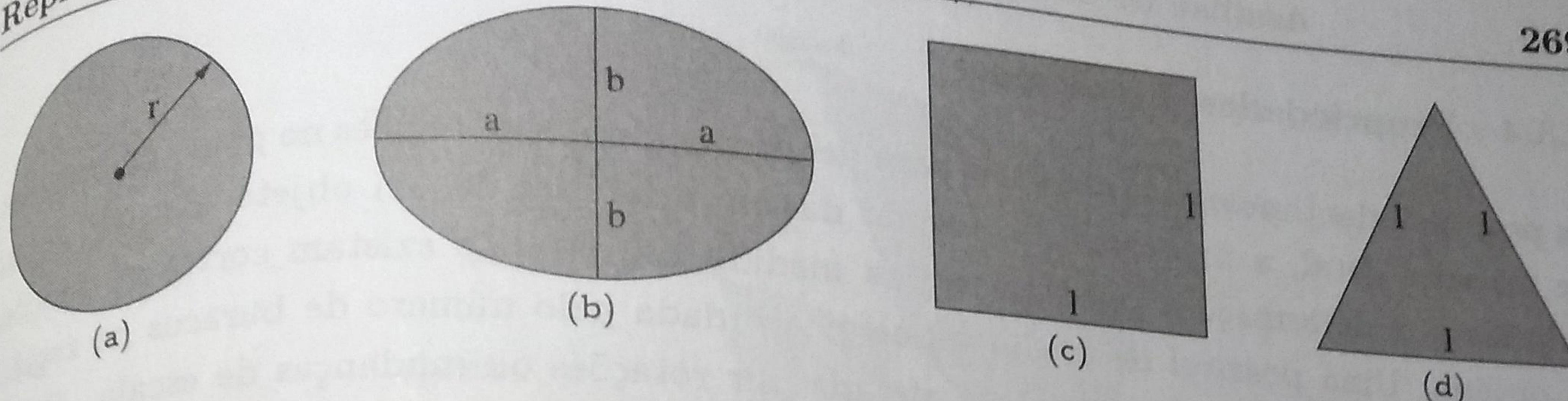


Figura 7.25: Exemplos de objetos para cálculo de compactade.

(7.26)
real do
gião.
, y_1 , ...
r

(c) quadrado

$$C = \frac{(4l)^2}{l^2} = 16$$

(7.27)
aplicada a
islação e

(d) triângulo

$$C = \frac{(3l)^2}{\frac{1}{2}\sqrt{l^2 - l^2/4}} = 12\sqrt{3}$$

(7.28) 7.3.3 Projeções Horizontal e Vertical

Projeções horizontal e vertical podem ser calculadas para uma região. Para uma imagem binária, em que os pixels $f(x, y)$ da região possuem valor igual a 1, a projeção horizontal P_h é definida como a soma dos pixels existentes em cada linha do objeto

$$P_h(y) = \sum_{x=0}^{M-1} f(x, y) \quad (7.29)$$

A projeção vertical P_v é definida como a soma dos pixels existentes em cada coluna do objeto

$$P_v(x) = \sum_{y=0}^{N-1} f(x, y) \quad (7.30)$$

Um exemplo de uso de projeções horizontal e vertical em linhas de texto contidas em uma imagem é apresentado na figura 7.26.

Exemplo de texto para ilustrar as projeções horizontal e vertical.

Figura 7.26: Projeções horizontal e vertical.

7.3.4 Propriedades Topológicas

As propriedades topológicas são úteis para descrições globais de regiões no plano da imagem. De maneira geral, a *topologia* é o estudo das propriedades de um objeto que não sejam afetadas por deformações geométricas, na medida em que não existam cortes ou junções no objeto. Uma possível descrição topológica é dada pelo número de buracos na região. Essa propriedade, obviamente, não é afetada por rotações ou mudanças de escala. Outra propriedade topológica útil para descrição de regiões é o número de componentes conexos, definidos conforme o conceito de vizinhança adotado entre os pixels.

O número de buracos H e de componentes conexos C em uma região de pixels podem ser usados na definição do *número de Euler* E

$$E = C - H \quad (7.31)$$

O número de Euler também é uma propriedade topológica. As regiões mostradas na figura 7.27, por exemplo, possuem números de Euler iguais a 0, -1 e 1, respectivamente, visto que o objeto "A" possui um componente conexo e um buraco, "B" possui um componente conexo e dois buracos, enquanto o objeto "C" possui um componente conexo e nenhum buraco.

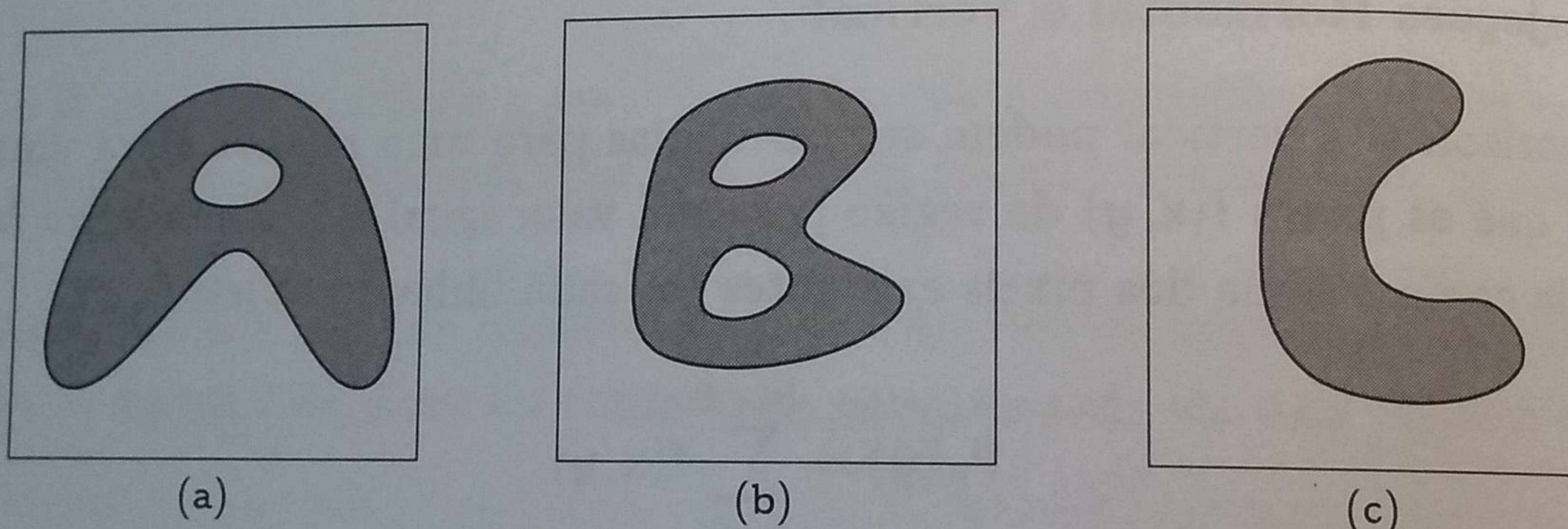


Figura 7.27: Regiões com diferentes números de Euler. (a) $E = 0$; (b) $E = -1$; (c) $E = 1$.

As regiões representadas por segmentos de retas (chamadas de *malhas poligonais*) possuem uma interpretação particularmente simples do número de Euler. A figura 7.28 mostra uma malha poligonal. Denotando o número de vértices por V , o número de arestas por A e o número de faces por F , tem-se a seguinte relação, conhecida como *fórmula de Euler*

$$V - A + F = C - H \quad (7.32)$$

que, em relação à equação 7.31, está relacionada ao número de Euler como

$$V - A + F = C - H = E \quad (7.33)$$

A malha poligonal mostrada na figura 7.28 possui 11 vértices, 18 arestas, 4 faces, 1 região conexa e 4 buracos. Portanto,

