Introdução ao Processamento e Análise de Imagens

Neucimar J. Leite

IC-UNICAMP

1 Histograma

O histograma dos níveis de cinza de uma imagem **f**, $h_f(z)$, constitui uma *operação global* indicando a freqüência de ocorrência dos níveis de cinza ou brilho z de **f** (função densidade de probabilidade). Assim, o histograma de uma imagem com L níveis de cinza é representado por um vetor de L elementos, cujos valores podem ser definidos a partir do seguinte algoritmo:

Início
  h(f(x,y))) = 0 {*zera contadores de níveis de cinza* }
  Para cada valor f(x,y) **faça**
   h(f(x,y)) = h(f(x,y)) + 1
  Fim-Para
Fim

O histograma dá uma idéia global da *dinâmica* de uma cena e tem aplicações em inúmeras transformações de imagens, tais como filtragem, segmentação, reconhecimento de padrões etc. As figuras 1 e 2 apresentam imagens com seus respectivos histogramas.

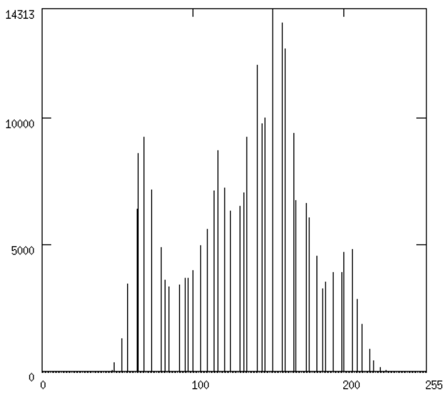



Figura 1:

Observe que um mesmo histograma pode estar associado a diferentes imagens e que a sua informação é invariante a operações de rotação e translação (objetos se movendo num mesmo background) da imagem original. Um histograma pode ser

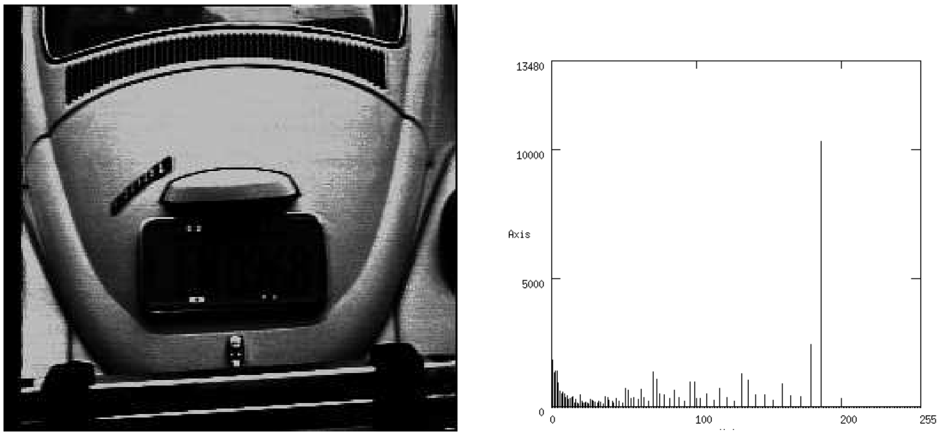


Figura 2:

definido igualmente para imagens multiespectrais. No caso de imagens coloridas, por exemplo, podemos definir um histograma para cada banda espectral ou um histograma 3-D referente às componentes RGB da imagem. Neste caso, a intensidade dos voxels (pixels num volume) é proporcional à freqüência de aparição dos níveis de cinza na cena original.

2 Transformações radiométricas ou de escala de cinza

As transformações radiométricas independem da localização dos pixels na imagem e podem, de maneira geral (e em termos de implementação), ser representadas por uma operação do tipo *look-up table* (LUT) que transforma um pixel de nível de cinza g_i em um nível de cinza g_f. Seja G_i a escala de cinza da imagem inicial $\mathbf{f}$: $\mathrm{G}_i = [0,1,...,\mathrm{N}_i]$, e G_f a escala de cinza da imagem resultante $\mathbf{g}$: $\mathrm{G}_f = [0,1,...,\mathrm{N}_f]$. Uma LUT é uma aplicação $\mathbf{f}(\mathrm{G}_i) \rightarrow \mathbf{g}(\mathrm{G}_f)$ tal que

$$\forall g_i \in \mathbf{G}_i, \quad \exists\ g_f \in \mathbf{G}_f, \quad g_f = r(g_i) \tag{1}$$

Alguns exemplos de funções r de transformações radiométricas são indicados na Fig. 3.

A função r_1 define o *negativo* ou *complemento* da imagem original; a função r_2 realça o *contraste* entre os pontos p_1 e p_2, e r_3 corresponde a uma *limiarização*

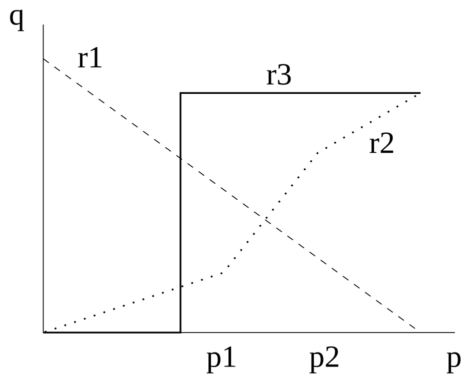


Figura 3:

(thresholding) dos pixels de entrada, resultando numa imagem *binária* (preta e branca). A Fig. 4 ilustra estas transformações para r_1, r_2 e r_3, respectivamente.

Figura 4:

2.1 Equalização histogrâmica

A equalização ou modelagem histogrâmica é uma transformação radiométrica que visa aumentar a dinâmica dos níveis de cinza melhorando, por exemplo, o contraste de imagens digitalizadas sob péssimas condições de iluminação.

De modo geral, o que se procura é obter um mapeamento não-linear dos níveis de cinza da imagem de entrada de tal forma que a imagem resultante contenha uma distribuição mais uniforme dos seus níveis de cinza, ou seja, um histograma planar (Fig. 5).

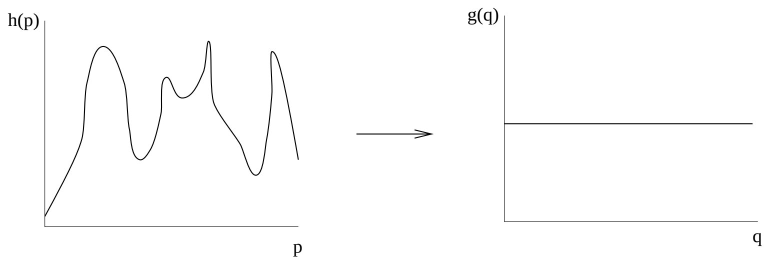


Figura 5:

Existem várias técnicas de equalização histogrâmica cujas abordagens consideram inicialmente funções contínuas e não discretas. A seguir, ilustraremos uma destas técnicas apresentada em [Woo93].

Seja f uma variável no intervalo [0, 1] (f representa, por exemplo, a intensidade normalizada dos níveis de cinza de uma imagem de tal forma que $0 \leq f(x,y) \leq 1$). Uma transformação T, no intervalo [0, 1], é tal que

$$g = T(f) \text{ (ou } g(x,y) = T(f(x,y)), \quad \text{para o caso de nossa imagem discreta)} \qquad (2)$$

onde g é o valor resultante obtido a partir do nível de entrada f. Visando a monotonicidade e preservação da escala de cinza original, a transformação T deve satisfazer as seguintes condições (Fig. 6):

- T deve ser monotonicamente crescente no intervalo [0, 1].
- $0 \leq T(f) \leq 1$ para $0 \leq f \leq 1$.

A primeira condição preserva a ordem na seqüência da escala de cinza, variando do preto ao branco, e a segunda garante que o intervalo desta escala continua o mesmo.

A transformação inversa, que também satisfaz as condições acima, é denotada por:

$$f = T^{-1}(g) \qquad (3)$$

Considerando f e g como variáveis contínuas, os níveis de cinza das imagens podem ser caracterizados por suas funções densidade de probabilidade (ou histogramas) denotadas por $p_f(f)$ e $p_g(g)$, respectivamente.

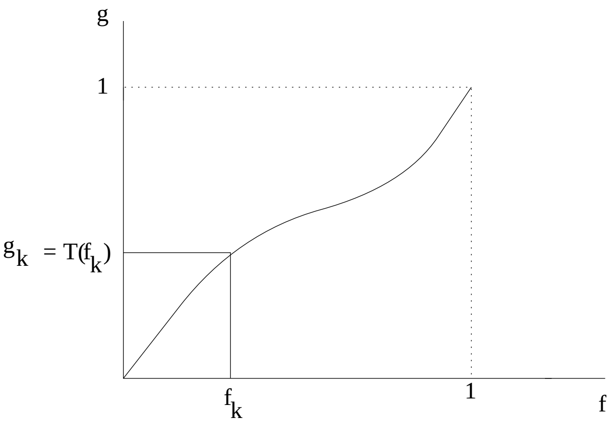


Figura 6:

Consideremos, agora, a seguinte função de transformação

$$g = T(f) = \int_0^f p_f(w)dw \qquad 0 \leq f \leq 1, \tag{4}$$

representando a função de distribuição cumulativa (FDC) de f. Esta função é monotonicamente crescente e varia de 0 a 1 em função de f.

Exercício: Mostre que a função densidade de probabilidade $p_g(g)$ resultante da transformação acima é, como desejamos, uniforme no intervalo [0, 1].

A partir do exerício acima, concluímos que se T(f) é uma FDC, então esta função pode ser empregada na definição de uma nova imagem com distribuição mais uniforme dos níveis de cinza. Como mencionado anteriormente, o efeito aqui será visível no que concerne o aumento do contraste dos níveis de cinza de uma imagem apresentando baixa dinâmica.

2.1.1 O caso discreto

Dada uma imagem contendo $n = M \times N$ pixels, assumindo valores discretos $k = 0, 1, ..., L-1$, a extensão destes conceitos para o caso discreto se dá a partir das seguintes definições de probabilidades:

$$p_f(f_k) = \frac{n_k}{n} \qquad 0 \leq f_k \leq 1, \tag{5}$$

onde n_k é o número de aparições do nível k, e $p_f(f_k)$ é a probabilidade deste nível. Finalmente, a forma discreta da Eq. 4 é

$$g_k = T(f_k) = \sum_{j=0}^{k} \frac{n_j}{n} = \sum_{j=0}^{k} p_f(f_j), \qquad 0 \leq f_k \leq 1 \quad e \quad k = 0, 1, ..., L-1. \tag{6}$$

Esta equação define a função de transformação discreta de equalização histográmica. A Fig. 7 ilustra esta operação.

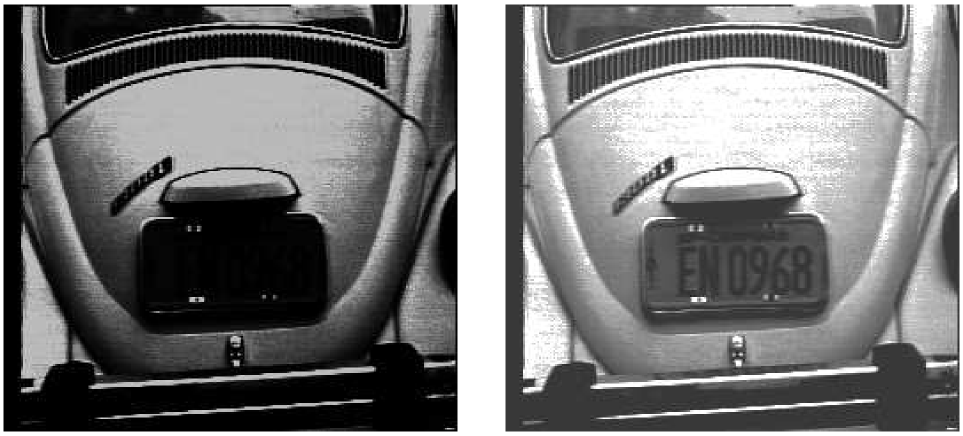


Figura 7:

3 Operações lineares: a Transformada de Fourier

A Transfomada de Fourier (FT) é uma ferramenta largamente empregada em processamento de sinais, de modo geral, e em processamento de imagens, em particular. Denominada assim em homenagem ao físico francês Jean Baptiste Fourier (1768-1830), a FT decompõe um sinal em suas componentes elementares seno e cosseno. A FT aplicada a uma imagem no *domínio espacial* gera uma informação no *domínio da freqüência*, em que cada ponto, definido por um vetor do tipo $(k.cosseno, k.seno)$, representa uma dada freqüência contida no domínio espacial da imagem.

As aplicações referentes à FT são inúmeras: filtragem, segmentação, reconhecimento de padrões, descrição de imagens, compressão e recontrução constituem algumas delas.

3.1 Fundamentos matemáticos

As próximas seções apresentam alguns conceitos matemáticos básicos relacionados com a definição de transfomadas lineares [Kak82].

3.1.1 Operações lineares em imagens

Seja Γ uma operação que transforma uma imagem $\mathbf{f}$ em uma outra imagem $\mathbf{g}$, isto é, $\Gamma : \mathbf{f} \rightarrow \mathbf{g}$. Γ é um operador linear se:

$$\Gamma[a\mathbf{f} + b\mathbf{g}] = a\Gamma[\mathbf{f}] + b\Gamma[\mathbf{g}], \tag{7}$$

para $\mathbf{f}$ e $\mathbf{g}$ imagens e a, b constantes quaisquer.

Uma imagem 2-D pode ser vista como uma coleção de *fontes pontuais* concentradas no espaço bidimensional. No caso linear, a análise da saída de uma fonte pontual pode ser utilizada para definir a *função de espalhamento pontual* de Γ associada à *função de transferência* do sistema.

