

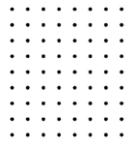
**Malha Digital:**

**Malha digital:** É o suporte no qual as imagens digitais são representadas e através do qual se estabelecem as conexões entre pontos vizinhos.

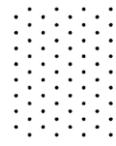
**Estrutura da malha digital:** antes de se fazer qualquer operação sobre uma imagem digital, deve-se conhecer primeiro a estrutura de relações espaciais entre os pixels, ou seja, definir quais as relações de vizinhança entre os pixels da malha digital, para se poder saber de que forma se distinguirão objectos constituídos por conjuntos de pixels, assim como definir as suas respectivas fronteiras.

- Para tal, há que estabelecer a geometria da malha digital de representação e caracterizá-la topologicamente.

Num largo número de sistemas digitais de aquisição de imagem, a sua estrutura espacial é definida segundo uma malha ortogonal e equidistante, sendo cada ponto dessa malha chamado de "pixel". Regra geral, os pixels têm forma quadrada. A geometria da malha é que pode variar.



Malha rectangular



Malha hexagonal

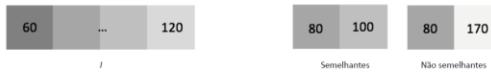
**Conectividade e Vizinhança:**

As regras de conectividade estabelecem a relação entre dois ou mais pixels.

De uma forma geral, para que dois pixels sejam considerados como conectados, têm que verificar condições de adjacência radiométrica e adjacência espacial.

Em primeiro lugar, para que dois pixels se considerem conectados radiometricamente, têm ambos que ter valores pertencentes a um mesmo intervalo numérico (classe numérica), ou conjunto de valores numéricos, podendo dizer-se que assim pertencerão a um mesmo tipo de objeto.

Por exemplo, numa imagem de cinzentos de 8-bits, estabelecendo o critério de similaridade de serem semelhantes todos os pixels com valores dentro do intervalo  $I = [60, \dots, 120]$ , dois pixels com valores respectivamente de 80 e 100 são semelhantes, estando espacialmente próximos, ou não. Outros dois com valores de 80 e 170, não são semelhantes.



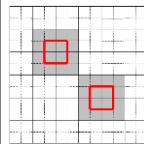
Para formular o critério de conectividade espacial, introduz-se primeiro o conceito de **vizinhança**.

A vizinhança de conectividade-4 do pixel de coordenadas  $(x, y)$  é o conjunto de pixels  $V_4$ .

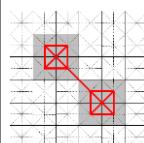
$$V_4(x, y) = \{(x, y), (x + 1, y), (x - 1, y), (x, y + 1), (x, y - 1)\}$$

A vizinhança de conectividade-8 do pixel de coordenadas  $(x, y)$  é o conjunto de pixels  $V_8$ .

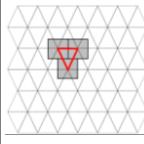
$$V_8(x, y) = V_4(x, y) \cup \{(x + 1, y + 1), (x + 1, y - 1), (x - 1, y + 1), (x - 1, y - 1)\}$$



Na malha de conectividade-4 há um máximo de quatro pixels vizinhos a que cada pixel se pode conectar.

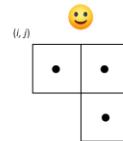
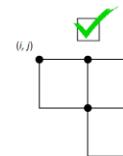


Na malha de conectividade-8 há um máximo de oito pixels vizinhos a que cada pixel se pode conectar.

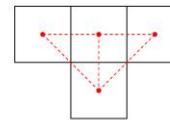


A malha hexagonal é de conectividade-6 e, neste caso, a vizinhança de cada pixel é constituída sempre por sete pixels.

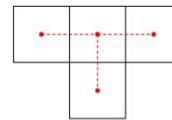
Por definição, as coordenadas **imagem** de um pixel devem referir-se ao seu canto superior esquerdo. Contudo, para facilitar a compreensão de vários conceitos, será por vezes ilustrada a sua posição no seu centro geométrico.