$$11 - 18 + 4 = 1 - 4 = -3 \quad (7.34)$$

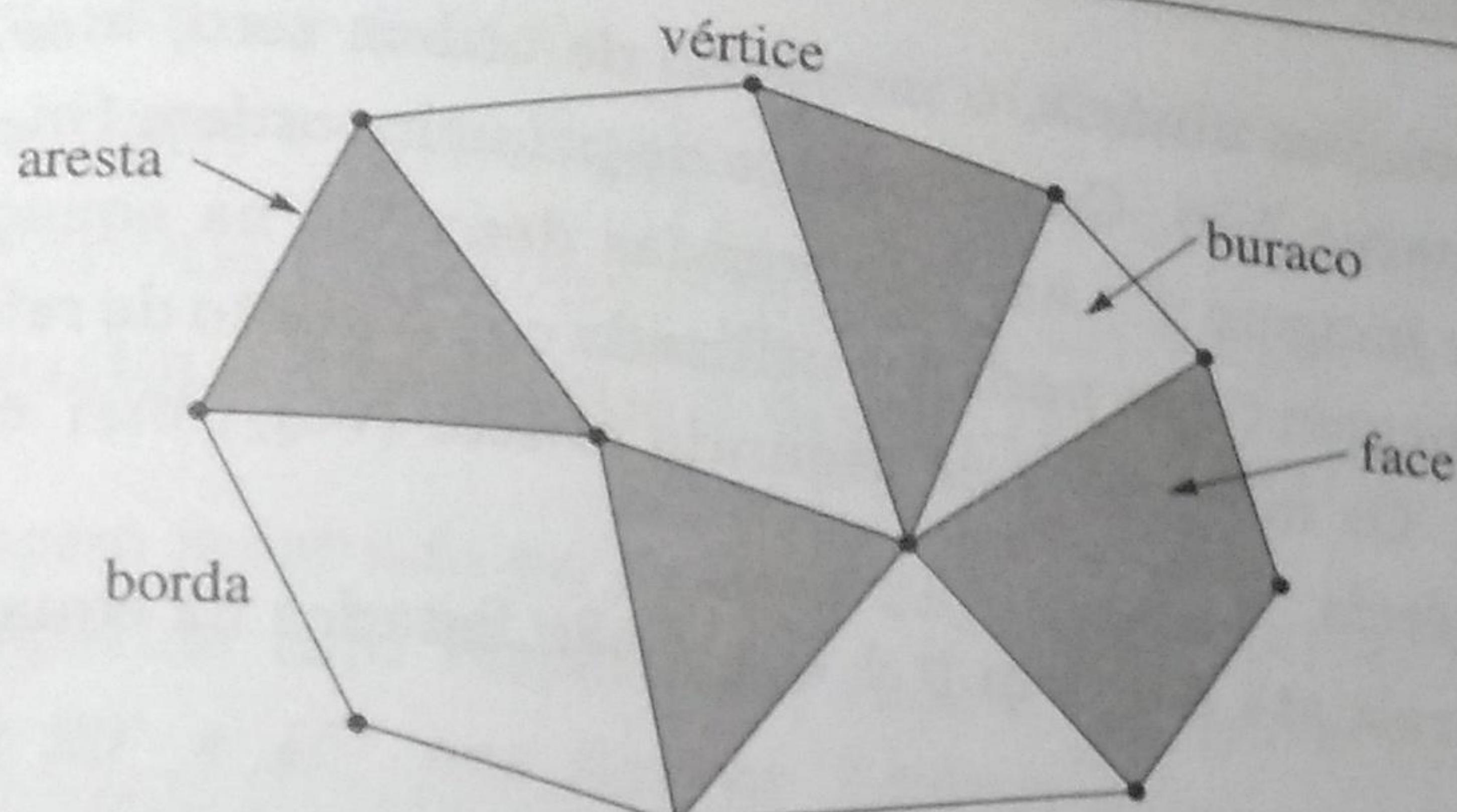


Figura 7.28: Malha poligonal.

7.3.5 Momentos

O conceito matemático de momentos tem sido utilizado em muitos campos de conhecimento, tais como física, estatística e análise de imagens. O uso de momentos em análise de imagens teve um grande avanço a partir dos trabalhos de Hu (1962) e Alt (1962).

Os momentos geométricos de ordem $p + q$ de uma função contínua bidimensional f são definidos como

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy \quad p, q = 0, 1, 2, \dots, \infty \quad (7.35)$$

A função pode ser representada, portanto, por um conjunto infinito de valores de momentos. Um conjunto de n momentos consiste em todos m_{pq} , para $p + q \leq n$, ou seja, o conjunto contém $1/2(n + 1)(n + 2)$ elementos. A aplicação prática dos momentos implica, obviamente, a utilização de um número finito de valores. A seleção de um conjunto representativo de valores de momentos que contenha informações suficientes para caracterizar a imagem de forma única é uma tarefa extremamente importante.

No caso de uma imagem com dimensões $M \times N$ pixels, a equação 7.35 pode ser reformulada usando a aproximação

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q f(x, y) \quad (7.36)$$

em que (x, y) são coordenadas de pontos das regiões.

Propriedades de invariância quanto às transformações de escala, translação e rotação podem ser derivadas para os momentos. Invariância com relação à translação pode ser alcançada com o uso de momentos centrais, μ_{pq} , definidos como

$$\mu_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (x - x_c)^p (y - y_c)^q f(x, y) \quad (7.37)$$

em que x_c, y_c são as coordenadas do centro de massa (centróide) da região, as quais podem ser obtidas usando as expressões

$$x_c = \frac{m_{10}}{m_{00}} \quad \text{e} \quad y_c = \frac{m_{01}}{m_{00}} \quad (7.38)$$

No caso de uma imagem binária, o momento de ordem zero, m_{00} , representa a área da região, definida na equação 7.26. Os momentos de primeira ordem (m_{01} e m_{10}) representam o centro de massa da imagem f . As coordenadas descritas na equação 7.38 definem uma localização única da imagem f que pode ser utilizada como ponto de referência para descrever a posição da imagem. Os momentos de segunda ordem (m_{02} , m_{11} e m_{20}) são conhecidos como momentos de inércia.

Os momentos centrais até a ordem $p + q = 3$ são listados na equação 7.39.

$$\begin{aligned}
 \mu_{00} &= m_{00} \\
 \mu_{01} &= \sum_x \sum_y (x - x_c)^0 (y - y_c)^1 f(x, y) = m_{01} - \frac{m_{01}}{m_{00}} m_{00} = 0 \\
 \mu_{10} &= \sum_x \sum_y (x - x_c)^1 (y - y_c)^0 f(x, y) = m_{10} - \frac{m_{10}}{m_{00}} m_{00} = 0 \\
 \mu_{11} &= \sum_x \sum_y (x - x_c)^1 (y - y_c)^1 f(x, y) = m_{11} - \frac{m_{01}}{m_{00}} m_{10} \\
 \mu_{02} &= m_{02} - \frac{m_{01}^2}{m_{00}} \\
 \mu_{20} &= m_{20} - \frac{m_{10}^2}{m_{00}} \\
 \mu_{12} &= m_{12} - 2 \frac{m_{01}}{m_{00}} m_{11} - \frac{m_{10}}{m_{00}} m_{02} + 2 \left(\frac{m_{01}}{m_{00}} \right)^2 m_{10} \\
 \mu_{21} &= m_{21} - 2 \frac{m_{10}}{m_{00}} m_{11} - \frac{m_{01}}{m_{00}} m_{20} + 2 \left(\frac{m_{10}}{m_{00}} \right)^2 m_{01} \\
 \mu_{03} &= m_{03} - 3 \frac{m_{01}}{m_{00}} m_{02} + 2 \left(\frac{m_{01}}{m_{00}} \right)^2 m_{01} \\
 \mu_{30} &= m_{30} - 3 \frac{m_{10}}{m_{00}} m_{20} + 2 \left(\frac{m_{10}}{m_{00}} \right)^2 m_{10}
 \end{aligned} \tag{7.39}$$

Os momentos centrais, embora invariantes quanto à translação, são dependentes da escala e rotação. A invariância quanto à escala pode ser obtida normalizando-se μ_{pq} como

$$\eta_{pq} = \frac{\mu_{pq}}{(\mu_{00})^\gamma} \tag{7.40}$$

em que

$$\gamma = \frac{p+q+2}{2} \quad p+q \geq 2 \tag{7.41}$$

Invariância quanto à rotação pode ser obtida se o sistema de coordenadas for escolhido tal que $\mu_{11} = 0$ (Cash e Hatamian 1987). Uma forma geral de invariância é mostrada por Hu (1962) e discutida por Wong e Hall (1978), Maitra (1979), Jain (1989) e Pratt (1991), em que sete momentos invariantes à translação, rotação e escala são calculados em função dos momentos de segunda e terceira ordens, listados na equação 7.42.

$$\begin{aligned}
 \phi_1 &= \eta_{20} + \eta_{02} \\
 \phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
 \phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
 \phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
 \phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] +
 \end{aligned} \tag{7.42}$$

Representação e Descriç

$$(3\eta_{21} - \eta_{03})(\eta_{30} - 3\eta_{12})$$

$$\phi_6 = (\eta_{20} - \eta_{02})(\eta_{30} - 3\eta_{12})$$

$$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} - 3\eta_{12})$$

Exemplo 7.3 A imagem na figura 7.29(b), reflet por uma rotação de 20 momentos invariantes, imagens.

Os resultados, apre mentos. A pequena di da imagem. Vale ress apenas os valores dos dos descritores. O va complexos resultantes

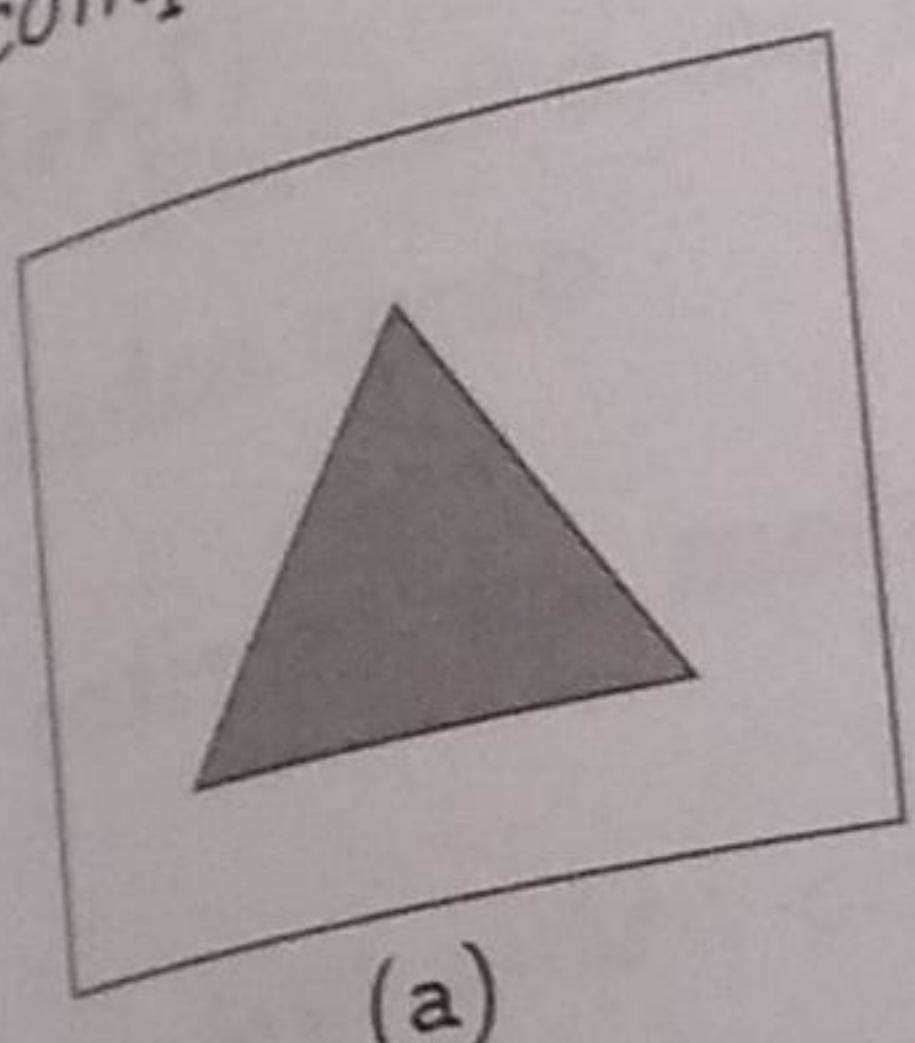


Figura 7.29:

Tabela 7.1: Mo

Momentos Invariantes (log)

ϕ_1

ϕ_2

ϕ_3

ϕ_4

ϕ_5

ϕ_6

ϕ_7

Embora os s gens em virtud não são invaria com respeito à

²Transformaç

$$\begin{aligned}\phi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ \phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - \\ &\quad (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]\end{aligned}$$

Exemplo 7.3 A imagem mostrada na figura 7.29(a) foi reduzida à metade do tamanho na figura 7.29(b), refletida com respeito ao eixo x na figura 7.29(c) e transformada por uma rotação de 20° e 45° nas figuras 7.29(d) e (e), respectivamente. Os sete momentos invariantes, descritos na equação 7.42, são calculados para cada uma das imagens.