Seja a função definindo a seguinte seqüência de imagens

$$rect(x, y) = \begin{cases} 1, & \text{para } \mid \text{x} \mid \leq \frac{1}{2}, \mid \text{y} \mid \leq \frac{1}{2} \\ 0, & \text{caso contrário} \end{cases} \tag{8}$$

e seja

$$\delta_n = n^2 rect(nx, ny), \quad n = 1, 2, ... \quad , \tag{9}$$

isto é, δ_n é zero fora da área $1/n \times 1/n$ definida por $|x| \leq 1/2n$, $|y| \leq 1/2n$, e tem valor constante n^2 dentro desta área. Assim,

$$\int\int_{-\infty}^{+\infty} \delta_n(x, y) dxdy = 1, \quad \text{para } n \text{ qualquer.} \tag{10}$$

A função definida pelo pseudo-limite da Eq. 10 acima, denotado por δ, quando $n \to \infty$, é conhecida como *função Delta de Dirac*. Neste caso, $\delta(x, y) = 0$ para todo (x,y) diferente de (0,0) e infinito neste domínio. Observe que $\delta(-x, -y) = \delta(x, y)$. Baseado em 10, definimos

$$\int\int_{-\infty}^{+\infty} \delta(x, y) dxdy = 1, \tag{11}$$

e podemos ver facilmente que

$$\int\int_{-\infty}^{+\infty} g(x, y)\delta(x, y) dxdy = g(0, 0), \tag{12}$$

ou ainda

$$\int\int_{-\infty}^{+\infty} g(x, y)\delta(x - \alpha, y - \beta) dxdy = g(\alpha, \beta) \tag{13}$$

que corresponde à *propriedade de deslocamento* da função. Outra propriedade interessante é dada por

$$\int\int_{-\infty}^{+\infty} e^{j2\pi(ux+vy)} dudv = \delta(x, y), \tag{14}$$

em que $j = \sqrt{-1}$ e $e^{-j\theta} = cos\theta - jsen\theta$, θ é a *variável de freqüência.*

A definição da função Delta de Dirac é considerada em várias aplicações como, por exemplo, na abordagem de problemas relacionados com quantização, filtragem (definição da função característica de um sistema), reconstrução a partir de projeções etc.

4 Invariância ao deslocamento

Seja a função de espalhamento $h(x,y)$ que define a saída de um ponto (x,y) de uma imagem. Uma operação linear é dita invariante ao deslocamento se a resposta à fonte pontual $\delta(x-\alpha, y-\beta)$, localizada no plano 2-D, é dada por $h(x-\alpha, y-\beta)$ (a saída é igualmente deslocada de α e β no plano).

Consideremos $f(x,y)$ uma imagem de entrada qualquer de um sistema. Pela Eq. 13, esta imagem pode ser vista como uma soma linear de fontes pontuais, isto é,

$$f(x,y) = \int\int_{-\infty}^{+\infty} f(\alpha,\beta)\delta(\alpha-x,\beta-y)d\alpha d\beta \tag{15}$$

Assumindo nosso operador linear Γ invariante ao deslocamento, obtemos

$$\Gamma[f(x,y)] = \Gamma[\int\int_{-\infty}^{+\infty} f(\alpha,\beta)\delta(\alpha-x,\beta-y)d\alpha d\beta] = \int\int_{-\infty}^{+\infty} f(\alpha,\beta)\Gamma[\delta(\alpha-x,\beta-y)]d\alpha d\beta, \tag{16}$$

o que significa que a resposta ao somatório de excitações é igual ao somatório da resposta de cada excitação.

Como a resposta a $\delta(\alpha-x,\beta-y) = \delta(x-\alpha,y-\beta)$, que representa uma fonte pontual em (α,β), é dada por $h(x-\alpha,y-\beta)$, então

$$\Gamma[f(x,y)] = g(x,y) = \int\int_{-\infty}^{+\infty} f(\alpha,\beta)h(x-\alpha,y-\beta)d\alpha d\beta \tag{17}$$

define a *convolução* de $\mathbf{f}$ e $\mathbf{h}$, $\mathbf{f} * \mathbf{h}$. Observe que $\mathbf{f} * \mathbf{h}$ é igual a $\mathbf{h} * \mathbf{f}$. A Fig. 8 ilustra esta operação.

5 Análise de Fourier

A Transformada de Fourier F(u,v) de uma função 2-D, f(x,y), é dada por

$$F(u,v) = \int\int_{-\infty}^{+\infty} f(x,y)e^{-j2\pi(ux+vy)}dxdy \tag{18}$$

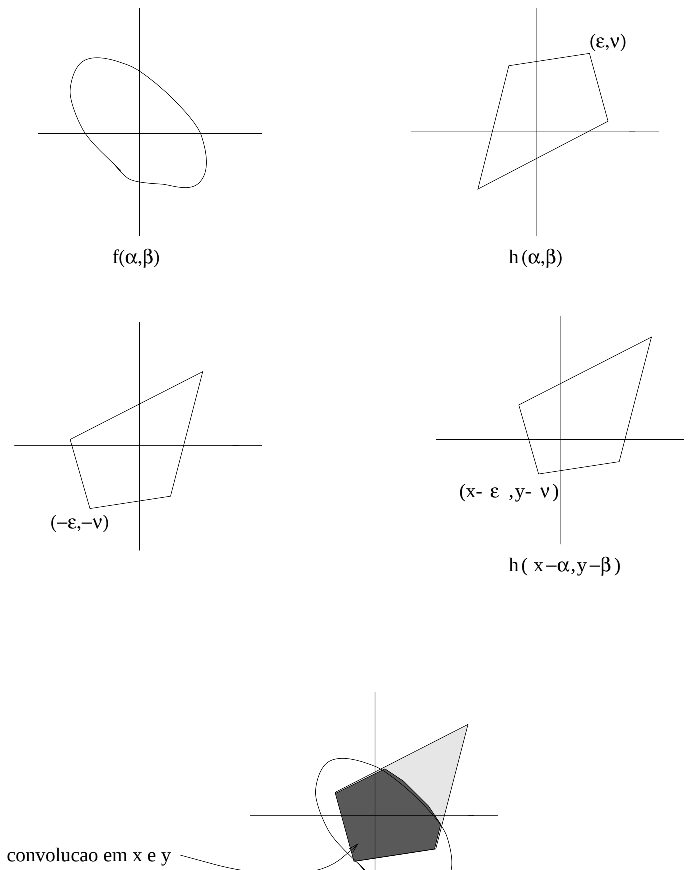


Figura 8: [Kak82]

Observe que a transformação apresenta valores complexos constituídos de uma parte real, R, e imaginária, I. Assim:

$$F(u,v) = R(u,v) + jI(u,v) \tag{19}$$

ou

$$F(u,v) = |F(u,v)|e^{-j\phi(u,v)} \tag{20}$$

em que

$$|F(u,v)| = \sqrt{R^2(u,v) + I^2(u,v)} \tag{21}$$

corresponde ao *espectro de Fourier*, e

$$\phi(u,v) = tg^{-1}[\frac{I(u,v)}{R(u,v)}] \tag{22}$$

define o *ângulo de fase*. A *energia* E de $f(x,y)$ é dada por

$$E(u,v) = |F(u,v)|^2 \tag{23}$$

Exercício 1: Mostre que a FT da função $rect(x,y)$ é a função $sinc(u,v)$ (Fig. 9), ou seja

$$F(u,v) = \frac{sen\pi u}{\pi u}\frac{sen\pi v}{\pi v} \tag{24}$$

Exercício 2: Generalize o resultado anterior mostrando que a FT de $rect(nx,ny)$ é

$$(1/n^2)sinc(u/n, v/n) \tag{25}$$

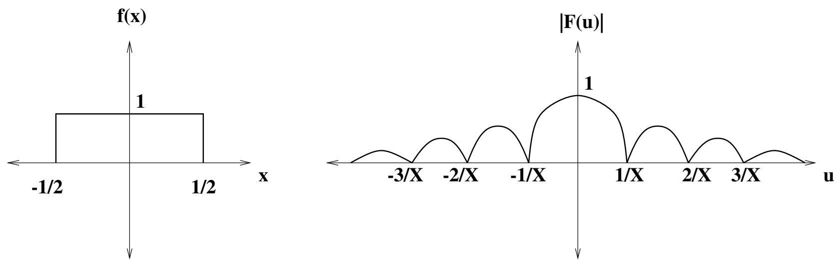


Figura 9:

Uma condição necessária para a existência de uma FT é que o sinal não tenha descontinuidades infinitas. Naturalmente, a função Delta de Dirac não obedece a esta condição. Na realidade, ela representa o limite da seqüência de funções $\delta_n = n^2 rect(nx,ny)$ bem comportada que, por sua vez, possui uma FT. Portanto, podemos definir o limite ($n \longrightarrow \infty$) desta FT como sendo a FT da função Delta de Dirac, ou seja, quando

$$f(x,y) = \delta(x,y) \tag{26}$$

então

$$F[f(x,y)] = F(u,v) = \lim_{n\to\infty} sinc(u/n, v/n) = 1 \tag{27}$$

Multiplicando ambos os lados da Eq. 18 por $e^{j2\pi(ux+vy)}$, integrando em relação a u e v, e depois considerando a equação 14, obtemos:

$$\int\int_{-\infty}^{+\infty} F(u,v)e^{-j2\pi(u\alpha+v\beta)}dudv = f(\alpha,\beta) \tag{28}$$

ou

$$f(x,y) = \int\int_{-\infty}^{+\infty} F(u,v)e^{j2\pi(ux+vy)}dudv \tag{29}$$

Este último resultado define a *FT inversa* que expressa um sinal f como uma combinação linear de senos e cossenos de magnitudes e fases diferentes. Esta equação, juntamente com a *FT direta* (Eq. 18), constitui o par de transformadas de Fourier.

5.1 Algumas propriedades da FT

Seja $\mathcal{F}[f]$ a FT de uma função $f(x,y)$, isto é, $\mathcal{F}[f]$ = F(u,v). Algumas propriedades interessantes da FT são [Pap62], [Kor88]:

1. Linearidade:

$$\begin{aligned}\mathcal{F}[af_1(x,y) + bf_2(x,y)] &= a\mathcal{F}[f_1(x,y)] + b\mathcal{F}[f_2(x,y)] \\ &= aF_1(u,v) + bF_2(u,v)\end{aligned}$$

2. Escalonamento:

$$\mathcal{F}[f(\alpha x, \beta y)] = \frac{1}{\alpha\beta}F(\frac{u}{\alpha}, \frac{v}{\beta})$$

3. Deslocamento no espaço:

$$\mathcal{F}[f(x-\alpha, y-\beta)] = F(u,v)e^{(-j2\pi(u\alpha+v\beta))}$$

4. Deslocamento na freqüência:

$$\mathcal{F}[e^{(j2\pi(u_0x+v_0y))}f(x,y)] = F(u-u_0, v-v_0)$$

5. Rotação de 180^0:

$$\mathcal{F}[\mathcal{F}[f(x,y)]] = f(-x,-y)$$

6. Convolução:

$$\begin{aligned}\mathcal{F}[\textstyle\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_1(\alpha,\beta)f_2(x-\alpha,y-\beta)d\alpha d\beta] &= \mathcal{F}[f_1(x,y)]\mathcal{F}[f_2(x,y)] \\ &= F_1(u,v)F_2(u,v)\end{aligned}$$

Este último resultado é de grande importância em processamento de sinais. Ele mostra que a convolução de duas funções no domínio espacial é equivalente à operação de multiplicação no domínio da freqüência. A equivalência para a convolução no domínio da freqüência pode ser expressa por

$$\mathcal{F}[f_1(x,y)f_2(x,y)] = \textstyle\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} F_1(u-s,v-t)F_2(s,t)dsdt$$

6 Transformadas discretas: o caso geral

Seja $\mathbf{f}$ uma imagem de dimensão $M \times N$ representada pela seguinte matriz de pixels:

$$[f]_{M\times N} = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \vdots & \vdots & \cdots & \vdots \\ f(M,0) & f(M,1) & \cdots & f(M,N) \end{bmatrix}$$

Uma transformada discreta geral de $[f]_{M\times N}$ é uma matriz $[F]_{M\times N}$ dada pelo seguinte produto

$$[F] = [P][f][Q] \tag{30}$$

onde $[P]$ e $[Q]$ são matrizes não-singulares (possuem inversa, determinante $\neq 0$) de dimensão $M \times M$ e $N \times N$, respectivamente:

$$[F]_{M\times N} = [P]_{M\times M}[f]_{M\times N}[Q]_{N\times N} \tag{31}$$

A transformada $[F]_{M\times N}$ pode ser escrita, ainda, da seguinte forma

$$F(u,v) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} P(u,m)f(m,n)Q(n,v) \quad \text{para u=0,1, ... M-1 e v=0,1, ... N-1} \tag{32}$$

Sejam, agora, $[P]^{-1}$ e $[Q]^{-1}$ as matrizes inversas de $[P]$ e $[Q]$, respectivamente. É fácil mostrar que a transformada inversa de $[F]$ é dada por

$$[f] = [P]^{-1}[F][Q]^{-1} \tag{33}$$

Exercício: Verifique que se $[P]$ e $[Q]$ são matrizes reais, simétricas e ortogonais, então

$$[F] = [P][f][Q], \qquad [f] = [P][F][Q] \tag{34}$$

6.1 Exemplo de uma caso particular: a Transformada Discreta de Fourier

Seja a matriz de transformação $[\Phi_{J\times J}]$ com um pixel na posição (m,n) dado por

$$\tfrac{1}{J}e^{-j(2\pi/J)mn}, \qquad m,n = 0,1,\cdots,J-1 \tag{35}$$

e

$$[P] = [\Phi_{M\times M}] \qquad e \qquad [Q] = [\Phi_{N\times N}] \tag{36}$$

A *Transformada Discreta de Fourier - DFT* de uma matriz **f** é dada por

$$[F] = [P][f][Q] \tag{37}$$

Considerando as definições de $[P]$ e $[Q]$ acima, e a expansão em 32, obtemos

$$F(u,v) = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f(m,n)e^{-j2\pi(\frac{mu}{M}+\frac{nv}{N})} \tag{38}$$

para $u = 0,1,\cdots,M-1$ e $v = 0,1,\cdots,N-1$.