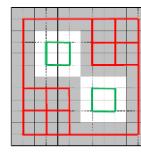
O **grafo de adjacência** de um certo conjunto de pixels é representado pelo conjunto de ligações elementares que conectam esses pixels.



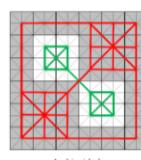
Grafo de conectividade 8



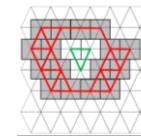
Grafo de conectividade 4



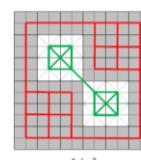
Para a malha de conectividade-4, os grafos correspondentes ao conjunto e ao seu complementar, não se interseccionam, o que não cria ambiguidades lógicas.



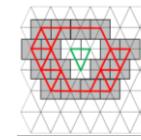
No caso da conectividade-8, há uma ambiguidade lógica.



A intersecção entre os grafos do conjunto e do seu complementar é diferente de conjunto vazio. Como tal, quando se usa a malha digital quadrada, constitui-se a regra de considerar-se a conectividade-8 para um conjunto e a conectividade-4 para o seu complementar, ou vice-versa.



Solução

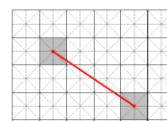


No caso da malha hexagonal não existe ambiguidade lógica entre um conjunto e o seu complementar.

**Distâncias:**

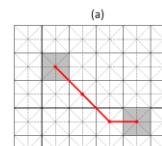
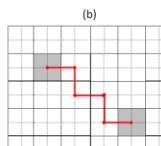
Numa malha digital, as distâncias entre pixels são geralmente duas: **distância euclidiana** e **distância grafológica**.

**Distância euclidiana:** calculada através das diferenças de coordenadas-pixel entre os dois pixels.



$$|AM| = 3; |AP| = 2; Dist = \sqrt{13} \approx 3.6$$

**Distância grafológica:** corresponde ao valor da menor soma de conexões que ligam os dois pixels dentro do grafo de adjacência da malha digital.

(a): distância grafológica numa malha de conectividade-8 (Chessboard Distance):  $Dist = 3$ ;(b): distância grafológica numa malha de conectividade-4 (City Block Distance):  $Dist = 5$ .

**Resoluções:**

Descreve o nível de detalhe de uma imagem. O termo aplica-se a imagens digitais, imagens em sequências de vídeo, ou outros tipos de imagem. Resoluções mais altas significam maior detalhe.

Tipos de resolução associados às imagens digitais:

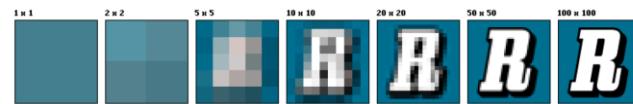
1. Número de pixels (*pixel count*)
2. Pixels por polegada (*PPI*)
3. Pontos por polegada (*DPI*)
4. Espacial (*spatial resolution*)
5. Espectral (*spectral resolution*)
6. Radiométrica (*radiometric resolution*)
7. Temporal (*temporal resolution*)

**Número de pixels (pixel count):** quando a contagem dos pixels de uma imagem é referida como "resolução", a convenção descreve-a como o conjunto de dois números inteiros positivos em que o primeiro corresponde ao número de colunas (*width*) e o segundo corresponde ao número de linhas (*height*).

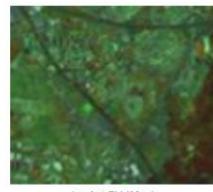
Outra convenção comum para quantificar o número de pixels consiste na contagem em mega-pixels, que pode ser calculada multiplicando o número de colunas pelo número de linhas, e dividindo por um milhão (1 mega-pixel =  $10^6$  pixels).

Por exemplo: uma imagem de 2048 colunas por 1536 linhas tem um total de  $2048 \times 1536 = 3145728$  pixels, ou seja, 3.1 mega-pixels.

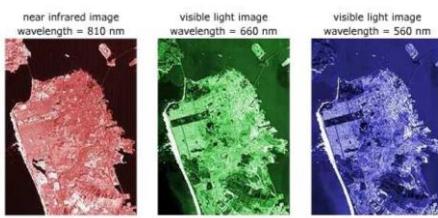
O tamanho de uma imagem digital está, pois, relacionado com a quantidade de pixels que tem e não com o tamanho da sua representação num ecrã.