Os resultados, apresentados na tabela 7.1, demonstram a invariância desses momentos. A pequena diferença entre os valores pode ser atribuída à natureza discreta da imagem. Vale ressaltar que o valor absoluto do logaritmo foi calculado em vez de apenas os valores dos momentos invariantes, de modo a reduzir o intervalo dinâmico dos descritores. O valor absoluto foi utilizado para evitar a manipulação de números complexos resultantes do logaritmo negativo dos momentos invariantes.

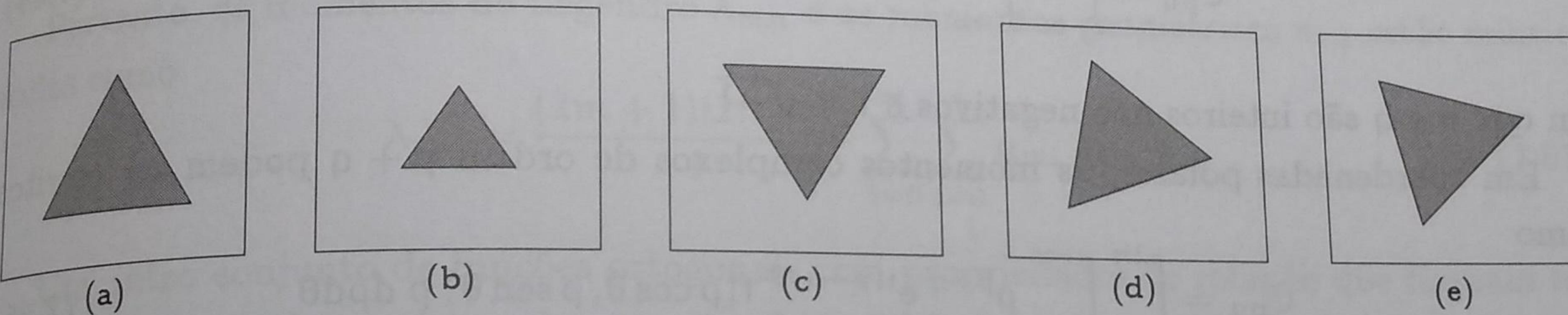


Figura 7.29: Imagens utilizadas para demonstrar a invariância dos momentos.

Tabela 7.1: Momentos invariantes para as imagens mostradas nas figuras 7.29(a) a (e).

Momentos Invariantes (log)	Imagen Original	Após Redução	Após Reflexão	Após Rotação 20°	Após Rotação 45°
ϕ_1	1.1258	1.1644	1.1246	1.1247	1.1253
ϕ_2	2.6248	2.6965	2.6223	2.6251	2.6255
ϕ_3	24.9650	24.1925	27.3290	18.4029	18.7755
ϕ_4	16.3450	18.9354	16.2082	17.4779	17.0769
ϕ_5	37.0217	40.6211	39.3020	35.4284	35.5113
ϕ_6	17.9348	20.5269	17.7990	18.8074	19.9473
ϕ_7	38.7063	45.5126	38.0134	37.5000	35.3282

Embora os sete momentos apresentados na equação 7.42 sejam úteis em análise de imagens em virtude da invariância com respeito à translação, rotação e escala, esses momentos não são invariantes sob transformações afins². Um conjunto de quatro momentos invariantes com respeito às transformações afins, derivados a partir de momentos de segunda e terceira

²Transformações afins são descritas no capítulo 10.

ordem, foi proposto por Flusser e Suk (1993) como

$$\begin{aligned} I_1 &= \frac{\mu_{20}\mu_{02} - \mu_{11}^2}{\mu_{00}^4} \\ I_2 &= \frac{\mu_{30}^2\mu_{03}^2 - 6\mu_{30}\mu_{21}\mu_{12}\mu_{03} + 4\mu_{30}\mu_{12}^3 + 4\mu_{21}^3\mu_{03} - 3\mu_{21}^2\mu_{12}^2}{\mu_{00}^{10}} \\ I_3 &= \frac{\mu_{20}(\mu_{21}\mu_{03} - \mu_{12}^2) - \mu_{11}(\mu_{30}\mu_{03} - \mu_{21}\mu_{12}) + \mu_{02}(\mu_{30}\mu_{12} - \mu_{21}^2)}{\mu_{00}^7} \\ I_4 &= (\mu_{20}^3\mu_{03}^2 - 6\mu_{20}^2\mu_{11}\mu_{12}\mu_{03} - 6\mu_{20}^2\mu_{02}\mu_{21}\mu_{03} + 9\mu_{20}^2\mu_{02}\mu_{12}^2 + 12\mu_{20}\mu_{11}^2\mu_{21}\mu_{03} + \\ &\quad 6\mu_{20}\mu_{11}\mu_{02}\mu_{03}\mu_{30} - 18\mu_{20}\mu_{11}\mu_{02}\mu_{21}\mu_{12} - 8\mu_{11}^3\mu_{30}\mu_{03} - 6\mu_{20}\mu_{02}^2\mu_{30}\mu_{12} + \\ &\quad 9\mu_{20}\mu_{02}^2\mu_{21}^2 + 12\mu_{11}^2\mu_{02}\mu_{30}\mu_{12} - 6\mu_{11}\mu_{02}^2\mu_{30}\mu_{21} + \mu_{02}^3\mu_{30}^2)/\mu_{00}^{11} \end{aligned} \quad (7.43)$$

Abu-Mostafa e Psaltis (1984, 1985) propuseram o uso de *momentos complexos* como descritores de objetos. Os momentos complexos bidimensionais de ordem $p + q$ para uma imagem f são definidos como

$$C_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x + iy)^p (x - iy)^q f(x, y) dx dy \quad (7.44)$$

em que p e q são inteiros não negativos e $i = \sqrt{-1}$.

Em coordenadas polares, os momentos complexos de ordem $p + q$ podem ser escritos como

$$C_{pq} = \int_0^{2\pi} \int_0^{+\infty} \rho^{p+q} e^{i(p-q)\theta} f(\rho \cos \theta, \rho \sin \theta) \rho d\rho d\theta \quad (7.45)$$

Assumindo que os momentos complexos da imagem original e da imagem após ter sofrido uma rotação em coordenadas polares sejam denotados por C_{pq} e C_{pq}^r , respectivamente, a relação entre esses momentos é dada por

$$C_{pq}^r = C_{pq} e^{-i(p-q)\theta} \quad (7.46)$$

em que θ é o ângulo com que a imagem original sofreu a rotação.

Embora os momentos complexos possam ser invariantes quanto à rotação, Abu-Mostafa e Psaltis concluíram que esses descritores apresentam deficiências em razão da redundância ou supressão de informação.

Uma forma de reduzir o conjunto de momentos requeridos para descrever um objeto é utilizar momentos ortogonais. Dentre os momentos ortogonais mais conhecidos estão os momentos de Legendre e de Zernike.

O polinômio de Legendre de ordem n é definido como

$$P_n(x) = \sum_{i=0}^n a_{ni} x^i = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n \quad (7.47)$$

em que $n \geq 0$.

Os polinômios de Legendre formam um conjunto completo ortogonal no intervalo $[-1, 1]$, expressos como

$$\int_{-1}^1 P_m(x) P_n(x) dx = \frac{2}{2m+1} \delta_{mn} \quad (7.48)$$

em que $-1 \leq x \leq 1$, $P_m(x)$ e $P_n(x)$ são polinômios de Legendre de ordem m e n , respectivamente, e δ_{mn} é o símbolo de Kronecker, o qual assume o valor 1 quando $m = n$ e o valor 0 quando $m \neq n$.

Os momentos de Legendre de ordem $m + n$ de uma função $f(x, y)$ definida no domínio $[-1, 1] \times [-1, 1]$ são

$$\lambda_{mn} = \frac{(2m+1)(2n+1)}{4} \int_{-1}^1 \int_{-1}^1 P_m(x) P_n(y) f(x, y) dx dy \quad (7.49)$$

em que $m, n \geq 0$.

A partir das equações 7.35, 7.47 e 7.49, tem-se que

$$\begin{aligned} \lambda_{mn} &= \frac{(2m+1)(2n+1)}{4} \int_{-1}^1 \int_{-1}^1 \sum_{i=0}^m a_{mi} x^i \sum_{j=0}^n a_{nj} y^j f(x, y) dx dy \\ &= \frac{(2m+1)(2n+1)}{4} \sum_{i=0}^m \sum_{j=0}^n a_{mi} a_{nj} \int_{-1}^1 \int_{-1}^1 x^i y^j f(x, y) dx dy \end{aligned} \quad (7.50)$$

Portanto, os momentos de Legendre λ_{mn} e os momentos geométricos m_{ij} estão relacionados como

$$\lambda_{mn} = \frac{(2m+1)(2n+1)}{4} \sum_{i=0}^m \sum_{j=0}^n a_{mi} a_{nj} m_{ij} \quad (7.51)$$

Um outro conjunto de funções ortogonais com propriedades de rotação que formam um conjunto ortogonal completo sobre um círculo unitário, $x^2 + y^2 \leq 1$, foi proposto por Zernike (1934). A forma desses polinômios é dada por

$$V_{mn}(x, y) = V_{mn}(\rho \sin \theta, \rho \cos \theta) = R_{mn}(\rho) e^{in\theta} \quad (7.52)$$

em que $m \geq 0$ e n pode assumir valores inteiros positivos e negativos sujeitos às restrições de que $|n| \leq m$, $m - |n|$ é par, ρ é o comprimento do vetor da origem ao pixel em (x, y) e θ é o ângulo entre o vetor ρ e o eixo x no sentido anti-horário.