Multiplicando-se $[\Phi_{J\times J}]$ pela matriz $[\Phi_{J\times J}]'$ definida por

$$e^{j(2\pi/J)mn}, \tag{39}$$

obtemos a matriz identidade, o que significa naturalmente que $[\Phi_{J\times J}]' = [\Phi_{J\times J}]^{-1}$.

Considerando-se $[P] = [\Phi_{M\times M}]$ e $[Q] = [\Phi_{N\times N}]$, então $[P]^{-1} = [\Phi_{M\times M}]^{-1} e [Q]^{-1} = [\Phi_{N\times N}]^{-1}$. Assim, a *Transformada Discreta Inversa de Fourier* é dada por

$$[f] = [P]^{-1}[F][Q]^{-1} \tag{40}$$

ou na sua forma expandida

$$f(m,n) = \sum_{u=0}^{M-1}\sum_{v=0}^{N-1} F(u,v)e^{j2\pi(\frac{mu}{M}+\frac{nv}{N})} \tag{41}$$

para $m = 0, 1, \cdots, M-1$ e $n = 0, 1, \cdots, N-1$.

6.2 DFT: visualização e implementação

Como vimos, a FT representa de forma bastante diferente a informação contida no domínio espacial. A Fig. 10 ilustra, para o caso 1-D, o modelo do processo reversível associado ao par de transformadas *direta* e *inversa* de Fourier.

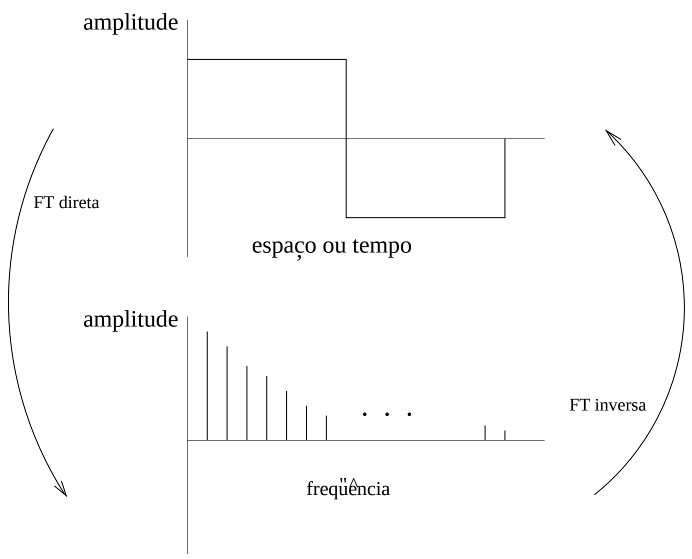


Figura 10:

A DFT é a versão digital da FT e, assim, não contém todas as freqüências constituintes da imagem mas um conjunto significativo destas capaz de descrevê-la.

Para o caso de imagens com dimensão $N \times N$, a DFT pode ser dada por:

$$F(u,v) = \frac{1}{N}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} f(m,n)e^{-j2\pi(\frac{mu}{N}+\frac{nv}{N})} \tag{42}$$

para $u,\ v = 0,1,\cdots,N-1$. $f(m,n)$ é a imagem no domínio espacial e os termos exponenciais constituem as funções de base associadas a cada ponto $F(u,v)$ no domínio da freqüência. Estas funções são componentes seno e cosseno de freqüências crescentes. Assim, $F(0,0)$ representa a menor freqüência (ou o nível DC do sinal) e $F(N-1,N-1)$, a maior destas freqüências.

Ainda para o caso de imagens de dimensão $N \times N$, a DFT inversa pode ser descrita por:

$$f(x,y) = \frac{1}{N}\sum_{u=0}^{N-1}\sum_{v=0}^{N-1} F(u,v)e^{j2\pi(\frac{xu}{N}+\frac{yv}{N})} \tag{43}$$

para $x = 0,1,\cdots,N-1$ e $y = 0,1,\cdots,N-1$.

Observe a consideração do termo $1/N$ tanto na transformada direta como inversa (comente isto em função da linearidade da DFT).

Geralmente, o intervalo de valores representando o espectro da FT é muito grande (intervalo entre a representação das altas freqüências, gerando valores de baixa intensidade, e a representação das baixas freqüências que predominam no sinal). Para amenizar tal dificuldade na visualização (display) do espectro, podemos considerar a seguinte função de escalonamento

$$D(u,v) = c\ log(1 + |F(u,v)|), \tag{44}$$

onde c é um fator de escala. Como podemos observar, a função acima realça o valor dos pixels associados às regiões de alta freqüência, melhorando a dinâmica e, conseqüentemente, a qualidade da informação visual do espectro. Por exemplo, para um espectro variando na faixa $[0,\ 2.0 \times 10^5]$, $log[1 + F(u,v)]$ define um novo valor s variando no intervalo $[0,\ 5.3]$. Um display deste intervalo num sistema de k bits deve considerar $c = \frac{s.(2^k-1)}{5.3}$. A Fig. 11 apresenta, respectivamente, uma imagem original, o display de $|F(u,v)|$ e uma melhor visualizção deste espectro (numa escala de 0 a 255).

De modo geral, e como na Fig. 11, o display do espectro de Fourier considera o ponto $F(0,0)$ centrado em $\mathrm{u}_0 = \mathrm{v}_0 = \mathrm{N}/2$. Isto pode ser realizado facilmente a partir das propriedades de translação ou deslocamento vistas anteriormente e que, para o caso discreto, podem ser indicadas por:

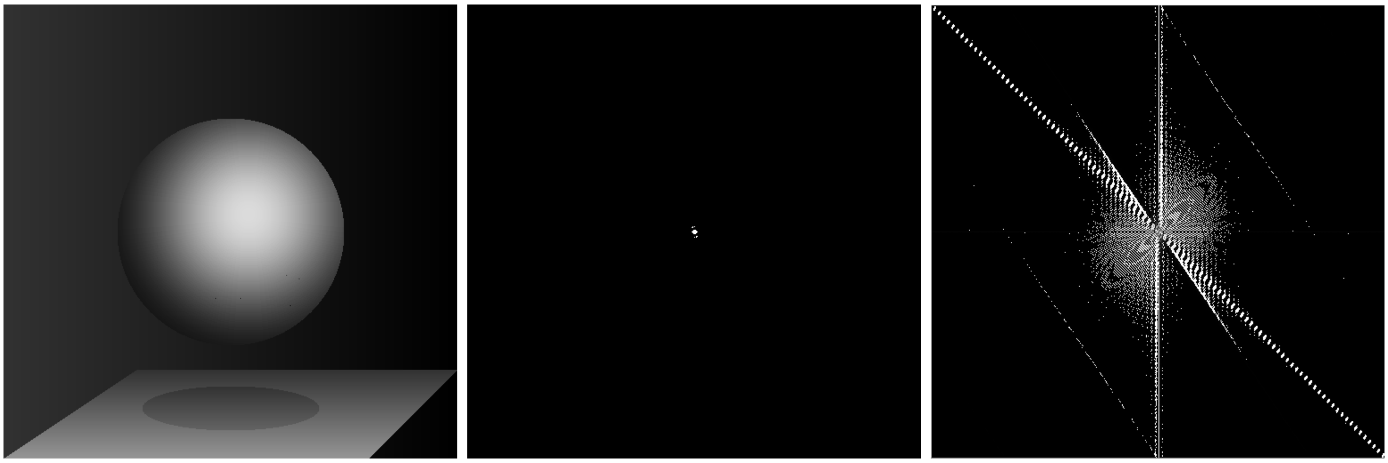


Figura 11:

$$f(x,y)e^{j2\pi(u_0x+v_0y)/N} \Leftrightarrow F(u-u_0, v-v_0) \tag{45}$$

e

$$f(x-x_0, y-y_0) \Leftrightarrow F(u,v)e^{-j2\pi(ux_0+vy_0)/N}, \tag{46}$$

ou seja, multiplicando-se f(x,y) pelo respectivo termo exponencial e calculando-se a DFT, obtemos um resultado com a origem deslocada para o ponto (u_0, v_0), no plano da freqüência. As considerações são semelhantes para a DFT inversa.

Observe que para o caso $u_0 = v_0 =$ N/2, temos

$$e^{j2\pi(u_0x+v_0y)/N} = e^{j\pi(x+y)} = (-1)^{x+y} \tag{47}$$

e

$$f(x,y)(-1)^{x+y} \Leftrightarrow F(u-N/2, v-N/2), \tag{48}$$

o que fornece facilmente a translação do espectro. Observe, ainda, que o deslocamento de f(x,y) não altera o valor do espectro já que:

$$|F(u,v)e^{-j2\pi(ux_0+vy_0)/N}| = |F(u,v)| \tag{49}$$

Em termos de uma primeira abordagem sobre a *implementação direta* da DFT, é importante considerar a propriedade de separabilidade da mesma. Neste caso, podemos expressar as equações 42 e 43 da seguinte forma:

$$F(u,v) = \frac{1}{N} \sum_{x=0}^{N-1} e^{-j2\pi ux/N} \sum_{y=0}^{N-1} f(x,y) e^{-j2\pi vy/N} \tag{50}$$

para u,v = 0,1,...,N-1, e

$$f(x,y) = \frac{1}{N} \sum_{u=0}^{N-1} e^{j2\pi ux/N} \sum_{v=0}^{N-1} F(u,v) e^{j2\pi vy/N} \tag{51}$$

para x,y = 0,1,...,N-1; ou ainda, para a Eq. 50:

$$F(u,v) = \frac{1}{N} \sum_{x=0}^{N-1} F(x,v) e^{-j2\pi ux/N}, \tag{52}$$

onde

$$F(x,v) = N[\frac{1}{N} \sum_{y=0}^{N-1} f(x,y) e^{-j2\pi vy/N}] \tag{53}$$

Este resultado mostra que $F(u,v)$ pode ser obtida considerando uma transformada 1-D, ao longo de cada linha de $f(x,y)$, definindo $F(x,v)$, seguida da mesma operação no sentido das colunas de $F(x,v)$. O mesmo desenvolvimento é válido para a Eq. 51.

Obs.: Consultar [Woo93], por exemplo, para outras propriedades interessantes da DFT.

> **Estudo dirigido:** Verifique que o número de operções necessárias à implementação de $[F]_{N\times N} = [P]_{N\times N}[f]_{N\times N}[Q]_{N\times N}$ é proporcional a $2N^3$. Leia a seção 3.4 do livro [Woo93] para uma boa compreensão da conhecida *"Fast Fourier Transform - FFT"* que a partir de uma decomposição adequada da DFT define um método de implementação em que o número de operações é proporcional a $Nlog_2N$. Este é o método mais comum de implementação do par da DFT.
>
> A seção 3.5 do mesmo livro aborda alguns aspectos relativos à propriedade de separabilidade e à definição geral de um transformada linear, vistas anteriormente. Leia esta seção para a análise de outros tipos de transformadas lineares. Em particular, a seção 3.5.3 aborda a *Transformada Discreta do Cosseno - DCT*. Após uma leitura sobre esta transformada, considere os comentários apresentados no exemplo da página 375 e aqueles abordados a partir do último parágrafo da página 378. Estes comentários estão relacionados com compressão de imagens e concernem diretamente a DCT.

7 Exemplo de aplicação: filtragem linear

Como mencionado anteriormente, a FT pode ser empregada em inúmeras aplicações de processamento de imagens, e um exemplo típico destas aplicações refere-se à *filtragem*. Com o objetivo de eliminar ou realçar determinadas componentes da imagem, tais como ruído ou contornos, a filtragem no domínio da freqüência explora a característica reversível da FT e a informação contida neste domínio.

Os filtros *passa-baixas*, por exemplo, associados às regiões homogêneas da imagem, "deixam passar" as componentes de baixa freqüência, atenuando as de alta freqüência relacionadas com as transições bruscas da imagem. Estas transições são representadas por ruídos ou contornos, o que significa que ao eliminarmos o ruído indesejável da imagem, a partir de um filtro passa-baixas, estamos atenuando os seus contornos nas mesmas proporções, caso típico do processamento linear. Os filtros *passa-altas*, por sua vez, atenuam ou eliminam as componentes de baixas freqüências. O efeito evidente desta filtragem é a obtenção de um realce nas zonas de alta freqüência do sinal, isto é, dos seus contornos e – infelizmente – ruído.

Num projeto de filtros lineares podemos considerar, ainda, os filtros *passa-faixas* associados às regiões compreendidas entre as baixas e altas freqüências. Estes filtros são empregados, por exemplo, em problemas de restauração de imagens. A Fig. 12 apresenta o modelo de filtros para o caso 1-D (uma rotação destas funções define os mesmos filtros para o caso 2-D).

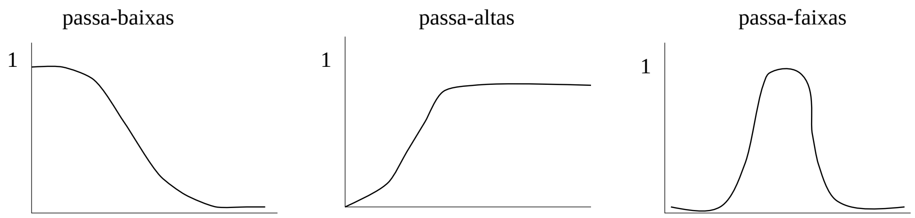


Figura 12:

O princípio geral da filtragem, no domínio da freqüência, pode ser indicado pela seguinte transformação

$$G(u,v) = H(u,v)F(u,v), \tag{54}$$

onde $F(u,v)$ é a transformada de Fourier da imagem de entrada, $G(u,v)$ é a imagem de saída filtrada, e $H(u,v)$ é a respectiva *função de transferência* do filtro. O problema consite, então, em definir a função $H(u,v)$ que conduza à imagem desejada

$G(u,v)$. A transformada inversa, $\mathcal{F}^{-1}[G(u,v)]$, define a imagem filtrada no domínio espacial $g(x,y)$.