**Resolução espacial:** consiste na capacidade de distinção espacial, na imagem, entre dois objectos próximos (não é o tamanho do menor objecto que pode ser visto na imagem). Quantifica-se pela área real que um pixel da imagem representa.



**Resolução espectral:** poder de resolução do espectro de frequência, ou seja, medida do poder de resolução dos valores do comprimento de onda, dentro do espectro electromagnético.

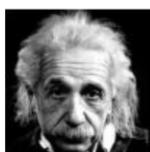


Actualmente, a maioria dos satélites mede a energia electromagnética em muitos comprimentos de onda diferentes. Tal é designado como imagem multiespectral.

Imagens tiradas em diferentes comprimentos de onda podem ser combinadas entre si, criando imagens compostas.

**Resolução radiométrica:** poder de distinção de diferenças de intensidade dos objectos. É geralmente expressa em níveis de cinzento, ou número de bits.

O número total de níveis de cinzento que uma imagem de  $n$  bits pode ter é igual a  $2^n$ . Por exemplo:



No caso de uma imagem colorida (RGB), esta é constituída por uma composição de três componentes de 8 bits cada, ou seja, a resolução radiométrica é de 24 bits.

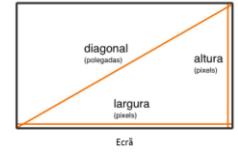
**"Pixels por polegada", ou "pixels per inch" (PPI):** é uma medida de densidade ou resolução de uma tela de computador, tela de televisão ou outro dispositivo de exibição. O valor de PPI indica quantos pixels estão em uma linha de uma polegada, num ecrã, independentemente se ela está na vertical, ou na horizontal (tratando-se de pixels quadrados).

O tamanho do ecrã não quer dizer absolutamente nada quando se fala em qualidade. Saber o número em polegadas de um ecrã, ou *display* de um *smartphone*, não diz nada acerca da sua resolução. Tratando-se de qualidade, o número que realmente importa é o que se expressa em pixels. Basicamente, quanto mais pixels houver em um mesmo espaço, melhor será a definição das imagens e, consequentemente, melhor será a qualidade visual.

Cálculo do valor da resolução PPI de um ecrã:

$$\text{diagonal (pixels)} = \sqrt{\text{largura}^2 + \text{altura}^2}$$

$$PPI = \frac{\text{diagonal (pixels)}}{\text{diagonal (polegadas)}}$$



Exemplo:

Resolução de um ecrã de 1920x1080 pixels (Full HD), com 15"

$$PPI = \frac{\sqrt{1920^2 + 1080^2}}{15} = 96$$

**"Pontos por polegada", ou "dots per inch" (DPI):** no momento da impressão de uma imagem em papel, o pixel deixa de existir e "o que se vê" no papel são apenas minúsculos pontos que a compõem. O valor de DPI corresponde à resolução com que a imagem será imprimida e depende dos valores em pixels, já que é necessário fazer uma conversão antes.

Nas fases, quer de impressão, quer de digitalização de uma imagem, deve-se ter em atenção a escolha do valor de DPI, pois tal influencia diretamente o tamanho e a qualidade da foto. Normalmente não é preciso um valor muito alto, podendo-se usar, para a maior parte dos casos, os valores de referência de 72, 100, ou 300.

Partindo da imagem digitalizada, coloca-se a questão ao contrário: conhece-se o valor de DPI e a sua dimensão (em pixels). Pode-se então determinar uma das suas dimensões de impressão em papel. Se este comprimento for ultrapassado, na impressão, ou na sua apresentação num ecrã, a qualidade pode ficar comprometida.



$$\text{diagonal (polegadas)} = \frac{\text{diagonal (pixels)}}{\text{DPI}}$$

Exemplo:

Imagen de 256x256 pixels

$$\text{diagonal (cm)} = \text{diagonal (polegadas)} \times 2.5400051 =$$

$$= \frac{\sqrt{256^2 + 256^2}}{96} \times 2.5400051 = 3.77 \times 2.5400051 = 9.57 \text{ cm}$$