O polinômio radial ortogonal $R_{mn}(\rho)$ é definido como

$$R_{mn}(\rho) = \sum_{s=0}^{(m-|n|)/2} (-1)^s \frac{(m-s)!}{s! \left(\frac{m+|n|}{2} - s\right)! \left(\frac{m-|n|}{2} - s\right)!} \rho^{m-2s} \quad (7.53)$$

em que $R_{m,n}(\rho) = R_{m,-n}(\rho)$.

Os seis primeiros polinômios radiais ortogonais são

$$R_{00}(\rho) = 1$$

$$R_{11}(\rho) = \rho$$

$$R_{20}(\rho) = 2\rho^2 - 1$$

$$R_{22}(\rho) = \rho^2$$

$$R_{31}(\rho) = 3\rho^3 - 2\rho \quad (7.54)$$

$$R_{33}(\rho) = \rho^3$$

Em contraste com o uso de coordenadas cartesianas para definir os momentos centrais, a representação polar é utilizada para obter a invariância quanto à rotação. Os momentos

complexos de Zernike, Z_{mn} , são definidos como

$$Z_{mn} = \frac{m+1}{\pi} \iint_{x^2+y^2 \leq 1} f(x, y) V_{mn}^*(x, y) dx dy \quad (7.55)$$

ou, em coordenadas polares

$$Z_{mn} = \frac{m+1}{\pi} \int_0^{2\pi} \int_0^1 f(\rho, \theta) R_{mn}(\rho) e^{-in\theta} \rho d\rho d\theta \quad (7.56)$$

em que R_{mn} é o polinômio radial definido na equação 7.53. O operador * na equação 7.55 denota o conjugado complexo de um polinômio de Zernike.

Por causa das condições $m - |n| = \text{par}$ e $|n| \leq m$ para os polinômios de Zernike, o conjunto de polinômios de Zernike contém $1/2(m+1)(m+2)$ polinômios linearmente independentes se o grau máximo for n .

Como $Z_{mn}^* = Z_{m,-n}$, então $|Z_{mn}| = |Z_{m,-n}|$. Portanto, é necessário considerar apenas $|Z_{mn}|$ com $n \geq 0$. Se a condição $m - |n| = \text{par}$ for eliminada dos polinômios de Zernike, definidos na equação 7.52, $\{V_{mn}\}$ torna-se o conjunto de polinômios pseudo-Zernike. O conjunto dos polinômios pseudo-Zernike foi derivado por Bhatia e Wolf (1954) e tem propriedades análogas àquelas dos polinômios de Zernike.

7.3.6 Eixos Maior e Menor

O *eixo maior* corresponde ao comprimento do maior segmento de reta que pode ser traçado ao longo do objeto. O *eixo menor* é o comprimento do maior segmento de reta perpendicular ao eixo maior que pode ser traçado ao longo do objeto. A figura 7.30 ilustra os eixos maior e menor de um objeto. Os eixos maior e menor são invariantes quanto à translação e rotação do objeto.

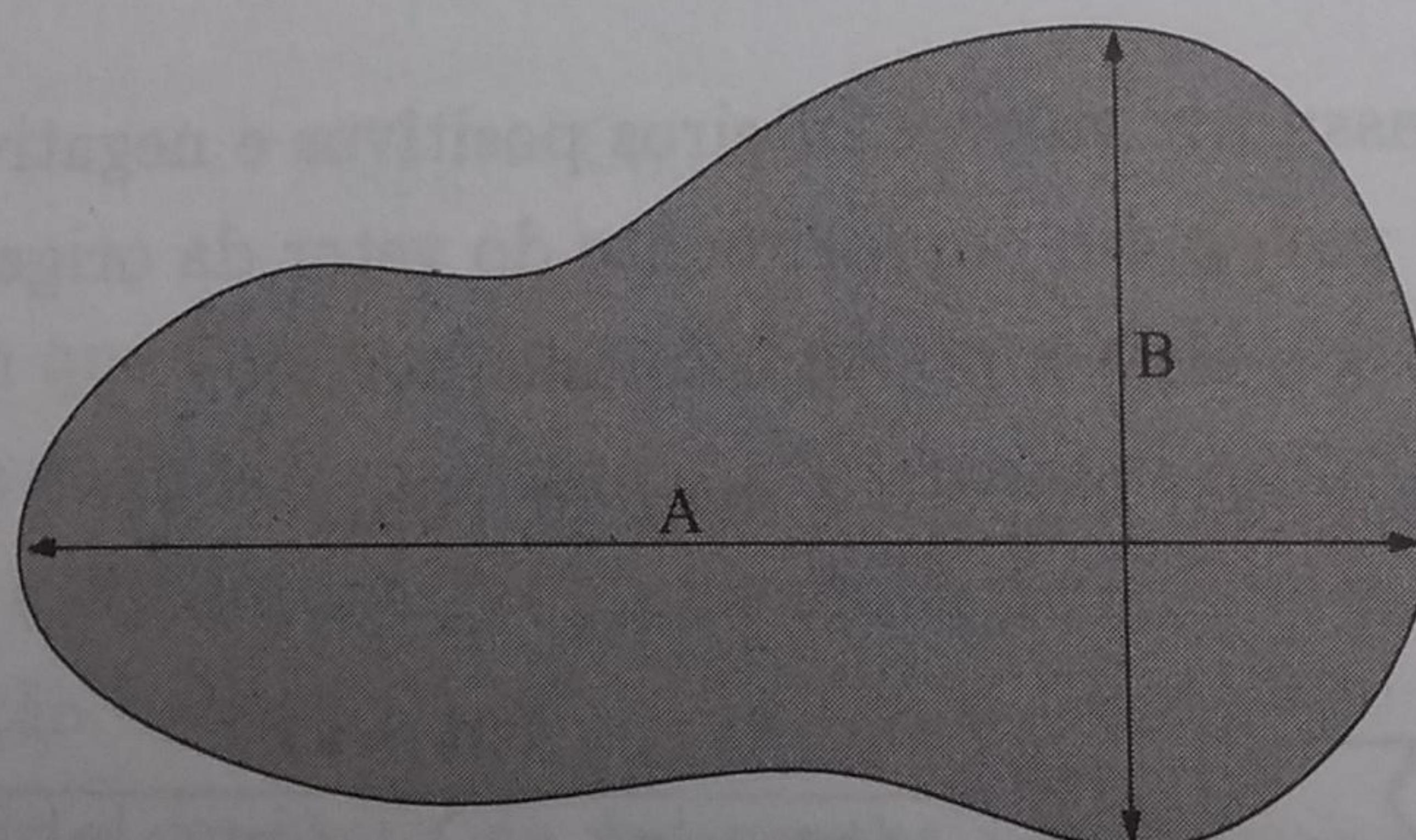


Figura 7.30: Eixos maior e menor.

7.3.7 Excentricidade

A excentricidade é a relação entre o comprimento do eixo maior A e o comprimento do eixo menor B , ou seja

$$\text{Excentricidade} = \frac{A}{B} \quad (7.57)$$

Exemplos de objetos apresentando valores de alta e baixa excentricidade são mostrados na figura 7.31

Uma outra mar
descritos na seção 7

7.3.8 Retângulo

O retângulo enu
mínimas que eng
do objeto. Test
exemplo, podem

Exemplos de
facilmente ester
cálculo do volu

7.3.9 C

A convexi
de convex

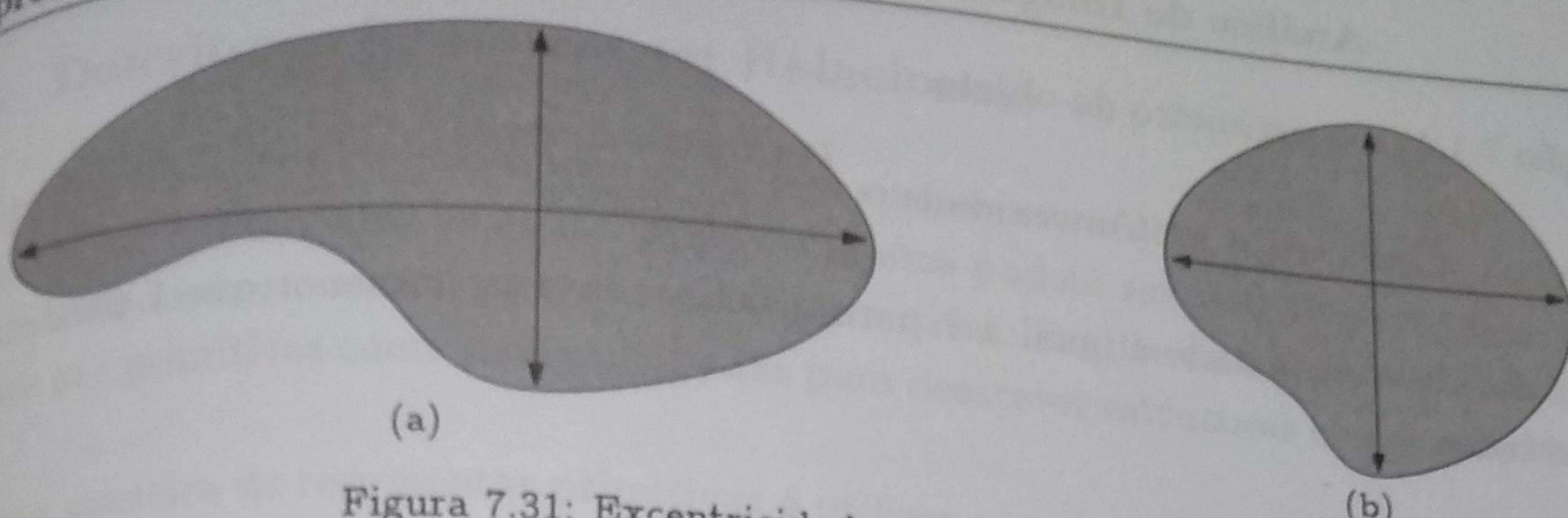


Figura 7.31: Excentricidade. (a) alta; (b) baixa.