7.1 Filtros passa-baixas

Um *filtro ideal passa-baixas* 2-D pode ser representado pela seguinte função de transferência

$$H(u,v) = \begin{cases} 1 & \text{se } D(u,v) \leq D_0 \\ 0 & \text{se } D(u,v) > D_0 \end{cases} \tag{55}$$

onde $D_0 \geq 0$ é a *freqüência de corte* e $D(u,v)$ é a distância do ponto (u,v) até a origem do plano da freqüência, ou seja,

$$D(u,v) = (u^2 + v^2)^{\frac{1}{2}} \tag{56}$$

Observe que de acordo com a Eq. 55, as freqüências contidas no círculo de raio D_0 não sofrem atenuações, daí o termo filtro ideal. A Fig. 13 apresenta uma imagem original com ruído, uma "máscara" correspondente à função $H(u,v)$, e a imagem resultante do filtro passa-baixas.

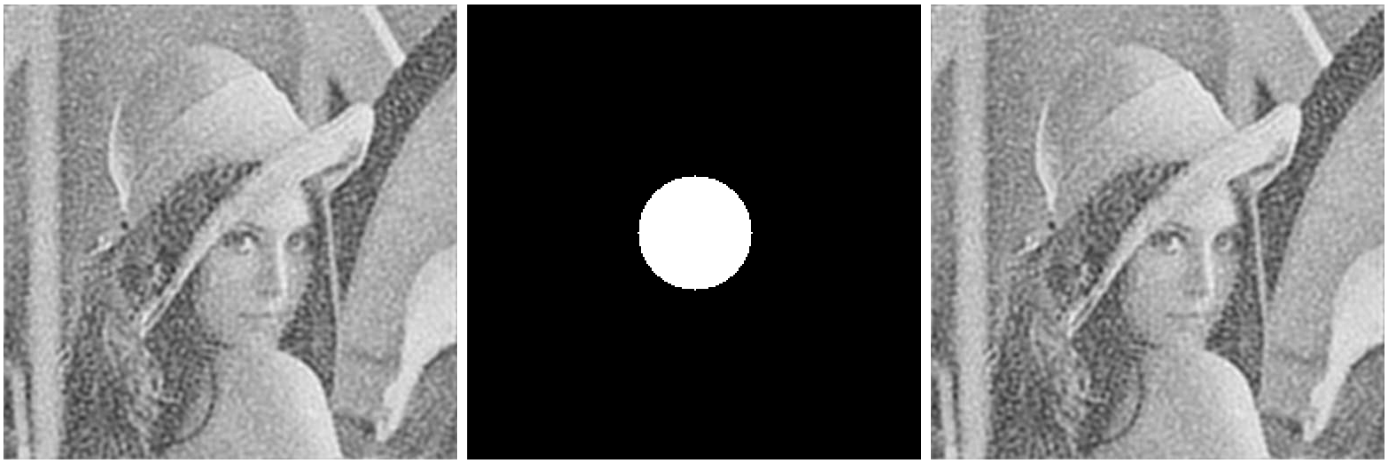


Figura 13:

Exercício: Verifique, em [Woo93], o exemplo apresentado na página 203 sobre filtro passa-baixas.

O *filtro passa-baixas de Butterworth* de ordem n é dado pela seguinte função de transferência:

$$H(u,v) = \frac{1}{1 + [D(u,v)/D_0]^{2n}} \tag{57}$$

Esta função define um filtro sem grande descontinuidade, não estabelecendo uma transição brusca na freqüência de corte. Nestes casos, costuma-se definir um valor para D_0 correspondendo a uma porcentagem do máximo valor $H(u,v)$. Podemos ver facilmente que quando $D_0 = D(u,v)$ o valor de $H(u,v)$ cai para 50% do seu valor inicial. Se quisermos, por exemplo, que este valor atinja $1/\sqrt{2}$, entao:

$$H(u,v) = \frac{1}{1 + [\sqrt{2} - 1][D(u,v)/D_0]^{2n}} = \frac{1}{1 + 0.414[D(u,v)/D_0]^{2n}} \tag{58}$$

7.2 Filtro passa-altas

Como mencionado anteriormente, as transições bruscas de um sinal estão associadas às componentes de alta freqüência do espectro de Fourier. Assim, um realce da imagem, com ênfase nestas transições, pode ser obtido deixando-se passar estas altas freqüências e atenuando-se as demais.

Um *filtro ideal passa-altas* 2-D é dado pela seguinte função de transferência:

$$H(u,v) = \begin{cases} 0 & \text{se } D(u,v) \leq D_0 \\ 1 & \text{se } D(u,v) > D_0 \end{cases} \tag{59}$$

D_0 é a freqüência de corte medida a partir da origem, no plano da freqüência, e $D(u,v)$ é definida como anteriormente. A Fig. 14 apresenta uma imagem original com ruído, uma "máscara" correspondente à função $H(u,v)$, e a imagem resultante do filtro passa-altas.

O *filtro passa-altas de Butterworth* de ordem n é dado por

$$H(u,v) = \frac{1}{1 + [D_0/D(u,v)]^{2n}} \tag{60}$$

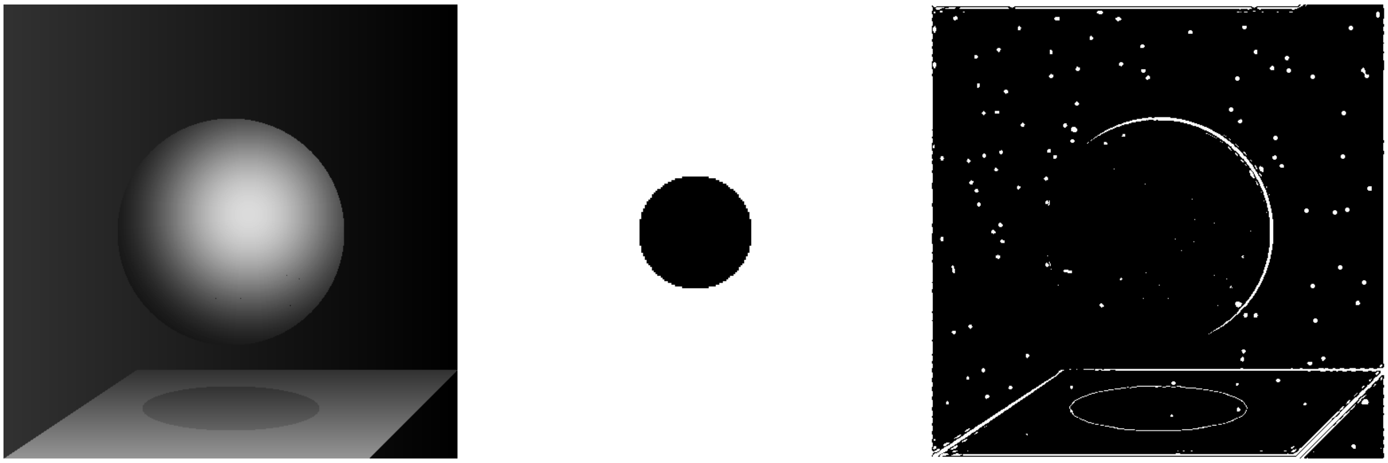


Figura 14:

Novamente, quando $D_0 = D(u,v)$, $H(u,v)$ cai 50% do seu valor máximo.

Exercício: Verifique os resultados da filtragem apresentada na página 213 em [Woo93].

8 Filtragem linear no domínio espacial

A filtragem linear, no domínio espacial, pode ser entendida facilmente a partir da propriedade da convolução relacionada com a FT.

Para o caso discreto, a convolução descrita pela Eq. 17, pode ser expressa por

$$g(x,y) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} h(i,j) f(x-i, y-j), \tag{61}$$

onde h corresponde à representação espacial da função de transferência H (Eq. 54), f é a imagem de entrada $N \times N$ com ruído, e g é a imagem resultante da filtragem no domínio espacial equivalente a $\mathcal{F}^{-1}[G(u,v)]$.

De maneira geral, a função $h(x,y)$ pode ser obtida diretamente a partir da DFT inversa de $H(u,v)$, de tamanho $N \times N$. Na prática, o que se considera é uma representação reduzida de $h(x,y)$, caracterizando o tipo de filtragem desejado. Esta filtragem é definida por uma combinação linear dos pixels numa vizinhança de $f(x,y)$ através das operações indicadas pela Eq. 61.

> **Estudo dirigido:** Leia a seção 4.5 em [Woo93] para uma abordagem sobre definição de uma máscara de convolução $h(x,y)$ a partir da DFT inversa de $H(u,v)$.

Sejam $\Gamma = (M+1)\times(N+1)$, M e N pares, uma vizinhança discreta centrada na origem, e uma imagem $f(x,y)$ com ruído. Uma *operação local* de filtragem linear, no domínio espacial, pode ser dada por

$$g(x,y) = \sum_{i=-m}^{m} \sum_{j=-n}^{n} h(i,j) f(x-i, y-j) \tag{62}$$

Esta equação representa uma convolução discreta da imagem f com uma *máscara* de coeficientes h. Assim, em termos de implementação, a operação de filtragem pode ser obtida a partir de um deslocamento de h sobre as linhas de f (ou vice-versa), no sentido da *varredura direta* da imagem (da esquerda para a direita, e de cima para baixo). Os coeficientes de h definem o tipo da filtragem: suavização ou realce de contornos, por exemplo.

8.1 Alguns exemplos: filtros passa-baixas

O *filtro da média* é um exemplo típico de uma filtragem passa-baixas no domínio espacial. Para uma vizinhança elementar 3×3, por exemplo, a máscara h pode ser dada por:

$$h = \frac{1}{9} \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} \tag{63}$$

É fácil constatar o efeito de redução do ruído após tal filtragem, assim como o problema de suavização dos contornos da imagem que, naturalmente, se torna mais grave à medida que a máscara h aumenta de dimensão. A Fig. 15 ilustra a filtragem linear de uma imagem com uma máscara 11×11.

Os valores de h podem ser definidos a partir da consideração do modelo de um determinado tipo de ruído, visando expressar aproximações deste modelo. As

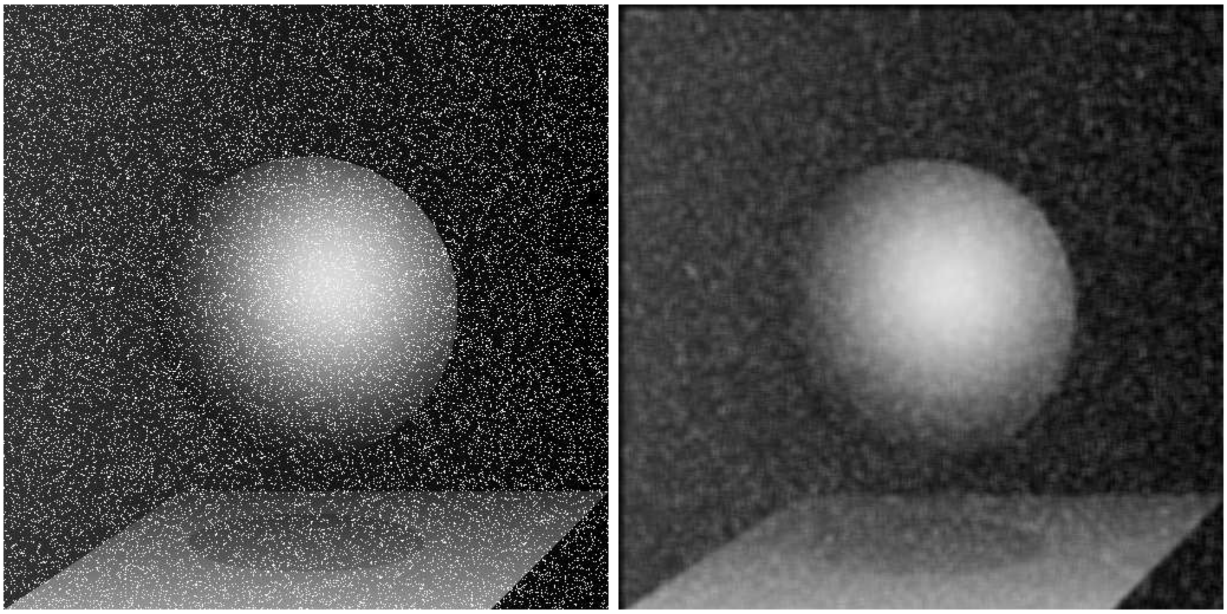


Figura 15:

máscaras, a seguir, são definidas considerando o ruído com uma distribuição de probabilidade gaussiana, distribuição esta muito utilizada na prática.

$$h_1 = \frac{1}{16} \begin{vmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{vmatrix} \tag{64}$$

8.2 Detectores de contorno

Os detectores de contorno identificam transições bruscas na função $f(x, y)$. A máscara 3×3, a seguir, é um exemplo de um detector de contorno espacial.

$$h = \begin{vmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{vmatrix} \tag{65}$$

Esta máscara, cujo somatório dos coeficientes é igual a zero, define valores nulos nas regiões homogêneas da imagem e valores mais elevados próximo aos seus contornos.