A distância máxima de impressão, expressa em centímetros (1 polegada = 2.5400051 cm), de um comprimento de  $k$  pixels a  $n$  DPI, pode calcular-se da seguinte forma:

$$\text{distância (cm)} = \frac{k}{n} \times 2.5400051$$

Por exemplo, para uma imagem com resolução de 1920x1080 pixels, a 72 DPI, tem-se:

- Largura da imagem no papel =  $1920/72 \times 2.5400051 = 67,73 \text{ cm}$
- Altura da imagem no papel =  $1080/72 \times 2.5400051 = 38,10 \text{ cm}$

Para uma imagem com resolução de 1920x1080 pixels tem-se, a 300 DPI, tem-se:

- Largura da imagem no papel =  $1920/300 \times 2.5400051 = 16,25 \text{ cm}$
- Altura da imagem no papel =  $1080/300 \times 2.5400051 = 9,14 \text{ cm}$

**Resolução temporal:** intervalo de tempo depois do qual é adquirida informação acerca de um mesmo objecto.

Séries de imagens que mostram um mesmo objecto, tiradas em momentos diferentes (ao longo do tempo) são referidas como informação multi-temporal.

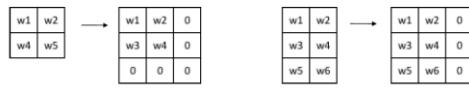
**Filtros Digitais 2D:**

Seja  $H$  um filtro bidimensional, caracterizado por uma matriz. Por conveniência, considere-se que  $H$  respeita as seguintes condições:

- $H$  é uma matriz quadrada.
- $H$  tem um número ímpar de elementos, ou seja, tem  $(2N+1) \times (2N+1)$  elementos, e que estes estão indexados desde  $-N$  até  $N$ , tal que o elemento central de  $H$  é  $H(0, 0)$ .
- Os valores numéricos ( $w_i$ ) de  $H$  designam-se por "coeficientes".

Por exemplo:

-1	0	1
-w1	w2	w3
w4	w5	w6
w7	w8	w9

**Correlação / Convolução:**

Sendo  $f$  uma imagem matricial, define-se a operação de **correlação** (" $\otimes$ ") entre  $H$  e  $f$  por:

$$H \otimes f(x, y) = \sum_{i=-N}^N \sum_{j=-N}^N H(i, j) \times f(x+i, y+j)$$

Por exemplo, para um  $H$  de  $3 \times 3$  tem-se:

$$\begin{aligned} H \otimes f(x, y) &= \sum_{i=-1}^1 \sum_{j=-1}^1 H(i, j) \times f(x+i, y+j) = \\ &= H(-1, -1) \times f(x-1, y-1) + H(-1, 0) \times f(x-1, y) + H(-1, 1) \times f(x-1, y+1) + \\ &\quad + H(0, -1) \times f(x, y-1) + H(0, 0) \times f(x, y) + H(0, 1) \times f(x, y+1) + \\ &\quad + H(1, -1) \times f(x+1, y-1) + H(1, 0) \times f(x+1, y) + H(1, 1) \times f(x+1, y+1) \end{aligned}$$

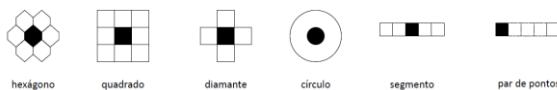
Em resumo, numa operação de filtragem espacial por convolução/correlação, uma das matrizes de entrada é geralmente uma imagem de níveis de cinzento ( $f$ ) e a outra, geralmente bastante mais pequena, é o filtro ( $H$ ). Dentro desta estabelece-se uma posição de referência (designado por "pixel central", mas que não tem que ser necessariamente o que se encontra no seu centro).

Para os pixels de fronteira da imagem há que fazer uma adaptação para a executar o processo de filtragem. Uma das quatro seguintes opções pode ser seguida:

1. Usa-se apenas a convolução que considere os subconjuntos de pixels de  $H$  que estejam dentro dos limites da imagem  $f$ .
2. São escolhidos valores iguais a zero para os pixels das regiões que estão fora da imagem, mas tal escolha pode distorcer a intensidade dos pixels de fronteira na imagem.
3. Acrescenta-se linhas e colunas à imagem. Cada pixel destas, terá um valor igual ao pixel da imagem que dele estiver mais próximo.
4. Acrescenta-se linhas e colunas à imagem, por forma a que reflita uma continuidade de carácter periódico, do interior para o exterior da imagem.