Uma outra maneira de calcular a excentricidade é por meio dos momentos centrais, descritos na seção 7.3.5, dada por

$$\text{Excentricidade} = \frac{(\mu_{02} - \mu_{20})^2 + 4\mu_{11}}{\text{Área}} \quad (7.58)$$

7.3.8 Retângulo Envolvente

O retângulo envolvente de um objeto bidimensional corresponde ao retângulo de dimensões mínimas que engloba o objeto, cujos lados são paralelos aos eixos principais (menor e maior) do objeto. Testes para verificação de intersecção entre objetos ou cálculos de área, por exemplo, podem ser simplificados pelo uso de retângulos envolventes.

Exemplos de retângulos envolventes são mostrados na figura 7.32. O conceito pode ser facilmente estendido para objetos tridimensionais, o que pode reduzir consideravelmente o cálculo do volume de objetos em muitas aplicações práticas.

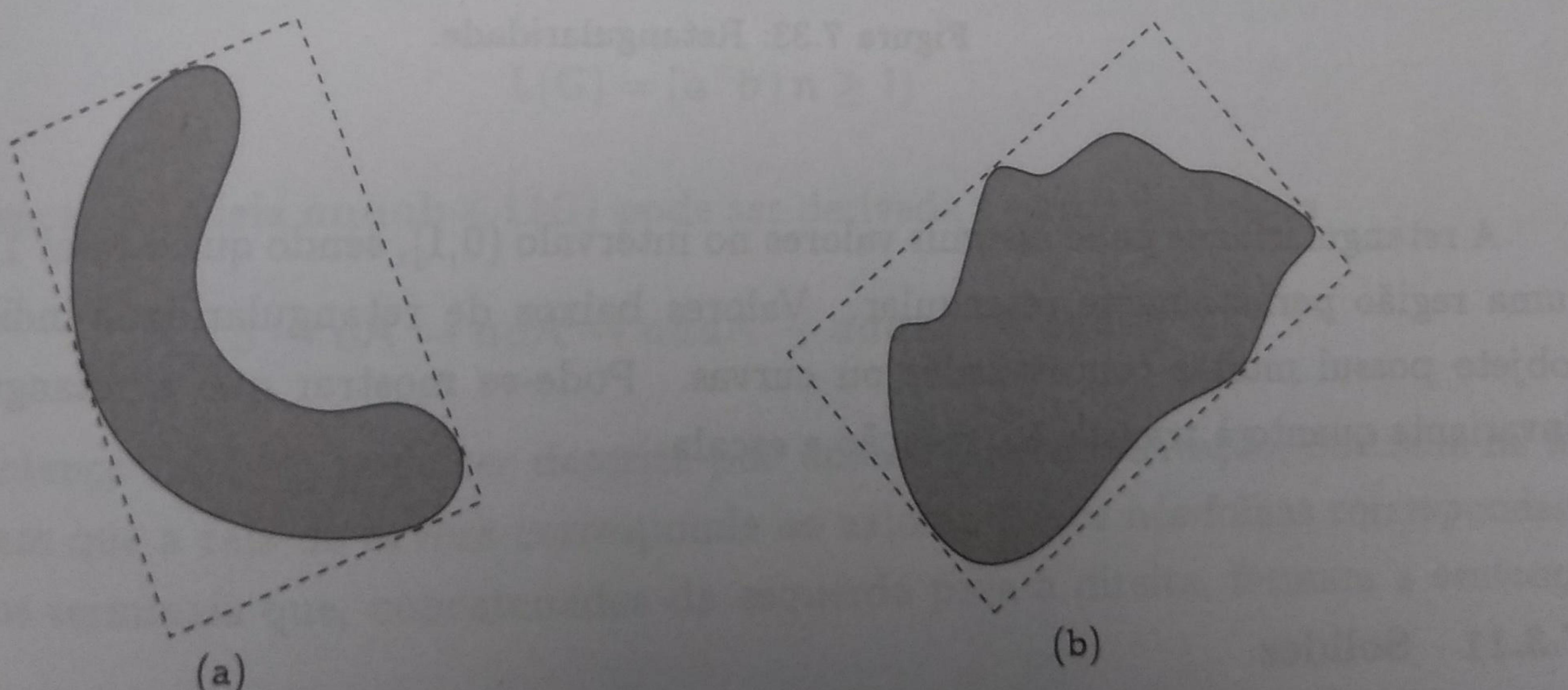


Figura 7.32: Retângulos envolventes.

7.3.9 Convexidade

A convexidade mede o grau com que um objeto difere de um objeto convexo. Uma medida de convexidade é dada pela relação entre o perímetro do fecho convexo do objeto (definido

na seção 7.1.4) e o perímetro do objeto.

$$\text{Convexidade} = \frac{\text{Perímetro convexo}}{\text{Perímetro}} \quad (7.59)$$

O valor de convexidade é igual a 1 para objetos convexos e menor que 1 para objetos não convexos.

7.3.10 Retangularidade

A retangularidade de um objeto é definida como a relação entre a área do objeto e a área do retângulo envolvente, ou seja

$$\text{Retangularidade} = \frac{\text{Área do objeto}}{L_1 L_2} \quad (7.60)$$

em que L_1 e L_2 são as dimensões do retângulo envolvente, conforme ilustrado na figura 7.33.

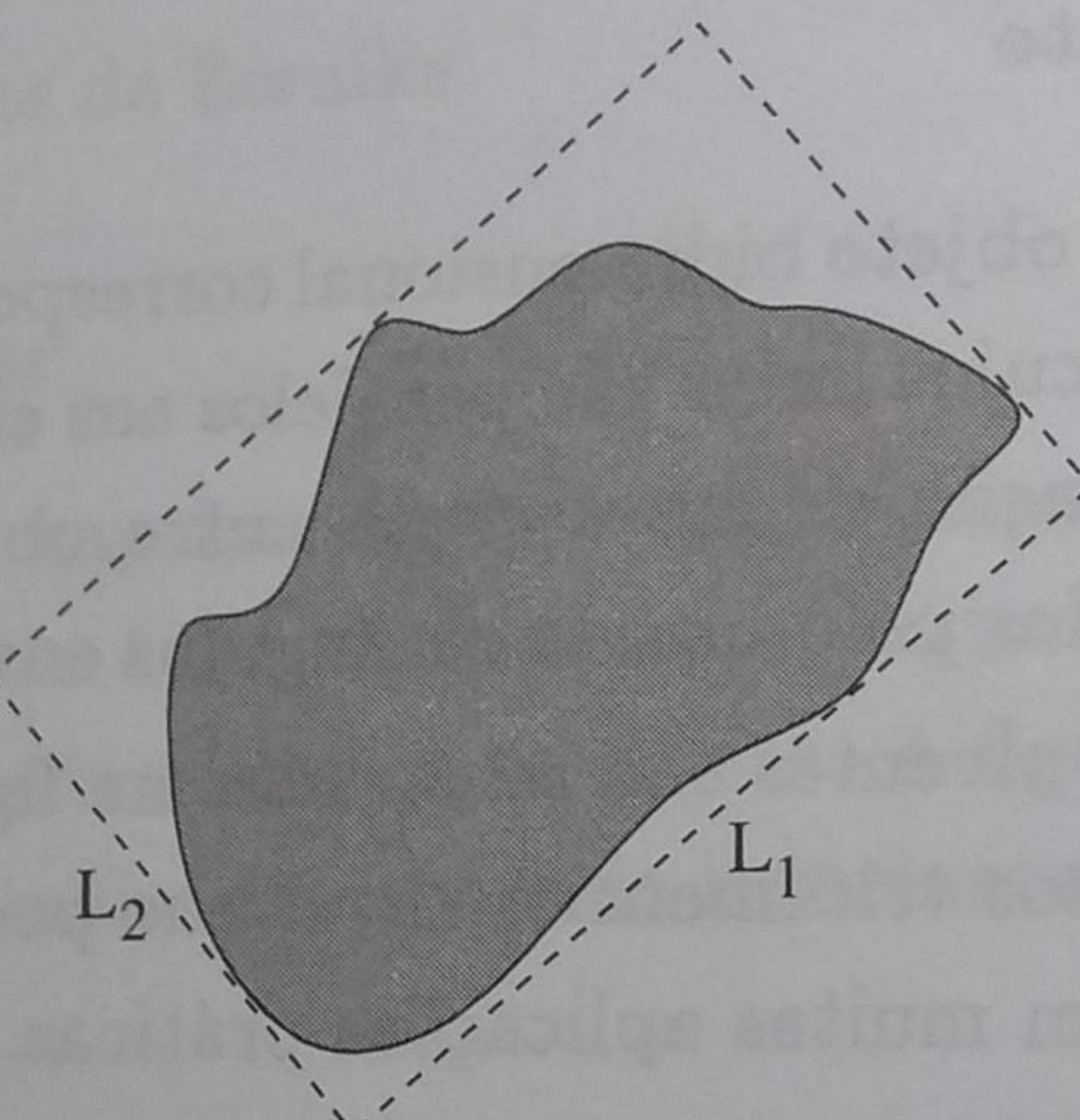


Figura 7.33: Retangularidade.

A retangularidade pode assumir valores no intervalo $(0,1]$, sendo que o valor 1 representa uma região perfeitamente retangular. Valores baixos de retangularidade indicam que o objeto possui muitas concavidades ou curvas. Pode-se mostrar que a retangularidade é invariante quanto à translação, rotação e escala.