8.2.1 Operadores diferenciais

Os operadores diferenciais, tais como o *gradiente*, são largamente empregados na detecção de contornos [Kak82]. Para um sinal 2-D, identificamos mudanças da função em várias direções (derivadas parciais em torno dos eixos x e y, por exemplo). Neste caso, definimos um vetor gradiente do tipo:

$$\bigtriangledown f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \tag{66}$$

A magnitude deste vetor é

$$|\bigtriangledown f| = \left[(\frac{\partial f}{\partial x})^2 + (\frac{\partial f}{\partial y})^2 \right]^{1/2} \tag{67}$$

Esta equação define a intensidade do gradiente cuja direção, perpendicular à direção do contorno, é dada por

$$\theta = tg^{-1} \left[\frac{\bigtriangledown y}{\bigtriangledown x} \right] \tag{68}$$

8.2.2 Exemplos

Para o caso discreto, o operador acima é definido a partir de diferenças entre pixels, numa determinada vizinhança. O conhecido *gradiente de Roberts*, por exemplo, é representado por (Fig. 16)

$$\bigtriangledown_{d1} f(x,y) = f(x,y) - f(x+1,y+1) \tag{69}$$

$$\bigtriangledown_{d2} f(x,y) = f(x,y+1) - f(x+1,y) \tag{70}$$

Este operador responde a mudanças nas direções diagonais (45^o e 135^o) da imagem, considerando pares de pontos centrados em (x + 1/2, y + 1/2). O gradiente de Roberts pode ser representado, ainda, pelas seguintes máscaras de convolução:

$$\bigtriangledown_{d_1} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \bigtriangledown_{d_2} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tag{71}$$

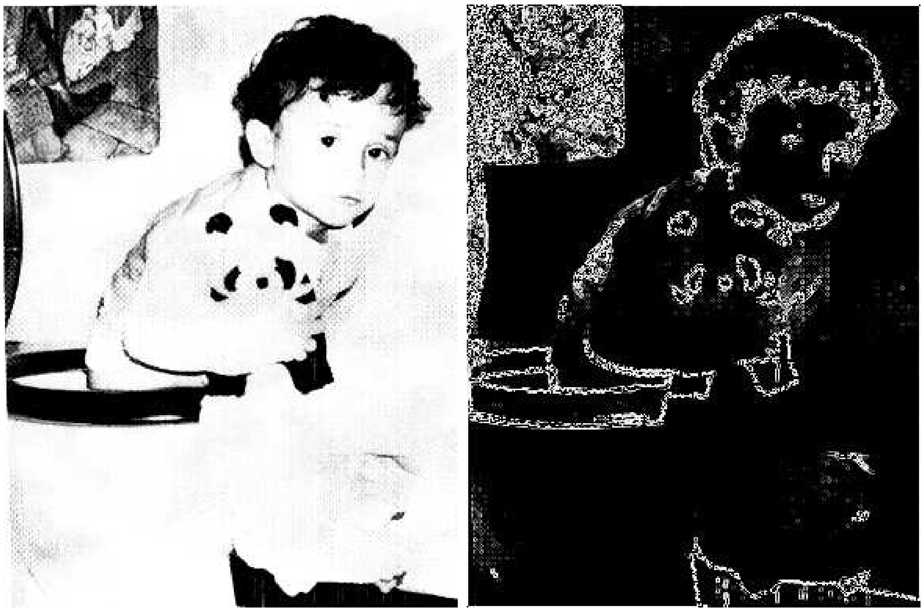


Figura 16:

Outras aproximações para a magnitude do gradiente, indicando descontinuidades nas direções principais x e y, podem ser obtidas a partir dos operadores de *Prewitt* e *Sobel* definidos, respectivamente, pelas seguintes máscaras de convolução:

$$\bigtriangledown_x = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad \bigtriangledown_y = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \tag{72}$$

e

$$\bigtriangledown_x = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad \bigtriangledown_y = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \tag{73}$$

Para as diferenças nas direções ortogonais $\bigtriangledown_x$ e $\bigtriangledown_y$, acima, a intensidade do contorno no ponto (x, y) é (Eq. 67):

$$| \bigtriangledown f(x,y)| = \sqrt{{\bigtriangledown_x}^2 + {\bigtriangledown_y}^2} \tag{74}$$

A magnitude do gradiente pode ser aproximada por

$$|\bigtriangledown f(x,y)| = max(|\bigtriangledown_x|, |\bigtriangledown_y|) \tag{75}$$

ou

$$|\bigtriangledown f(x,y)| = |\bigtriangledown_x| + |\bigtriangledown_y| \tag{76}$$

Estas aproximações evitam o cálculo da raiz quadrada e, para o caso da Eq. 75, a introdução de novos valores de níveis de cinza fora do intervalo original. A Fig. 17 apresenta uma imagem seguida da detecção de seus contornos nas direções x e y, utilizando o operador de Prewitt e o cálculo da magnitude indicado pela Eq. 75.

O Laplaciano

O *Laplaciano* 2-D de uma função f é um operador diferencial linear dado por

$$\bigtriangledown^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \tag{77}$$

Para o caso discreto, podemos mostrar que

$$\bigtriangledown^2 f(x,y) = {\bigtriangledown_x}^2 f(x,y) + {\bigtriangledown_y}^2 f(x,y) = [f(x+1,y)+f(x-1,y)+f(x,y-1)+f(x,y+1)] - 4f(x,y), \tag{78}$$

o que corresponde a uma convolução de f com a máscara

$$h = \begin{vmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{vmatrix} \tag{79}$$

Observe que, como esperado, o Laplaciano é zero para uma rampa linear e que o mesmo responde na entrada e na saída dos pontos de transição com valores positivos e negativos que se alternam. O módulo, $|\bigtriangledown^2 f|$, ou apenas a parte positiva, $|\bigtriangledown^2 f| \geq 0$, podem ser considerados como resultado do Laplaciano. No caso em que as respostas positivas e negativas são consideradas, um dado interessante associado a este operador é a *passagem por zero* (zero-crossing) da função.

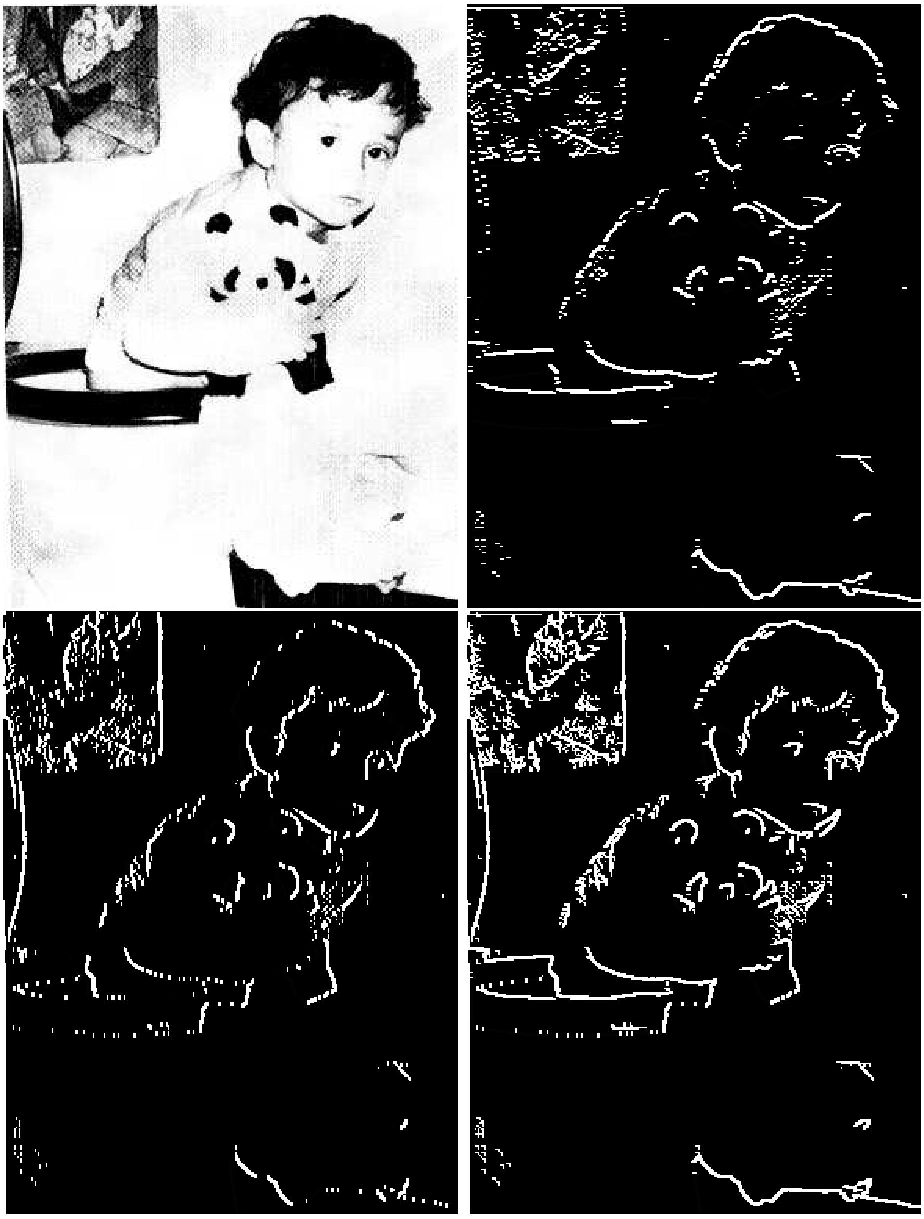


Figura 17:

Exercício 1: Mostre que o Laplaciano é proporcional à diferença entre uma imagem f e sua versão filtrada, $\overline{f}$, considerando os valores do ponto (x, y) e dos seus quatro vizinhos mais próximos.

Exercício 2: Verifique ???????

9 Filtragem não-linear

Como vimos, um dos grandes problemas relacionados com a eliminação do ruído de uma imagem, através de uma filtragem linear, refere-se à suavização dos seus contornos. O processamento não-linear aborda este problema tentando evitar uma filtragem homogênea ao longo das regiões próximas a estes contornos. Uma classe de filtros não-lineares bastante empregada na eliminação de ruídos, com preservação de contornos, são os *filtros estatísticos da ordem* [Kak82] [ea83]. Dentre estes, um dos mais importantes é o *filtro da mediana* que consiste em substituir o valor de um pixel m, centrado numa determinada vizinhança, pelo valor mediano dos pixels desta vizinhança ordenados de acordo com suas magnitudes [Hey82]. O filtro da mediana elimina eficientemente *ruído impulsivo*, do tipo *sal e pimenta*, representando descontinuidades abruptas e isoladas na imagem. Além disto, ele não introduz valores de níveis de cinza diferentes daqueles contidos na imagem original e, por afetar menos os contornos, pode ser aplicado iterativamente. A Fig. 18 apresenta uma imagem com ruído e uma iteração do filtro da mediana com uma vizinhança 5×5.

Outros casos específicos dos filtros estatísticos da ordem são os valores *máximo* e *mínimo* da seqüência ordenada de pixels. Como veremos, isto representa uma generalização das operações de *dilatação* e *erosão* morfológicas discutidas a seguir.

10 Morfologia Matemática

Um dos grandes problemas em processamento e análise de imagens refere-se à extração das componentes de interesse contidas na cena original. A análise do conteúdo desta cena depende de duas etapas principais. Uma relativa à *segmentação*, consistindo de transformações geralmente dependentes da aplicação, e outra referente à *análise quantitativa* dos objetos segmentados. Esta etapa visa a extração de características dos objetos a serem consideradas num processo de *classificação*.

Morfologia Matemática (MM) representa uma técnica não-linear de processamento de imagens e define um conjunto de ferramentas que auxiliam na realização

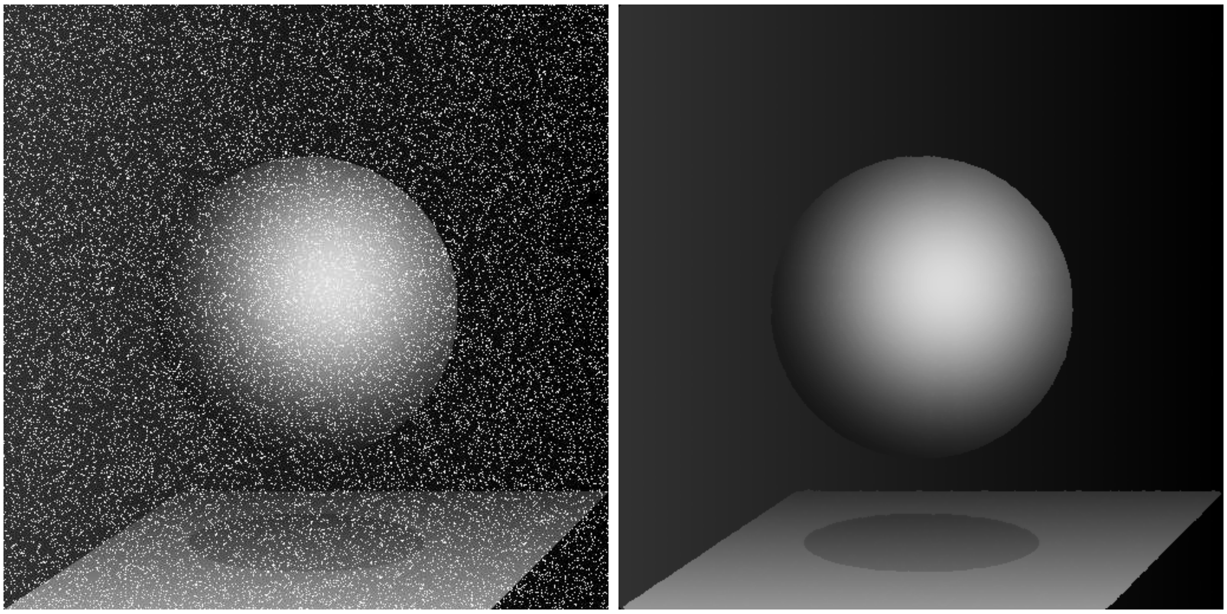


Figura 18:

destas etapas. Formalizada a partir dos anos 60 com os trabalhos de G. Matheron e J. Serra [Mat75], [Ser82], [Ser88], [Bar94], do Centro de Morfologia Matemática da Escola de Minas de Paris, ela tem como fundamento a teoria dos conjuntos e a geometria integral, e é utilizada atualmente no formalismo de diversos problemas práticos e teóricos de processamento e análise de imagens.

Como veremos, a idéia básica da MM consiste da comparação do conteúdo de uma imagem com outra menor e de forma conhecida, denominada *elemento estruturante*. De modo geral, este elemento contém características geométricas e/ou topológicas relacionadas com a informação que pretendemos extrair da imagem de interesse.

11 As transformações morfológicas

As transformações morfológicas podem ser consideradas sobre imagens binárias ou em níveis de cinza, vistas como conjuntos e funções, respectivamente. A estas imagens podemos associar os operadores booleanos de *união* $\cup$ e *interseção* $\cap$, no caso de conjuntos, e os operadores *max* e *min*, no caso de funções.