**Propriedades dos filtros digitais:**

Para além dos valores dos coeficientes associados, a forma e a dimensão são características segundo as quais  $H$  também pode variar. Eis alguns exemplos de geometrias de  $H$ .



No caso de  $H$  não ter, à partida, um número ímpar de elementos, as anteriores condições podem ser estabelecidas, pois pode-se pegar em qualquer janela e preencher-lá com zeros, por forma a que passe a ser quadrada e com um número ímpar de elementos. Esta operação não muda o comportamento do filtro  $H$  original.

A operação de **convolução** (" $*$ ") é semelhante à de correlação. A diferença consiste em primeiro rodar  $H$  de 180 graus, e só então aplicar a operação de correlação.

$$H * f(x, y) = \sum_{i=-N}^N \sum_{j=-N}^N H(i, j) \times f(x-i, y-j)$$

Note-se que a correlação e a convolução são operações idênticas se  $H$  for simétrico.

A diferença entre a correlação e a convolução é a de que a última respeita a propriedade associativa. Ou seja, se  $G$  e  $H$  são filtros, então,

$$G * (H * f) = (G * H) * f$$

- A verificação desta propriedade torna-se bastante conveniente quando, por exemplo, se pretende aplicar mais do que um filtro a uma imagem.
- Como geralmente a dimensão da imagem é significativamente maior do que a do filtro, o esforço de cálculo é reduzido, executando-se a convolução entre os dois filtros, seguida da convolução entre o filtro resultante e a imagem.

A classificação das janelas de convolução faz-se segundo duas propriedades: **convexidade** e a **isotropia**.

	Isotrópico	Anisotrópico
Convexo	Disco	Segmento
Não convexo	Contorno do disco	Par de pontos

**Filtragem Espacial:**

A frequência espacial de uma imagem é uma característica que pode ser definida pelo número de variações de níveis de cinzento por unidade de distância.

Se os valores numéricos de uma certa área oscilam pouco, diz-se que essa área tem variações de baixa frequência; se oscilam muito diz-se que essa zona é de alta frequência.

Os **filtros espaciais** 2D são operadores que permitem alterar a frequência espacial de uma imagem, modificando o valor do tom de cinzento de cada pixel em função dos valores dos tons de cinzento dos pixels da sua vizinhança.

Os filtros podem ser lineares ou não-lineares. Nos **filtros lineares** cada pixel resulta de uma combinação linear entre os pixels da sua vizinhança, com coeficientes que correspondem aos pesos a atribuir às parcelas.

$$f_1 = H * f(x, y) = \sum_{i=-N}^N \sum_{j=-N}^N H(i, j) \times f(x-i, y-j)$$

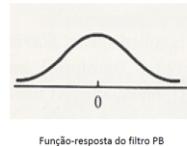
$f_1$ : imagem filtrada  
 $f$ : imagem inicial  
 $H$ : filtro

Quaisquer outros filtros são designados por **filtros não lineares**.

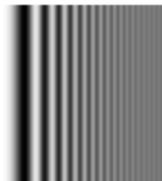
**Filtro Passa-Baixa:**

Filtro que suaviza o aspecto da imagem atenuando eventos de elevada frequência (transições abruptas), isto é, as zonas de fronteira radiométrica. Tende a minimizar ruídos e o resultado apresenta um efeito de desfocagem.

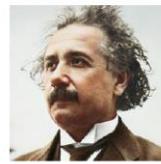
A forma da função resposta necessária para implementar um filtro passa-baixa indica que este deve ter todos os seus coeficientes positivos, devendo a respectiva soma ser igual a 1.



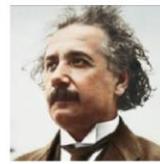
Função-resposta do filtro PB



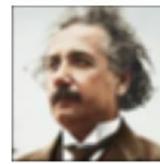
**Média:** É o mais simples filtro linear passa-baixa. Todos os coeficientes são iguais. Calcula-se a média dos tons de cinzento no interior da janela (H) e substitui-se o pixel central da janela pelo valor resultante (por esta razão as janelas são normalmente quadradas com dimensão ímpar (3x3, 5x5, etc.).