7.3.11 Solidez

A solidez mede a densidade de um objeto. Uma medida de solidez pode ser obtida pela relação entre a área de um objeto e a área do fecho convexo do objeto.

$$\text{Solidez} = \frac{\text{Área do objeto}}{\text{Área do fecho convexo}} \quad (7.61)$$

O valor de solidez é igual a 1 para um objeto sólido e menor que 1 para objetos apresentando borda irregular ou contendo buracos.

7.4 Descritores Sintáticos ou Relacionais

Descritores sintáticos ou relacionais permitem que as estruturas segmentadas da imagem sejam organizadas de uma maneira formal. Objetos podem ser descritos recursivamente a partir de elementos mais simples, chamados primitivas. Bordas ou segmentos de bordas de objetos são primitivas comumente utilizadas para descrever estruturas e seus relacionamentos.

Uma maneira de representar primitivas é utilizar uma seqüência ordenada ou cadeia de caracteres. Cada primitiva é interpretada como um alfabeto de uma determinada gramática, tal que um conjunto de regras de sintaxe determinam a geração das seqüências formadas pelos símbolos do alfabeto. Um exemplo de uso dos descritores sintáticos é na análise de seqüências genéticas.

Em teoria de linguagens formais (Chomsky 1957, 1959), uma gramática é definida pela quádrupla

$$G = (N, T, P, S) \quad (7.62)$$

em que N é um conjunto finito de símbolos não terminais, T é um conjunto finito de símbolos terminais, P é um conjunto finito de produções ou regras de reescrita e S é o símbolo inicial.

A linguagem $L(G)$ de uma gramática G é definida como o conjunto de cadeias que podem ser derivadas a partir de um símbolo inicial por repetidas aplicações de regras de reescrita. Seja, por exemplo, a gramática $G = (N, T, P, S)$ definida como

$$N = \{S, A\}$$

$$T = \{a, b\}$$

$$P = \{S \rightarrow aA; A \rightarrow aA; A \rightarrow b\}$$

A gramática gera a linguagem

$$L(G) = \{a^n b \mid n \geq 1\}$$

A sentença ou cadeia $aaaab \in L(G)$ pode ser derivada a partir das regras

$$S \rightarrow aA \rightarrow aaA \rightarrow aaaA \rightarrow aaaaA \rightarrow aaaaab$$

Essa sentença também pode ser descrita por uma árvore de derivação, ilustrada na figura 7.34, em que a raiz da árvore corresponde ao axioma S e os nós-folhas correspondem aos símbolos terminais que, concatenados da esquerda para a direita, formam a sentença original.

A gramática formal G permite a descrição apropriada de um conjunto infinito de seqüências, de acordo com a linguagem $L(G)$. A construção de elementos complexos a partir de primitivas mais simples é realizada pelas regras de reescrita. As primitivas servem como elementos básicos dos objetos que possibilitam uma descrição mais compacta dos dados em termos de suas relações estruturais.

Para padrões de objetos em imagens, não há um conjunto universal de primitivas para utilização em diferentes aplicações. De modo a permitir uma descrição adequada dos objetos, as primitivas devem conter informação relativa sobre a aplicação em questão, tal como a

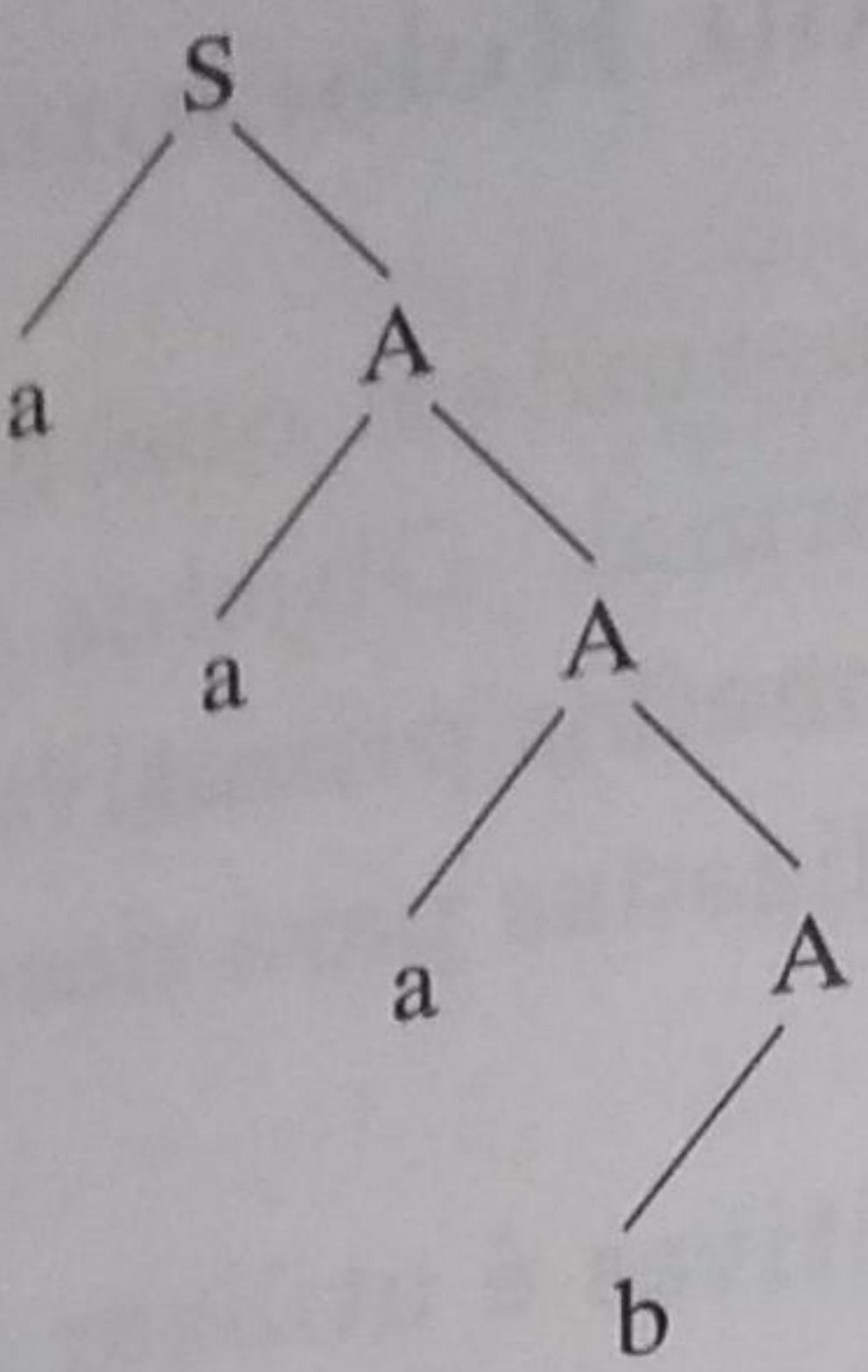


Figura 7.34: Árvore de derivação.

forma ou a localização dos elementos. Freqüentemente, o uso de informação semântica é também requerida na descrição de primitivas.

Exemplo 7.4 Sejam as primitivas gráficas definidas na figura 7.35(a).

e Kwok (1988). Técnicas de afinamento para imagens tridimensionais foram propostas por Ma (1994), Ma e Sonka (1996), Gagvani e Silver (1999) e Chuang et al. (2000).

Diversas referências descrevem métodos para geração do diagrama de Voronoi, algumas sugestões de textos incluem os trabalhos de Preparata e Shamos (1985), Aurenhammer (1991), Ohya et al. (1984) e O'Rourke (1993).

Descritores sintáticos são analisados em Fu (1982), Truve (1990), Henderson e Samal (1986), Bunke e Sanfeliu (1990) e Biswas e Dubes (1984).

7.6 Problemas

1. Descreva o conceito de transformada do eixo médio.
2. Determine o número de Euler para os objetos formados pelos caracteres A, B, C, R e S.
3. Encontre o código da cadeia de vizinhança-8, definido no sentido anti-horário, para o contorno mostrado na figura 7.36. Note que o ponto inicial está marcado com um círculo vazio.