As transformações morfológicas dispõem de algumas propriedades algébricas e topológicas que são importantes caracterizar. Seja ψ uma transformação agindo sobre imagens binárias ou em níveis de cinza. Mais distintamente, para imagens binárias, consideramos subconjuntos $\mathcal{P}(\Re^2)$ de $\Re^2$ e $\psi : \mathcal{P}(\Re^2) \to \mathcal{P}(\Re^2)$. Neste caso, a relação de ordem é a inclusão. Para imagens em níveis de cinza, ψ opera sobre funções de $\Re^2$ em $\Re$. Se $\mathcal{F}$ designa o conjunto destas funções, ou seja, $\mathcal{F} = \{f : \Re^2 \to \Re\}$, então $\psi : \mathcal{F} \to \mathcal{F}$. Neste caso, a relação de ordem é:

$$\forall f, g \in \mathcal{F}, f \leq g \quad \Leftrightarrow \quad \forall x \in \Re^2, f(x) \leq g(x) \tag{80}$$

É importante lembrar que, na prática, as imagens são processadas no plano discreto $\mathcal{Z}^2$, com funções assumindo valores em $\mathcal{Z}$. Uma seqüência de imagens pode ser definida como uma função em $\mathcal{Z}^3$ em que a terceira variável representa o tempo, por exemplo.

A transformação morfológica ψ é dita *extensiva* se, e somente se

$$\forall X \in \mathcal{P}(\Re^2), X \subseteq \psi(X) \quad ou \quad \forall f \in \mathcal{F}, f \leq \psi(f) \tag{81}$$

e *anti-extensiva* se, e somente se

$$\forall X \in \mathcal{P}(\Re^2), \psi(X) \subseteq X \quad ou \quad \forall f \in \mathcal{F}, \psi(f) \leq f \tag{82}$$

A mesma transformação é dita *crescente* se ela preserva a ordem no espaço $\mathcal{P}(\Re^2)$, ou seja,

$$\forall X, Y \in \mathcal{P}(\Re^2), X \subseteq Y \Rightarrow \psi(X) \subseteq \psi(Y) \quad ou \tag{83}$$
$$\forall f, g \in \mathcal{F}, f \leq g \Rightarrow \psi(f) \leq \psi(g) \tag{84}$$

Em caso contrário ($X \subseteq Y \Rightarrow \psi(Y) \subseteq \psi(X)$ ou $f \leq g \Rightarrow \psi(f) \geq \psi(g)$), ela é dita *decrescente*. A operação ψ é *idempotente* se

$$\psi \circ \psi = \psi, \tag{85}$$

ou seja, o resultado de várias iterações do operador ψ sobre a imagem equivale ao resultado de uma única aplicação.

Duas transformações ψ_1 e ψ_2 são *duais* se, e somente se

$$\forall X \in \mathcal{P}(\Re^2), \psi_1(X) = (\psi_2(X^c))^c \quad ou \quad \forall \in \mathcal{F}, \psi_1(f) = -(\psi_2(-f)) \tag{86}$$

Finalmente, uma transformação é dita *homotópica* se ela não modifica o número de conexidade N_n de um conjunto X, isto é:

$$N_n(\psi(X)) = N_n(X) \tag{87}$$

12 Erosão e dilatação

Erosão e dilatação são as operações elementares da MM e formam a base para a construção das transformações mais complexas. Assim, numa cadeia morfológica de processamento de imagens, podemos encontrar um grande número de operadores encadeados, todos definidos a partir destas funções elementares.

Seja X o conjunto inicial dos pontos que constituem a imagem binária a ser analisada, e B uma pequena imagem representando um certo conhecimento geométrico e/ou topológico de X. Ao elemento estruturante B associamos um ponto *origem* representando, na imagem, a posição de referência para o resultado das transformações. A Fig. 19 apresenta dois exemplos de elementos estruturantes de dimensão 3×3 com origem no seu centro. A Fig. 19(a) define um elemento *4-conexo*, e a Fig. 19(b), um elemento *8-conexo*.

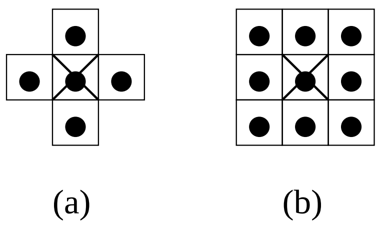


Figura 19: Exemplos de elementos estruturantes numa vizinhança 3×3.

A dilatação de X por B, $\delta_B(X)$, é dada pela união de todos os pontos x de $\Re^2$ tal que o elemento estruturante B, centrado em x, B$_x$, intercepta X. Assim:

$$\delta_B(X) = \{x \in \Re^2, B_x \cap X \neq \emptyset\} \tag{88}$$

A Fig. 20 apresenta o resultado de duas iterações de dilatação de uma imagem binária com o elemento estruturante 3×3 indicado na Fig. 19(b). Aqui, as imagens contêm apenas dois valores (imagens binárias): o valor "1" representando o conjunto X de pontos claros da *forma* (foreground), e o valor "0" correspondente ao complemento de X, X^c, e representando os pontos escuros do *fundo* (background) da imagem.

A erosão de X por B, ϵ_B, é definida pelo conjunto dos pontos de x em $\Re^2$ tal que B, ao ser transladado para a posição x, esteja totalmente incluído no conjunto X. Assim:

$$\epsilon_B(X) = \{x \in \Re^2, B_x \subset X\} \tag{89}$$

(a) (b)

Figura 20: Imagem original (a), e imagem dilatada (b).

A Fig. 21 apresenta o resultado de duas iterações de erosão de uma imagem binária com o elemento estruturante da Fig. 19(b).

(a) (b)

Figura 21: Imagem original (a), e imagem erodida (b).

No caso de funções, e considerando o exemplo prático de elementos estruturantes com valores iguais a zero no seu domínio $k \in \Re^2$ (*elemento estruturante planar*), a dilatação é dada simplesmente por:

$$\delta_B(f)(x) = max\{x_k, k \in B\}, \tag{90}$$

e a erosão:

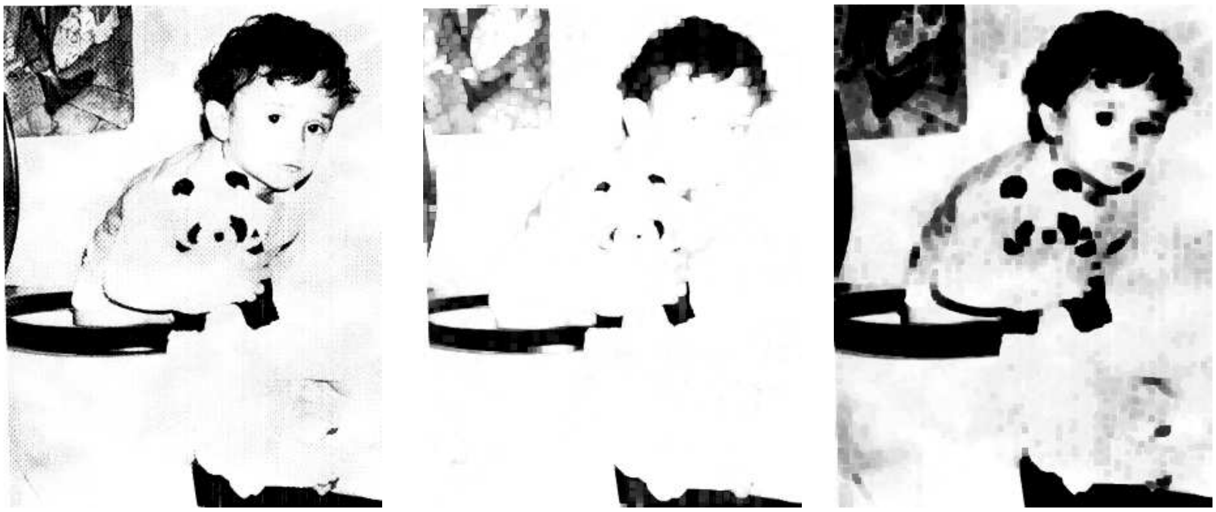
$$\epsilon_B(f)(x) = min\{x_k, k \in B\} \tag{91}$$

A Fig. 22 apresenta um exemplo de dilatação e erosão de uma imagem em níveis de cinza com um elemento estruturante de dimensão 5×5. Como podemos observar, a dilatação amplia as zonas claras da imagem e a erosão realiza o processo inverso, aumentando o domínio das regiões escuras.

É fácil constatar que erosão e dilatação são operações crescentes e duais, e que ambas não são nem idempotentes nem homotópicas. Observe, ainda, que para um elemento B de tamanho n, B_n, definido a partir de n iterações de dilatação com um elemento estruturante de tamanho unitário, B_1 (por exemplo, os elementos da Fig 19), então a erosão (dilatação) de um conjunto X com B_n pode ser decomposta em uma seqüência de n erosões (dilatações) de X com B_1. Assim:

$$\delta_{B_1} \circ \delta_{B_2} = \delta_{\delta_{B_1}(B_2)} \quad e \tag{92}$$

$$\epsilon_{B_1} \circ \epsilon_{B_2} = \epsilon_{\delta_{B_1}(B_2)} \tag{93}$$



(a) (b) (c)

Figura 22: Imagem original (a), imagem dilatada (b), e imagem erodida (c).

13 Abertura e fechamento morfológicos

As operações de *abertura* e *fechamento* com um elemento estruturante B, denotadas respectivamente por γ_B e φ_B, são obtidas combinando-se erosão e dilatação da seguinte forma:

$$\gamma_B = \delta_{\check{B}} \circ \epsilon_B, \tag{94}$$

e

$$\varphi_B = \epsilon_{\check{B}} \circ \delta_B, \tag{95}$$

onde $\check{B}$ é o transposto de B, ou seja, o simétrico de B em relação a sua origem. Naturalmente, no caso de B simétrico, temos

$$\gamma_B = \delta_B \circ \epsilon_B$$

e

$$\varphi_B = \epsilon_B \circ \delta_B$$

Abertura e fechamento morfológicos são operações duais, crescentes e idempotentes. A abertura é anti-extensiva e o fechamento, extensivo.

Em termos visuais, a abertura regulariza os contornos e elimina pequenas "ilhas" e "cabos" estreitos de uma imagem binária. O fechamento suprime pequenos "lagos" e "canais" estreitos (Fig. 23). Da mesma forma, para imagens em níveis de cinza, a abertura elimina estruturas claras, enquanto o fechamento atua sobre as estruturas escuras. Tudo isto em função da forma e dimensão do elemento estruturante considerado. Observe, finalmente, que estas operações não são homotópicas.

(a) (b) (c)

Figura 23: Imagem binária (a), abertura (b), e fechamento de tamanho 5 (c).

14 Filtros morfológicos

Os filtros morfológicos constituem alternativas interessantes à filtragem não-linear e são caracterizados por uma transformação ψ *crescente* e *idempotente* [Ser94]. Assim:

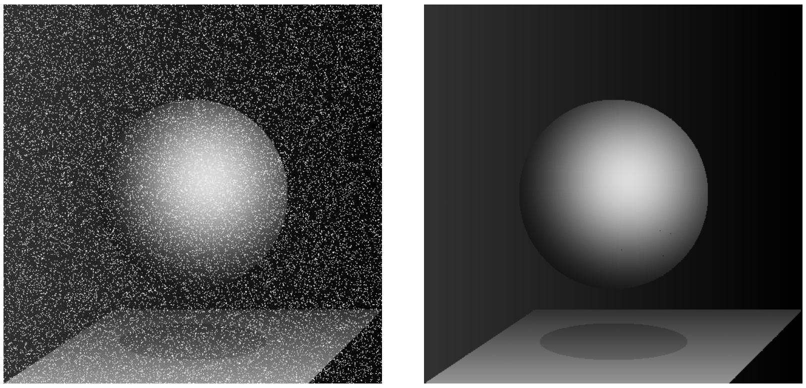
$$\forall (f,g) \in \mathcal{F} \times \mathcal{F}, f \leq g \quad \Rightarrow \quad \psi(f) \leq \psi(g) \tag{96}$$

$$\forall f \in \mathcal{F}, \psi(f) = \psi(\psi(f)) \tag{97}$$

Neste sentido, a erosão e a dilatação, por não serem idempotentes, não constituem filtros morfológicos. Por sua vez, a abertura e o fechamento são exemplos típicos de operadores de filtragem, cuja composição também define filtros. Esta composição pode ser considerada a partir de aberturas e fechamentos com elementos estruturantes de tamanho variável o que conduz, entre outros, à definição de filtros com comportamento simétrico em relação às estruturas claras e escuras da imagem [Ser88], [Che89]. A Fig. 24 ilustra uma filtragem por abertura.