Inicial



Média 3x3



Média 9x9

$$H_{3 \times 3} = \frac{1}{9} \times \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$H_{9 \times 9} = \frac{1}{81} \times \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}_{13}$$

**Mediana:** é um filtro não-linear de suavização. O resultado é geralmente melhor que o do filtro da média, quando usado em imagens com ruído do tipo Sal-e-Pimenta, ou Speckle. Ainda, tende a preservar melhor as fronteiras dos objectos.

123	125	126	130	140
122	124	126	127	135
118	120	150	125	134
119	115	119	123	133
111	116	110	120	130

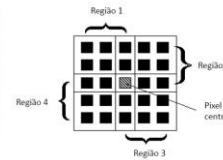
Valores da vizinhança

115, 119, 120, 123, 124,  
125, 126, 127, 150

Mediana: 124

**Kuwahara:** Suaviza uma imagem sem perturbar a nitidez e a posição das fronteiras.

Embora possa ser implementado em janelas de formas diversas, considera-se aqui uma janela quadrada de dimensão ímpar. Esta janela é dividida em 4 regiões e em cada uma delas calcula-se a intensidade média  $m_i$  e a variância  $s_i^2$ , ( $i = 1, 2, 3, 4$ ). O valor atribuído ao pixel central da janela corresponde ao valor médio da janela que tem menor variância.



Por exemplo, no filtro passa-baixa da média aritmética  $3 \times 3$ , tem-se:

$$\begin{bmatrix} f_{i-1,j-1} & f_{i-1,j} & f_{i-1,j+1} \\ f_{i,j-1} & f_{i,j} & f_{i,j+1} \\ f_{i+1,j-1} & f_{i+1,j} & f_{i+1,j+1} \end{bmatrix}$$

1/9 ×

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Coeficientes

$$\begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}$$

$$\begin{bmatrix} & & \\ & & \\ & & \end{bmatrix}$$

PB<sub>ij</sub>

Valor calculado

Por convolução, tem-se:

$$PB_{i,j} = \frac{1}{9} \times f_{i-1,j-1} + \frac{1}{9} \times f_{i-1,j} + \frac{1}{9} \times f_{i-1,j+1} + \frac{1}{9} \times f_{i,j-1} + \frac{1}{9} \times f_{i,j} + \frac{1}{9} \times f_{i,j+1} + \frac{1}{9} \times f_{i+1,j-1} + \frac{1}{9} \times f_{i+1,j} + \frac{1}{9} \times f_{i+1,j+1}$$

**Gaussiano 2-D:** é um operador de convolução que utiliza um kernel representado sob a forma de uma "bossa" gaussiana (em forma de sino).

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} \times e^{-\left(\frac{(x-x_0)^2}{2\sigma^2}\right)}$$

$$x, y = np.meshgrid(np.linspace(-1, 1, 2*el+1), np.linspace(-1, 1, 2*el+1))$$

$$G = \lambda \text{ambda } x, y: 1/(2*el*pi*sigma**2)*np.exp(-(x**2+y**2)/(2*sigma**2))$$

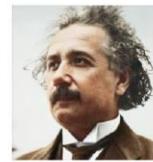
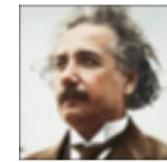
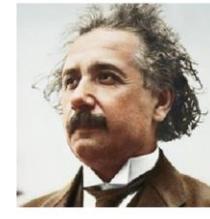
Exemplo de  $7 \times 7$ , centrada em  $(0, 0)$ , com  $\sigma^2 = 1$ :

$$\begin{bmatrix} 0.000192 & 0.000208 & 0.000224 & 0.000239 & 0.000254 & 0.000268 & 0.000282 \\ 0.000208 & 0.000225 & 0.000242 & 0.000259 & 0.000275 & 0.000292 & 0.000306 \\ 0.000224 & 0.000242 & 0.000260 & 0.000277 & 0.000294 & 0.000310 & 0.000326 \\ 0.000239 & 0.000259 & 0.000277 & 0.000294 & 0.000312 & 0.000329 & 0.000346 \\ 0.000254 & 0.000275 & 0.000294 & 0.000312 & 0.000331 & 0.000350 & 0.000368 \\ 0.000268 & 0.000292 & 0.000312 & 0.000331 & 0.000351 & 0.000370 & 0.000389 \\ 0.000282 & 0.000306 & 0.000329 & 0.000350 & 0.000370 & 0.000390 & 0.000410 \end{bmatrix}$$