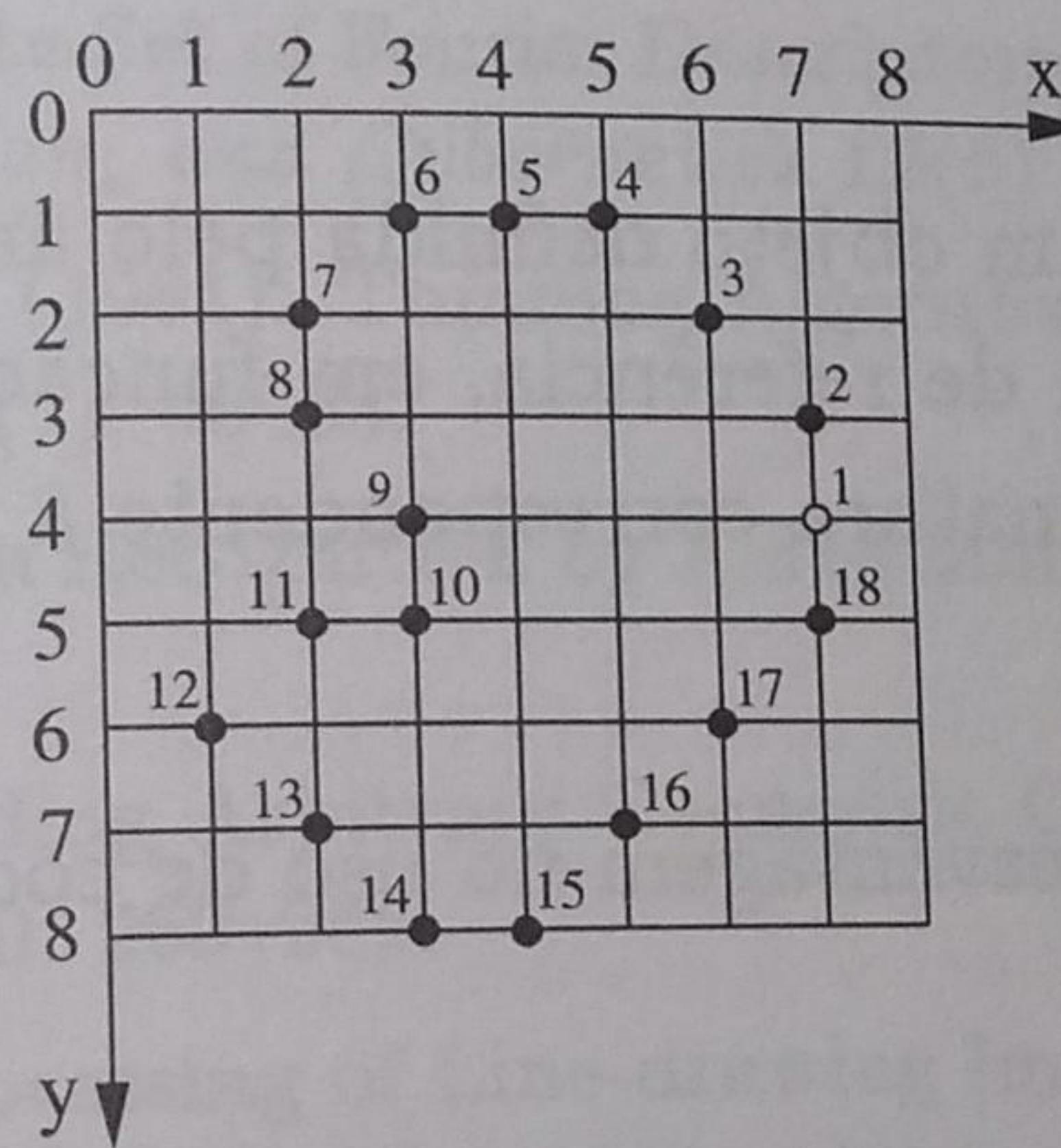


Figura 7.36: Contorno de um objeto.

4. Descreva três descritores invariantes quanto às transformações geométricas de escala, translação e rotação.
5. Escreva um programa para calcular a área englobada por um contorno definido pelo código da cadeia. Considere que os pixels do contorno estão conectados por meio da vizinhança-4.
6. Dada uma elipse com eixos maior a e menor b , mostre o gráfico de sua assinatura definida como a distância da borda ao centróide em função do ângulo.
7. Escreva um programa para implementar uma técnica de afinamento em imagens tridimensionais.
8. Cite vantagens e desvantagens no uso de modelos deformáveis para representar objetos em imagens.
9. Descreva um conjunto de descritores que poderiam ser utilizados em um sistema de reconhecimento automático de placas de veículos a partir de imagens de vídeo.

- 282**

Análise de Aplicações

 10. A precisão com que uma curva pode ser representada por meio de uma seqüência de segmentos de reta depende do processo utilizado na seleção dos pontos. Discuta quais critérios poderiam ser utilizados na escolha de pontos para a aproximação da curva.
 11. Encontre o esqueleto do objeto mostrado na figura 7.37, a partir da transformada do eixo médio.

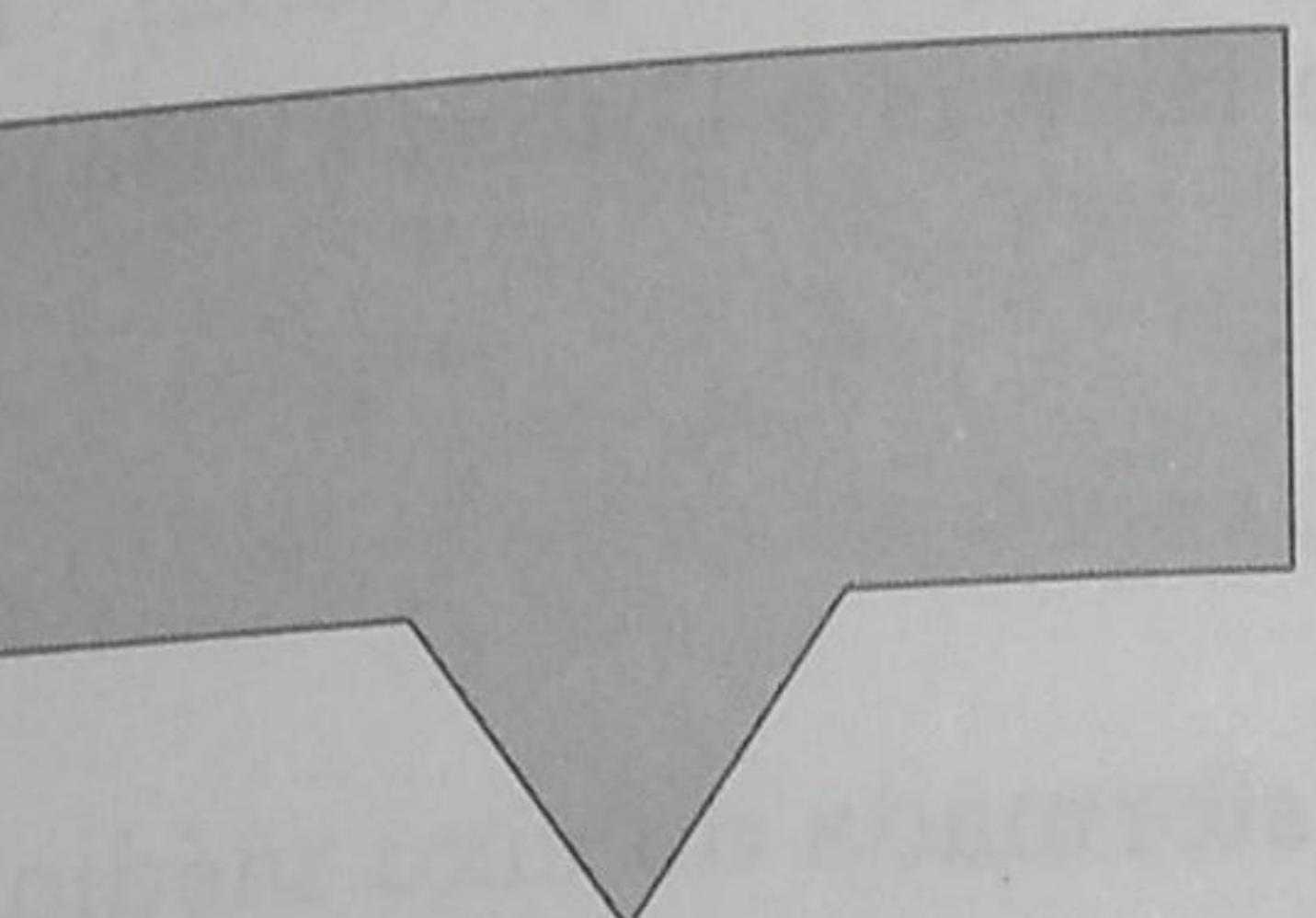


Figura 7.37: Esqueleto de objeto.

12. Prove que a compacidade, definida na equação 7.28, é mínima para um círculo.
 13. Dê dois exemplos de descritores que poderiam ser utilizados para caracterizar objetos tridimensionais.
 14. Considere a assinatura de um objeto definida pelo ângulo entre uma reta tangente à borda do objeto e uma reta de referência, em função da posição ao longo da borda. Desenhe o diagrama da assinatura correspondente a um círculo, um quadrado e um triângulo equilátero.
 15. Cite uma vantagem e uma desvantagem no uso de códigos da cadeia para representar bordas de objetos.

Referências Bibliográficas

- Abu-Mostafa, Y. S. e Psaltis, D. (1984). Recognitive Aspects of Moment Invariants, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**(6): 698–706.

Abu-Mostafa, Y. S. e Psaltis, D. (1985). Image Normalization by Complex Moments, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **7**(1): 46–55.

Alt, F. L. (1962). Digital Pattern Recognition by Moments, *Journal of the ACM* **9**(2): 240–258.

Amini, A. A., Weymouth, T. e Jain, R. C. (1990). Using Dynamic Programming for Solving Variational Problems in Vision, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **12**(9): 855–867.

Aurenhammer, F. (1991). Voronoi Diagrams - A Survey of a Fundamental Geometric Data Structure, *ACM Computing Surveys* **23**(3): 345–405.

Bhatia, A. B. e Wolf, E. (1954). On the Circle Polynomials of Zernike and Related Orthogonal Sets, *Proceedings of Cambridge Philosophical Society*, Vol. 50, Londres, Reino Unido, pp. 40–48.

Biswas, G. e Dubes, R. C. (1984). Some Experiments in Two-Dimensional Grammatical Inference, *Pattern Recognition Letters* **2**: 173–177.

Blum, H. (1967). A Transformation for Extracting New Descriptors for Shape, in W. Wathen-Dunn (ed.), *Models for the Perception of Speech and Visual Form*, MIT Press, Cambridge, MA, Estados Unidos.