14.1 Outros exemplos de operações elementares

O conjunto das transformações morfológicas pode ser obtido a partir da combinação dos operadores apresentados anteriormente. Como primeiro exemplo destas transformações, podemos citar o resíduo da dilatação e erosão que define o *gradiente*



(a) (b)

Figura 24: Imagem original com ruído (a), e exemplo de filtragem por abertura (b).

morfológico simétrico (Fig. 25):

$$grad(f) = \delta_B(f) - \epsilon_B(f) \tag{98}$$

Da mesma forma, os gradientes interno e externo são dados, respectivamente, por

$$grad^-(f) = f - \epsilon_B(f) \tag{99}$$
$$grad^+(f) = \delta_B(f) - f \tag{100}$$

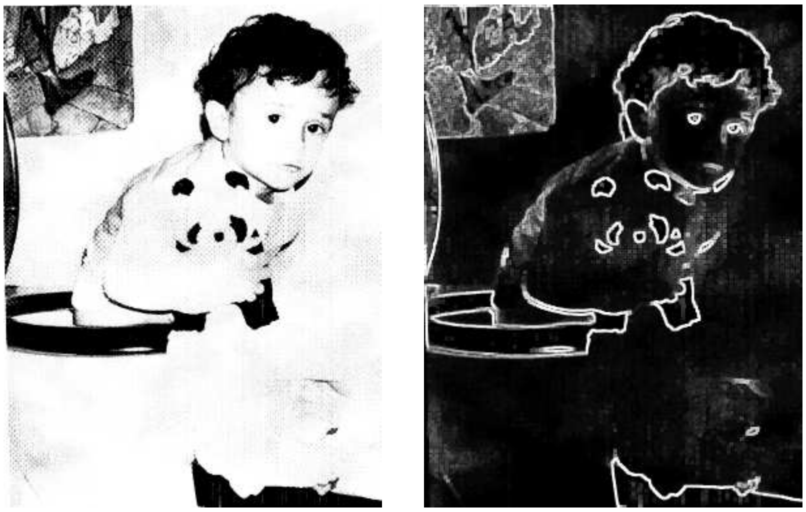
Outra transformação importante é o *chapéu mexicano claro* (white top-hat) que detecta as estruturas claras da imagem eliminadas por uma abertura de tamanho n (n iterações com um elemento estruturante elementar).

$$CM_n^+(f) = f - \gamma_n(f) \tag{101}$$

O *chapéu mexicano escuro* (black top-hat), associado às estruturas escuras da imagem, é dado por

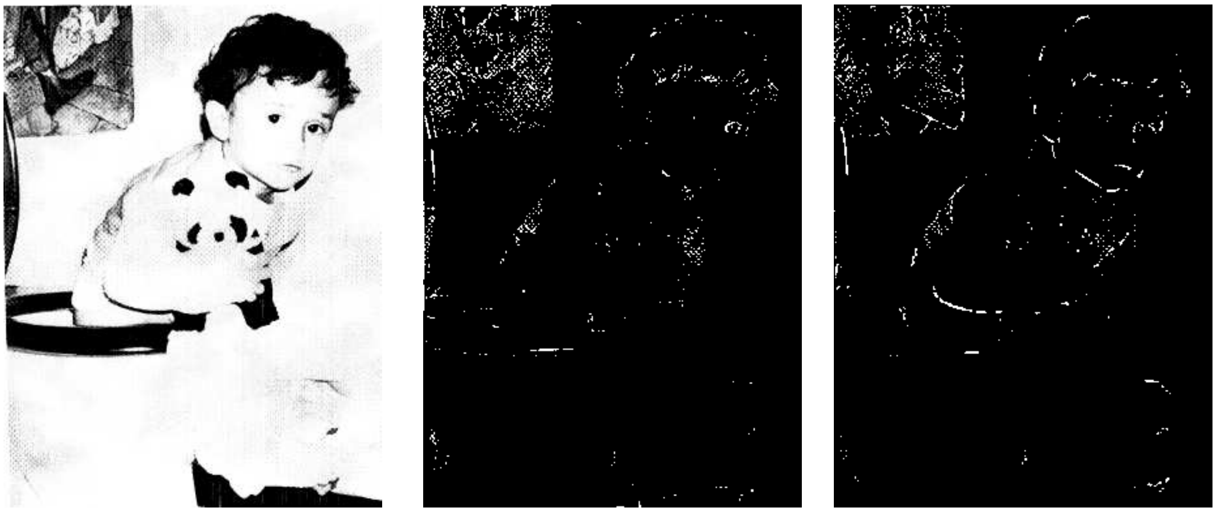
$$CM_n^-(f) = \varphi_n(f) - f \tag{102}$$

A Fig. 26 ilustra estas transformações que são empregadas, entre outros, na definição do esqueleto morfológico de uma imagem [Ser88].



(a) (b)

Figura 25: Imagem original (a), e o seu gradiente morfológico (b).



(a) (b) (c)

Figura 26: Imagem original (a), o chapéu mexicano claro (b), e o escuro (c).

$*$	0	$*$
1	1	1
$*$	1	$*$

$*$	0	0
1	1	$*$
$*$	1	$*$

Figura 27: Exemplo de elementos estruturantes homotópicos de afinamento.

15 A transformação Tudo ou Nada

A transformação Tudo ou Nada (Hit or Miss) considera as duas fases (pontos "1" e pontos "0") do elemento estruturante B. Sejam B_x^1 e B_x^2 estas fases com a mesma origem centrada em $x \in X$. A transformação Tudo ou Nada de X com B, X $\otimes$ B, é dada por

$$X \otimes B = \{x \in \Re^2 : B_x^1 \subset X, B_x^2 \subset X^c\} \tag{103}$$

ou ainda

$$X \otimes B = \epsilon_{B_x^1}(X) \backslash \delta_{B_x^2}(X) = \epsilon_{B_x^1}(X) \cap \epsilon_{B_x^2}(X^c) \tag{104}$$

15.1 Afinamento e espessamento homotópicos

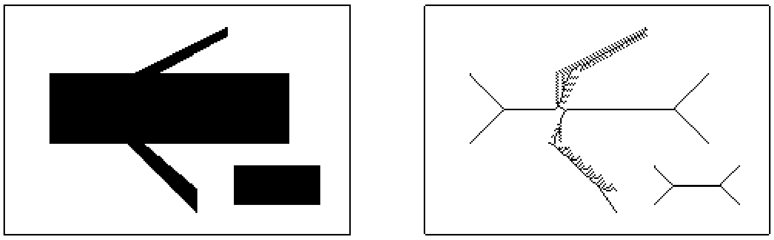
O afinamento e o espessamento homotópicos são exemplos de aplicações obtidas a partir da transformação do tipo Tudo ou Nada. O afinamento de uma imagem X com um elemento estruturante B, X $\ominus$ B, define uma representação homotópica do esqueleto de X (Fig. 28), e é dado por

$$X \circ B = X \backslash (X \otimes B) \tag{105}$$

O afinamento elimina pontos da borda de X e deve ser aplicado iterativamente até a idempotência, considerando-se uma seqüência de elementos estruturantes que preservem a topologia da imagem. A Fig. 27 apresenta exemplos destes elementos que, juntamente com suas rotações de 90^o, devem ser utilizados alternadamente no afinamento em uma malha retangular (o $*$ indica que o ponto pode assumir o valor "0" ou "1" na respectiva posição).

O espessamento é a operação dual do afinamento e é dado por

$$X \odot B = X \cup (X \otimes B) \tag{106}$$



(a) (b)

Figura 28: Imagem original (a), e o resultado do afinamento com as máscaras da Fig. 27 (b).

O espessamento homotópico adiciona pontos às componentes da imagem sem que a sua topologia seja alterada. É fácil verificar que os elementos estruturantes empregados aqui são o complemento daqueles utilizados na operação de afinamento e que, por dualidade, esta transformação define o esqueleto do fundo da imagem.

A abordagem anterior pode ser estendida a funções. Assim, seja f uma imagem em níveis de cinza e B_x^1, B_x^2 os subconjuntos do elemento estruturante planar representando, respectivamente, as fases "1" e "0" de B. O afinamento de f por B, $f \circ B$, é dado por

$$(f \circ B) = \begin{cases} \delta_{B_x^2}(f), \text{se } \delta_{B_x^2}(f) < f \leq \epsilon_{B_x^1}(f), \\ f, \text{ caso contrário} \end{cases} \tag{107}$$

O afinamento numérico define a zona de influência dos *mínimos regionais* de uma imagem, salientando os picos das "montanhas" que constituem o seu relevo. Esta operação pode ser associada à transformação que calcula a linha divisora de águas de uma função e que, como veremos posteriormente, constitui a base dos algoritmos morfológicos de segmentação.

16 Operações geodésicas e reconstrução

Sejam x e y dois pontos quaisquer de um conjunto X. O arco geodésico de extremidade xy é definido como o arco de menor distância $d_X(x, y)$ entre x e y, contido em X [Mai84]. Caso este arco não exista, esta distância é considerada infinita (Fig. 29). A função d_X é denominada distância geodésica.

Como discutido em [Mai84], as transformações morfológicas básicas podem ser

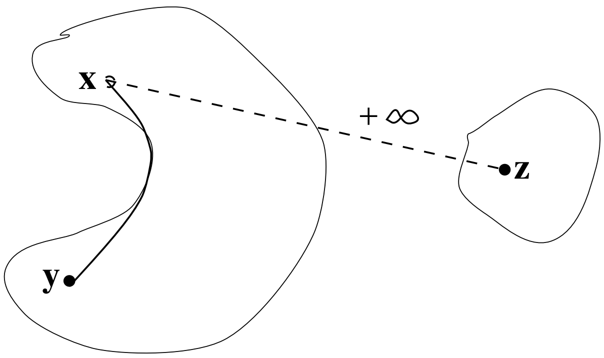


Figura 29: Noção de distância geodésica.

definidas no espaço métrico (X, d_X). Por exemplo, se Y $\subset$ X, pontos x de X tal que B$_X(x,\, \lambda)$ constituem o conjunto $\lambda - dilatado$ de Y em X. Este conjunto é denotado por

$$\delta_X^{\lambda}(Y) = \{x \in X : B_X(x, \lambda) \cap Y \neq \emptyset\} \tag{108}$$

onde B$_X(x,\, \lambda)$ = $\{x \in$ X: d$_X(x,\, y) \leq \lambda\}$

Na prática, a dilatação geodésica de tamanho n, no espaço $\mathcal{Z}^2$, pode ser obtida a partir de n iterações da dilatação geodésica de tamanho 1 definida por

$$\delta_X^1(Y) = (Y \oplus B) \cap X, \tag{109}$$

e assim:

$$\delta_X^n(Y) = \underbrace{\delta_X^1(\delta_X^1(...\delta_X^1(Y)))}_{n\ vezes} \tag{110}$$

O elemento estruturante B pode ser aqueles indicados na Fig. 19.

A erosão geodésica de Y em X é dada por

$$\epsilon_X^{\lambda}(Y) = \{x \in X : B_x(x, \lambda) \subseteq Y\} \tag{111}$$

As transformações geodésicas são empregadas na reconstrução de um conjunto conexo X marcado por outro conjunto de referência Y, não vazio, denominado *marcador*. O conjunto X é a máscara representando a componente binária que pretendemos preservar. Uma dilatação geodésica de Y em X, até a idempotência, permite a reconstrução total desta componente. Como exemplo, consideremos o problema de eliminação de partículas parcialmente incluídas numa imagem [Che89] (Fig. 30).

Se $\partial\Sigma$ representa o subconjunto de pontos pertencentes à fronteira de uma imagem Σ, então o novo subconjunto X', contendo componentes totalmente incluídas em X, é dado por

$$X' = X \backslash \delta_X^{\infty}(\partial\Sigma) \tag{112}$$

$\delta^{\infty}(\cdot)$ indica a execução da dilatação até a idempotência.

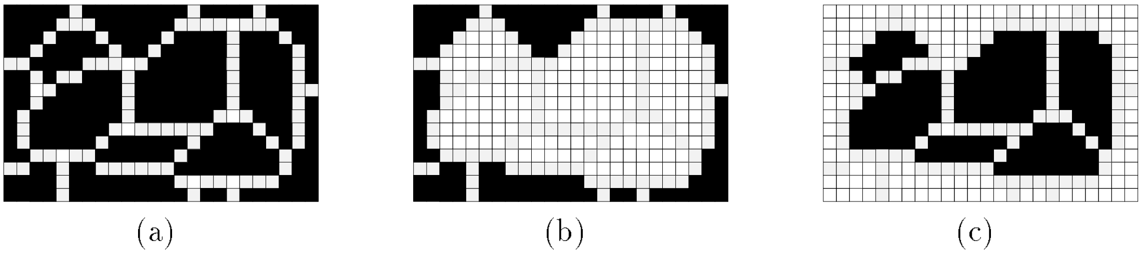


Figura 30: Imagem Σ contendo partículas parcialmente incluídas (a); dilatação geodésica de $\partial\Sigma$ em X (b); subconjunto das partículas totalmente incluídas em Σ (c).

A noção de reconstrução geodésica é muito importante para o processamento morfológico, e constitui a base de inúmeras transformações mais complexas relacionadas, por exemplo, com separação de partículas, definição de operadores conexos e funções de extinção morfológicas [Vin93], [JC93], [Vac95].

No caso numérico, a dilatação geodésica de uma função g "marcando" uma função f ($g \leq f$) é dada por (Fig. 31)

$$\delta_f^1(g) = min(\delta_1(g), f) \tag{113}$$

$$\delta_f^n(g) = \underbrace{\delta_f^1(\delta_f^1(...\delta_f^1(g)))}_{n\ vezes} \tag{114}$$

A realização desta função até a idempotência define a *reconstrução geodésica* de f por g. A erosão de uma função $g \geq f$, $\epsilon_f^n(g)$, é definida de maneira dual. Observe na Fig. 31(a) que apenas os domos da função f, marcados pela função g, são parcialmente reconstruídos. O mesmo acontece com os vales no caso da operação dual (Fig. 31(b)). Observe, ainda, que quando os valores da função marcadora g coincidem com os máximos regionais de f, a dilatação geodésica conduz a uma reconstrução total da máscara f.

A seção seguinte apresenta uma aplicação interessante das operações geodésicas relacionada com os algoritmos morfológicos de segmentação.

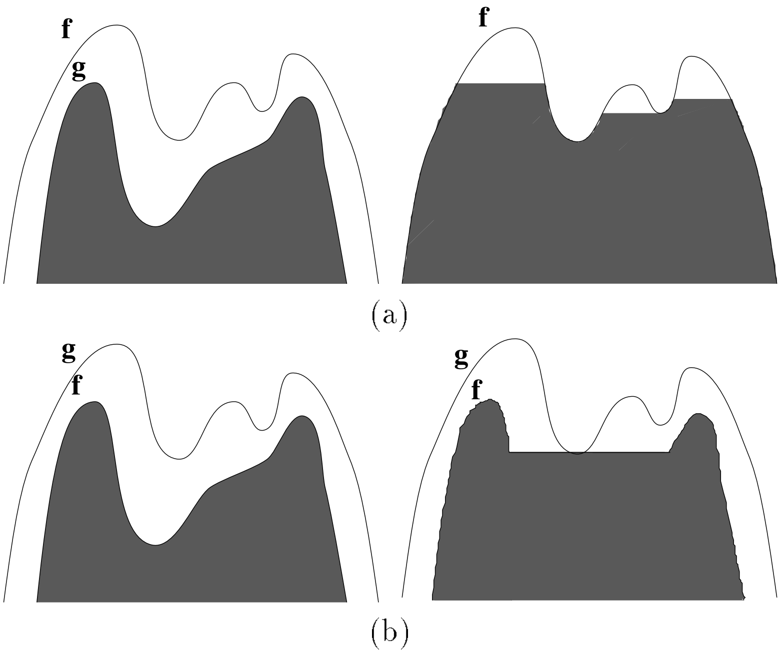


Figura 31: Exemplo de reconstrução geodésica (a), e reconstrução dual (b).