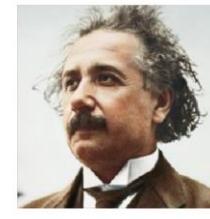
$$= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 1 & 1 & 0 \\ 0 & 1 & 3 & 2 & 2 & 2 & 0 \\ 0 & 2 & 5 & 4 & 4 & 4 & 0 \\ 0 & 1 & 5 & 3 & 3 & 3 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \times \frac{1}{49}$$

O filtro gaussiano é uma média ponderada, com maior peso aplicado ao pixel central, diminuindo progressivamente para o exterior. Como tal, proporciona uma suavização mais "delicada" que o da média, preservando melhor as fronteiras entre objectos.

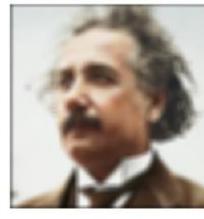
O grau de suavização é determinado pelo valor da variância (ou do desvio-padrão) da função de Gauss (funções com desvios-padrão mais altos requerem janelas de convolução maiores no sentido de as funções ficarem mais bem representadas).

Gaussiana ( $\sigma^2 = 0.5$ )Gaussiana ( $\sigma^2 = 5$ )

Inicial



Gaussiano 9x9



Média 9x9

**Filtros de borda:****Detecção de borda:****Segmentação:**

**Filtro Passa-Alta:**

A contrário da filtragem passa-baixa, que esbate/elimina os eventos de detalhe contidos nas imagens, a diferenciação vai ter o efeito contrário, ou seja, evidenciar o detalhe. Definem-se assim outros filtros designados por filtros derivativos (operadores de gradiente).

**Gradiente:** O gradiente de uma função  $f$ , no ponto  $(x,y)$ , define-se por  $\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$

A magnitude é dada por  $mag(\nabla f) = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}$ .

Estes conceitos constituem a base de diversas abordagens de diferenciação da imagem.

Filtro que atenua, ou elimina, os eventos da imagem com baixa frequência, pelo que os filtros tornam mais nítidas as fronteiras radiométricas e os detalhes.

A forma da função resposta necessária para implementar um filtro passa-alta indica que este deve ter os coeficientes positivos na vizinhança do centro e negativos na periferia, devendo a respectiva soma ser igual a 0.



**Roberts:** este filtro executa o gradiente cruzado, isto é, em vez de calcular as diferenças de valores de brilho na direção vertical e horizontal, fá-lo numa direção rodada de 45º, onde as janelas de convolução são as seguintes:

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

$$f(z_5) = mag(\nabla f) \approx \sqrt{(z_5 - z_9)^2 + (z_6 - z_8)^2}$$

$$\begin{array}{c} \rightarrow x \\ \downarrow y \\ \begin{bmatrix} z_1 & z_2 & z_3 \\ z_4 & z_5 & z_6 \\ z_7 & z_8 & z_9 \end{bmatrix} \end{array}$$

$$f(z_5) = mag(\nabla f) \approx |z_5 - z_9| + |z_6 - z_8|$$

**Sobel:** este operador realça linhas verticais e horizontais mais escuras que o fundo, sem realçar pontos isolados. Consiste na aplicação de duas máscaras, descritas a seguir, que compõem um resultado único.

$$\begin{array}{c} \rightarrow x \\ \downarrow y \\ \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \end{array}$$

$$S_x = (z_1 + 2 \times z_4 + z_7) - (z_3 + 2 \times z_6 + z_9)$$

$$S_y = (z_1 + 2 \times z_2 + z_3) - (z_7 + 2 \times z_8 + z_9)$$

$$f(z_5) = |S_x| + |S_y|$$

$$\begin{array}{c} \rightarrow x \\ \downarrow y \\ \begin{bmatrix} z_1 & z_2 & z_3 \\ z_4 & z_5 & z_6 \\ z_7 & z_8 & z_9 \end{bmatrix} \end{array}$$