- Brigger, P. e Unser, M. (1998). Multi-Scale B-Spline Snakes for General Contour Detection, Proceedings of the SPIE Conference on Mathematical Imaging: Wavelet Applications in Signal and Image Processing VI, Vol. 3458, San Diego, CA, Estados Unidos, pp. 92-102.
- Bunke, H. e Sanfeliu, A. (1990). Syntactic and Structural Pattern Recognition: Theory and Applications.
- Cash, G. L. e Hatamian, M. (1987). Optical Character Recognition by the Method of Moments, Computer Vision, Graphics, and Image Processing 39: 291-310.
- Chomsky, N. (1957). *Syntactic Structures*, The Hague: Mouton & Co.
- Chuang, J. H., Tsai, C. H. e Ko, M. C. (2000). Skeletonization of Three-Dimensional Object Using Generalized Potential Field, IEEE Transactions on Pattern Analysis and Machine Intelligence 22(11): 1241-1251.
- Cohen, L. D. e Cohen, I. (1992). Using Deformable Surfaces to Segment 3D Images and Infer Differential Structures, Computer Vision Graphics and Image Processing: Image Understanding 56(2): 242-263.
- Cohen, L. D. e Cohen, I. (1993). Finite-element Methods for Active Contour Models and Balloons for 2-D and 3-D Images, IEEE Transactions on Pattern Analysis and Machine Intelligence 15(11): 1131-1147.
- Costa, L. F. e Cesar Jr., R. M. (2001). *Shape Analysis and Classification: Theory and Practice*, CRC Press, Boca Raton, FL, Estados Unidos.
- Crimmins, T. R. (1982). A Complete Set of Fourier Descriptors for Two-Dimensional Shapes, IEEE Transactions on Systems, Man, and Cybernetics 12(6): 848-855.
- Wies, E. R. e Plummer, A. P. N. (1981). Thinning Algorithms: A Critique and a New Methodology, Pattern Recognition 14(1): 53-63.
- Ster, J. e Suk, T. (1993). Pattern Recognition by Affine Moment Invariants, Pattern Recognition 26(1): 167-174.
- Li, H. (1961). On the Encoding Arbitrary Geometric Configurations, IRE Transactions on Electronic Computers EC-10: 260-268.
- Li, L. (1974). Computer Processing of Line-drawing Images, Computer Surveys 6(1): 57-97.
- Li, L. (1982). *Syntactic Pattern Recognition and Applications*, Prentice Hall, Inc., Englewood Cliffs, NJ, Estados Unidos.
- Gagvani, N. e Silver, D. (1999). Parameter-Controlled Volume Thinning, Graphical Models and Image Processing 61(3): 149-164.
- Geman, S. e Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images, IEEE Transactions on Pattern Analysis and Machine Intelligence 6(6): 721-741.
- Gonzalez, R. C. e Woods, R. E. (2002). *Digital Image Processing*, Prentice Hall, Inc., Upper Saddle River, NJ, Estados Unidos.
- Gunn, S. R. e Nixon, M. S. (1997). A Robust Snake Implementation: A Dual Active Contour, IEEE Transactions on Pattern Analysis and Machine Intelligence 19(1): 63-68.
- Henderson, T. C. e Samal, A. (1986). Shape Grammar Compilers, Pattern Recognition 19(4): 279-288.
- Holt, C. M., Stewart, A., Clint, M. e Perrott, R. H. (1987). An Improved Parallel Thinning Algorithm, Communications of the ACM 30(2): 156-160.
- Hu, M. K. (1962). Visual Pattern Recognition by Moment Invariants, IRE Transactions on Information Theory IT-8: 179-187.
- Ivins, J. e Porritt, J. (1994). Statistical Snakes: Active Region Models, Proceedings of Fifth British Machine Vision Conference, Vol. 2, York, Reino Unido, pp. 377-386.
- Jain, A. K. (1989). *Fundamentals of Digital Image Processing*, Prentice Hall, Inc., Englewood Cliffs, NJ, Estados Unidos.

- Kass, M., Witkin, A. P. e Terzopoulos, D. (1988). Snakes: Active Contour Models, *International Journal of Computer Vision* 1(4): 321-331.
- Kim, W. Y. e Kim, Y. S. (2000). A Region-based Shape Descriptor Using Zernike Moments, *Signal Processing: Image Communication* 16(1-2): 95-102.
- Kuhl, F. P. e Giardina, C. R. (1982). Elliptic Fourier Features of a Closed Contour, *Computer Graphics and Image Processing* 18: 236-258.
- Kwok, P. (1988). A Thinning Algorithm by Contour Generation, *Communications of the ACM* 31(11): 1314-1324.
- Lam, K.-M. e Yan, H. (1994). Fast Greedy Algorithm for Active Contours, *Electronic Letters* 30(1): 21-23.
- Lam, L., Lee, S.-W. e Suen, C. Y. (1992). Thinning Methodologies: A Comprehensive Survey, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14(9): 869-885.
- Li, X. e Zhiying, Z. (1988). Group Direction Difference Chain Codes for the Representation of the Border, *Digital and Optical Shape Representation and Pattern Recognition*, Orlando, FL, Estados Unidos, pp. 372-376.
- Liao, S. X. e Pawlak, M. (1996). On Image Analysis by Moments, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18(3): 254-266.
- Lin, C. C. e Chellappa, R. (1987). Classification of Partial 2-D Shapes Using Fourier Descriptors, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 9(5): 686-690.
- Loncaric, S. (1998). A survey of Shape Analysis Techniques, *Pattern Recognition* 31(8): 983-1001.
- Ma, C. M. (1994). On Topology Preservation in 3D Thinning, *Computer Vision, Graphics and Image Processing* 59(3): 328-339.
- Ma, C. M. e Sonka, M. (1996). A Fully Parallel 3D Thinning Algorithm and Its Applications, *Computer Vision and Image Understanding* 64(3): 420-433.
- Maitra, S. (1979). Moment Invariants, *Proceedings of IEEE* 67(4): 697-699.
- Malladi, R., Sethian, J. A. e Vemuri, B. C. (1995). Shape Modeling with Front Propagation: A Level Set Approach, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17(2): 158-175.
- McInerney, T. e Terzopoulos, D. (1995). Topologically Adaptable Snakes, *Proceedings of the Fifth International Conference on Computer Vision*, Cambridge, MA, Estados Unidos, pp. 840-845.
- Menet, S., Saint-Marc, P. e Medioni, G. (1990). B-Snakes: Implementation and Application to Stereo, *Proceedings of the DARPA Image Understanding Workshop*, Pittsburgh, PA, Estados Unidos, pp. 720-726.
- Ohya, T., Iri, M. e Murota, K. (1984). Improvements of the Incremental Method for the Voronoi Diagram with Computational Comparison of Various Algorithms, *Journal of the Operational Research Society* 27: 306-337.
- O'Rourke, J. (1993). *Computational Geometry in C*, Cambridge University Press.
- Pavlidis, T. (1977). *Structural Pattern Recognition*, Springer-Verlag, Berlim, Alemanha.
- Pavlidis, T. (1980a). A Thinning Algorithm for Discrete Binary Images, *Computer Graphics and Image Processing* 13: 142-157.
- Pavlidis, T. (1980b). Algorithms for Shape Analysis of Contours and Waveforms, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2(4): 301-312.
- Persoon, E. e Fu, K. S. (1977). Shape Discrimination Using Fourier Descriptors, *IEEE Transactions on Systems, Man, and Cybernetics* 7(3): 170-179.
- Persoon, E. e Fu, K. S. (1986). Shape Discrimination Using Fourier Descriptors, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8(3): 388-397.
- Pratt, W. K. (1991). *Digital Image Processing*, John Wiley & Sons, Inc., Nova York, NY, Estados Unidos.

- Preparata, F. P. e Shamos, M. I. (1985). *Computational Geometry - An Introduction*, Springer-Verlag, Nova York.
- Rutovitz, D. (1966). Pattern Recognition, *Journal of the Royal Statistical Society* 129: 504-530.
- Schalkoff, R. J. (1989). *Digital Image Processing and Computer Vision*, John Wiley & Sons, Inc., Nova York, NY, Estados Unidos.
- Shridhar, M. e Badreldin, A. (1984). High Accuracy Character Recognition Algorithm using Fourier and Topological Descriptors, *Pattern Recognition* 17(5): 515-524.
- Sonka, M., Hlavac, V. e Boyle, R. (1999). *Image Processing, Analysis, and Machine Vision*, Brooks/Cole Publishing Company, Pacific Grove, CA, Estados Unidos.
- Stentiford, F. W. M. e Mortimer, R. G. (1983). Some New Heuristics for Thinning Binary Hand-printed Characters for OCR, *IEEE Transactions on Systems, Man, and Cybernetics SMC-13*(1): 81-84.
- Storvik, G. (1994). A Bayesian Approach to Dynamic Contours Through Stochastic Sampling and Simulated Annealing, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16(10): 976-986.
- Teague, M. R. (1980). Image Analysis via the General Theory of Moments, *Journal of the Optical Society of America* 70: 920-930.
- Truve, S. (1990). Image Interpretation Using Multi-Relational Grammars, *International Conference on Computer Vision*, pp. 146-155.
- Tsai, W. H. e Yu, S. S. (1985). Attributed String Matching with Merging for Shape Recognition, *IEEE transactions on Pattern Analysis and Machine Intelligence* 7(4): 453-462.
- Wallace, T. P. e Wintz, P. A. (1980). An Efficient Three-Dimensional Aircraft Recognition Algorithm Using Fourier Descriptors, *Computer Graphics and Image Processing* 13(1): 99-126.
- Williams, D. J. e Shah, M. (1992). A Fast Algorithm for Active Contours and Curvature Estimation, *Computer Vision Graphics and Image Processing: Image Understanding* 55(1): 14-26.
- Wong, R. Y. e Hall, E. L. (1978). Sequential Hierarchical Scene Matching, *IEEE Transactions on Computers* C-27(4): 359-366.
- Xu, C. e Prince, J. L. (1998). Snakes, Shapes, and Gradient Vector Flow, *IEEE Transactions on Image Processing* 7(3): 359-369.
- Zahn, C. T. e Roskies, R. Z. (1972). Fourier Descriptors for Plane Closed Curves, *IEEE Transactions on Computers* 21(3): 269-281.
- Zernike, F. (1934). Beugungstheorie des Schneidenverfahrens und seiner verbesserten Form, der Phasenkontrastmethode (Diffraction theory of the cut procedure and its improved form, the phase contrast method), *Physica* 1: 689-704.
- Zhang, D. e Lu, G. (2004). Review of Shape Representation and Description Techniques, *Pattern Recognition* 37(1): 1-19.
- Zhang, D. S. e Lu, G. (2002). A Comparative Study of Fourier Descriptors for Shape Representation and Retrieval, *Proceedings Fifth Asian Conference on Computer Vision*, Melbourne, Austrália, pp. 646-651.
- Zhang, T. Y. e Suen, C. Y. (1984). A Fast Parallel Algorithm for Thinning Digital Patterns, *Communications of the ACM* 27(3): 236-239.