17 Esqueleto por zona de influência

Seja Y uma imagem binária composta de várias componentes conexas Y_1, Y_2, $\cdots$,Y_k. Seja X uma outra imagem binária tal que $Y \subset X$. A zona de influência geodésica, $iz_X(Y_i)$, de um conjunto conexo Y_i é o conjunto de pontos p de X cuja distância geodésica, em relação a Y_i, é menor que a distância geodésica de qualquer outra componente de Y:

$$iz_X(Y_i) = \{p \in X : \forall j \in [1,k], j \neq i, d_X(p, Y_i) < d_X(p, Y_j)\} \tag{115}$$

Os pontos de X que não pertencem a nenhuma zona de influência constituem o *esqueleto por zona de influência - SKIZ* de Y em X (Fig. 32), ou seja,

$$SKIZ_X(Y) = X \backslash IZ_X(Y) \tag{116}$$
$$IZ_X(Y) = \bigcup_{i \in [1,k]} iz_X(Y_i) \tag{117}$$

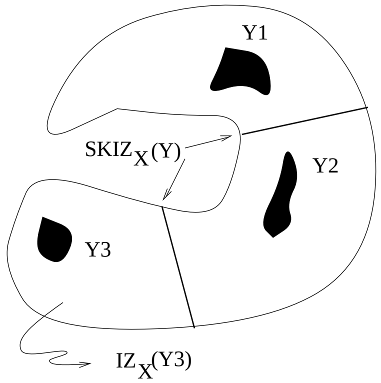


Figura 32: Esqueleto por zona de influência de Y em X.

18 Segmentação morfológica

A principal ferramenta de segmentação morfológica é baseada na transformação que define a linha divisora de águas - LDA ou *watershed* de uma função [Mey93], [Soi91]. Intuitivamente, esta LDA pode ser obtida da seguinte forma. Considere uma imagem em níveis de cinza representada por uma superfície topográfica em que os níveis indicam a altitude do ponto no relevo. Suponha, agora, que os mínimos em cada uma das bacias (mínimos regionais) que constituem a superfície sejam perfurados, e que a imergimos em um lago com velocidade vertical constante. A água que penetra regularmente pelos orifícios preenche a superfície topográfica e, durante o preenchimento, dois ou mais fluxos vindos de mínimos diferentes podem se unir. Os diques construídos na superfície para evitar que tal junção de águas ocorra constituem a LDA da imagem.

Existem algoritmos eficientes que calculam a função LDA a partir da simulação deste processo de preenchimento de bacias [Mey93], [Soi91]. Um algoritmo iterativo baseado na definição do SKIZ considera os seguintes passos (ver [Vin93] para uma implementação do método). Seja o conjunto X_i obtido por limiarizações sucessivas de uma imagem f com N níveis de cinza:

$$X_i = \{x \in \mathcal{Z}^2, f(x) \leq i\} \quad i = 1, \cdots, N \tag{118}$$

Se o conjunto $SKIZ_X(Y)$ indica, como anteriormente, o esqueleto por zona de influência de Y em X (seção 17), então o conjunto Z, representando a LDA de f, pode ser dado iterativamente por

$$\begin{aligned} Z_j &= SKIZ_{X_j}(X_{j-1} \backslash Z_{j-1}) \quad j = 1, \cdots, N \\ Z &= \bigcup_j Z_j \end{aligned} \tag{119}$$

O algoritmo de LDA define a zona de influência de todos os mínimos regionais de uma imagem e, neste sentido, a sua aplicação resulta em uma super-segmentação da mesma. A Fig. 34(c) apresenta a imagem de LDA do gradiente, Fig. 34(b), calculado sobre a imagem de grãos de feijão na Fig. 34(a). Cada região da Fig. 34(c), delimitada por uma LDA, corresponde à zona de influência do mínimo mais profundo nesta região.

De modo geral, o modelo de segmentação utilizando o algoritmo de LDA é constituído de duas etapas principais: uma "inteligente" que consiste da obtenção de marcadores associados às regiões que pretendemos segmentar, e a outra, automática, representada pelo cálculo da LDA sobre uma imagem f que considera a informação destes marcadores. Esta imagem pode ser o chapéu mexicano, a própria imagem original ou o seu gradiente, por exemplo. O objetivo da primeira etapa, específica à aplicação, é eliminar mínimos não significativos da imagem diminuindo, assim, o efeito da super-segmentação. A segunda etapa define uma imagem de "relevo" adequado à segmentação das regiões. O gradiente, por exemplo, que atribui valores elevados aos pontos de contorno das regiões e mínimos no seu interior pode ser empregado na segmentação da Fig. 34(a).

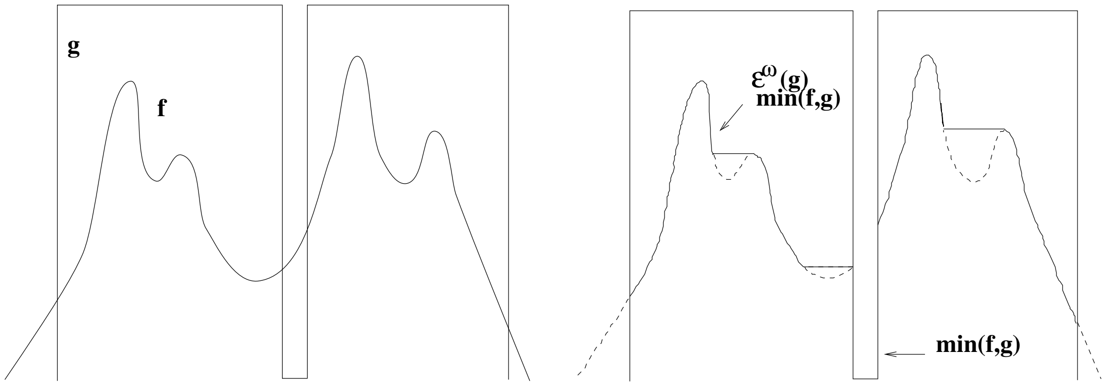
De posse destas imagens, o próximo passo consiste em impor mínimos na imagem gradiente, na posição indicada pelos marcadores obtidos na primeira etapa. Esta imposição especifica as regiões da imagem a serem perfuradas no processo de inundação. As bacias que não forem marcadas não definirão zonas de influência e, conseqüentemente, não constituirão LDA.

Para a imposição dos mínimos, precisamos considerar uma imagem g definida da seguinte forma:

$$g(x) = \begin{cases} +\infty & \text{se} \quad x \notin M \\ 0 & \text{se} \quad x \in M \end{cases} \tag{120}$$

onde M representa o conjunto dos marcadores. Para o cálculo da LDA das regiões marcadas por M, efetuamos inicialmente uma reconstrução geodésica dual (secão 16) de g em $min(f,g)$: $\epsilon^{\infty}_{min(f,g)}$ (g). A Fig. 33 ilustra estas operações. Finalmente, o algoritmo de LDA pode ser aplicado nesta imagem reconstruída, originando uma segmentação mais adequada. Um algoritmo eficiente de segmentação por LDA considerando marcadores é apresentado em [Mey93].

A Fig. 34 apresenta uma aplicação simples deste método. Os marcadores conexos M (Fig. 34(d)), associados a cada um dos grãos, foram obtidos facilmente por limiarização da imagem original [Kak82]. Por exemplo, todos os pixels com nível de cinza abaixo de um determinado valor (30, neste caso) fazem parte do conjunto M. A este conjunto, precisamos acrescentar um marcador para o fundo da imagem gradiente (por quê?). Observe a separação dos grãos que se tocam no resultado da



(a) (b)

Figura 33: Imposição de mínimos (a), e reconstrução dual (b).

segmentação (Fig. 34(e)).

Referências

[Bar94] G. J. F. Banon & J. Barrera. *Bases da Morfologia Matemática para a Análise de Imagens Binárias*. IX Escola de Computação, Recife, PE, Julho 1994.

[Che89] M. Coster & J. L. Chermant. *Précis d'Analyse d'Images*. Presses du CNRS, 1989.

[ea83] A. C. Borik et alli. A Generalization of Median Filtering using Combination of Order Statistics. In *IEEE Proceedings, 71(31)*, pages 1342–1350, 1983.

[Hey82] G. Heygster. Rank filters in digital image processing. *Computer Graphics and Image Processing*, 17:148–164, 1982.

[JC93] J. Serra & R. W. Schafer J. Crespo. Image Segmentation using Connected Filter. In *Int. Workshop on Mathematical Morphology and Its Application to Signal Processing*, pages 52–57, Barcelona, Spain, May 1993.

[Kak82] A. Rosenfeld & A. Kak. *Digital Picture Processing*, volume 1. Academic Press, 1982.

[Kor88] T. W. Korner. *Fourier Analysis*. Cambridge Univ., 1988.

[Mai84] C. Lantuejoul & F. Maisonneuve. Geodesic Methods in Quantitative Image Analysis. *Pattern Recognition*, 17(2):177–187, 1984.

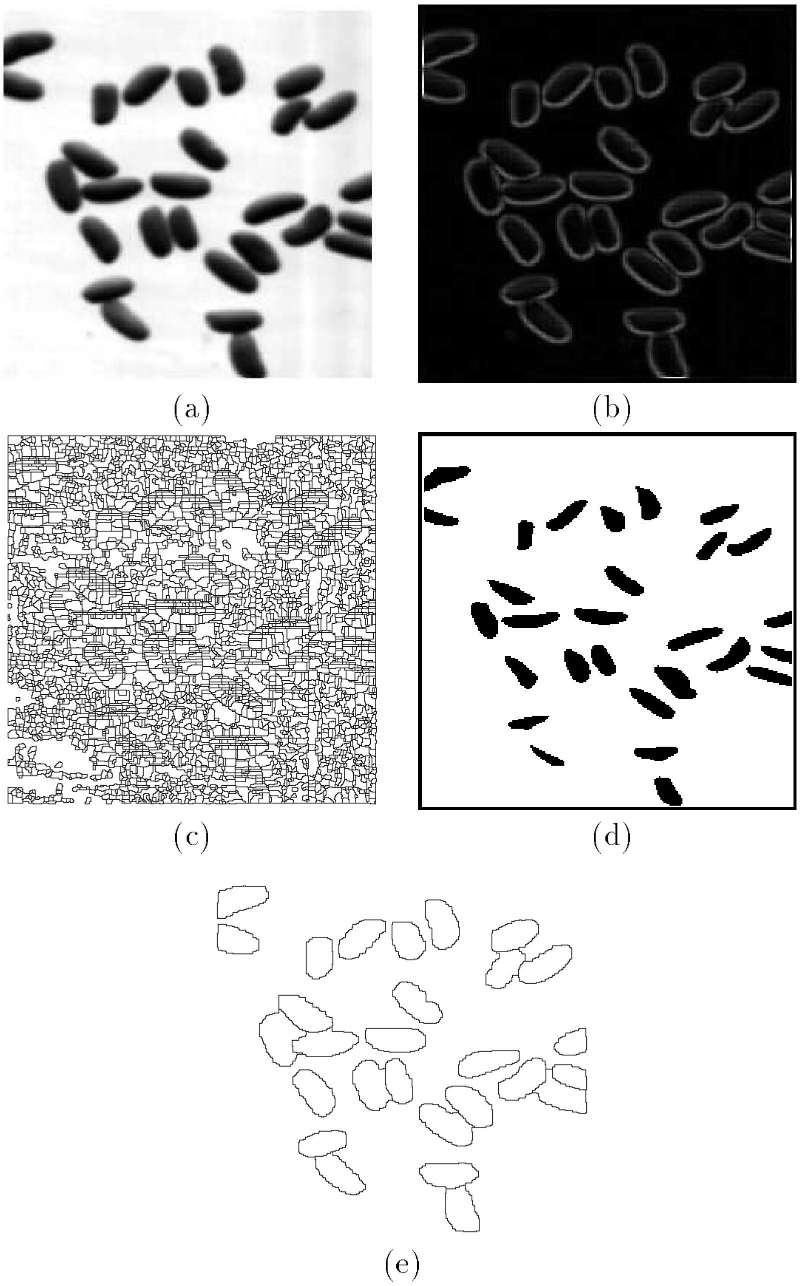


Figura 34: Imagem de grãos (a), imagem gradiente (b), e sua LDA (c). Marcadores obtidos por limiarização (d), e segmentação por LDA (e).

[Mat75] G. Matheron. *Random Sets and Integral Geometry.* John Wiley & Sons, 1975.

[Mey93] S. Beucher & F. Meyer. *Mathematical Morphology in Image Processing*, chapter 12. Marcel Dekker, Inc, 1993.

[Pap62] A. Papoulis. *The Fourier Integral and its Applications.* McGraw-Hill, 1962.

[Ser82] J. Serra. *Image Analysis and Mathematical Morphology.* Academic Press, 1982.

[Ser88] J. Serra. *Image Analysis and Mathematical Morphology, volume 2: Theoretical Advances.* Academic Press, 1988.

[Ser94] J. Serra. Morphological filtering: An overview. *Signal Processing*, (38):3–11, 1994.

[Soi91] Luc Vincent & P. Soile. Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations. *IEEE Transactions on PAMI*, 13(6):583–598, 1991.

[Vac95] Corinne Vachier. *Extraction de Caractéristiques, Segmentation d'Images et Morphologie Mathématique.* PhD thesis, ENSMP - Fontainebleau, França, 1995.

[Vin93] Luc Vincent. Morphological Grayscale Reconstruction in Image Analysis: Applications and Efficient Algorithms. *IEEE Transactions on Image Processing*, 2(2):176–201, 1993.

[Woo93] R. Gonzalez & R. R. Woods. *Digital Image Processing.* Addison-Wesley, 1993.