**Prewitt:** este operador realça linhas verticais e horizontais mais escuras que o fundo, sem realçar pontos isolados. Consiste na aplicação de duas máscaras, descritas a seguir, que compõem um resultado único.

$$\begin{array}{c} \rightarrow x \\ \downarrow y \\ \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \end{array}$$

$$S_x = (z_1 + z_4 + z_7) - (z_3 + z_6 + z_9)$$

$$S_y = (z_1 + z_2 + z_3) - (z_7 + z_8 + z_9)$$

$$f(z_5) = |S_x| + |S_y|$$

$$\begin{array}{c} \rightarrow x \\ \downarrow y \\ \begin{bmatrix} z_1 & z_2 & z_3 \\ z_4 & z_5 & z_6 \\ z_7 & z_8 & z_9 \end{bmatrix} \end{array}$$

Considerando a janela da figura, pode aproximar-se a equação anterior no ponto  $z_5$  de diversas formas. A mais simples é utilizar a diferença  $(z_5 - z_6)$  para definir a derivada parcial na direcção x e a diferença  $(z_5 - z_8)$  para definir a derivada parcial na direcção y.

$$\begin{array}{c} \rightarrow x \\ \downarrow y \\ \begin{bmatrix} z_1 & z_2 & z_3 \\ z_4 & z_5 & z_6 \\ z_7 & z_8 & z_9 \end{bmatrix} \end{array}$$

$$f(z_5) = mag(\nabla f) \approx \sqrt{(z_5 - z_6)^2 + (z_5 - z_8)^2}$$

ou

$$f(z_5) = mag(\nabla f) \approx |z_5 - z_6| + |z_5 - z_8|$$

Filtro Passa-Alta = Imagem original - Filtro Passa-Baixa

Por exemplo, a partir do filtro da média do slide 14, tem-se:

$$\begin{aligned} PA_{i,j} &= f_{i,j} - PB_{i,j} = \\ &= f_{i,j} - \left( \frac{1}{9} \times f_{i-1,j-1} + \frac{1}{9} \times f_{i-1,j} + \frac{1}{9} \times f_{i-1,j+1} + \frac{1}{9} \times f_{i,j-1} + \frac{1}{9} \times f_{i,j} + \frac{1}{9} \times f_{i,j+1} + \frac{1}{9} \times f_{i+1,j-1} + \frac{1}{9} \times f_{i+1,j} + \frac{1}{9} \times f_{i+1,j+1} \right) = \\ &= -\frac{1}{9} \times f_{i-1,j-1} - \frac{1}{9} \times f_{i-1,j} - \frac{1}{9} \times f_{i-1,j+1} - \frac{1}{9} \times f_{i,j-1} + \frac{8}{9} \times f_{i,j} - \frac{1}{9} \times f_{i,j+1} - \frac{1}{9} \times f_{i+1,j-1} - \frac{1}{9} \times f_{i+1,j} - \frac{1}{9} \times f_{i+1,j+1} \end{aligned}$$

$t_{4,4,4}$	$t_{4,4}$	$t_{4,4,4}$
$t_{4,4}$	$t_{4}$	$t_{4,4,4}$
$t_{4,4,4}$	$t_{4,4}$	$t_{4,4,4}$

$$\begin{array}{c} \text{Função} \\ 1/9 \times \end{array}$$

-1	-1	-1
-1	8	-1
-1	-1	-1

$$\begin{array}{c} \text{Coeficientes} \\ -1/9 \quad 8/9 \quad -1/9 \end{array}$$

PA <sub>ij</sub>
------------------

$$\begin{array}{c} \text{Valor a calcular} \\ \end{array}$$

Roberts (exemplo 1):

$$\begin{array}{c} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \\ \text{ou} \\ \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad g_1(i,j) = f(i,j) - f(i+1,j+1) \\ \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad g_2(i,j) = f(i,j+1) - f(i+1,j) \\ g_1(i,j) = f(i,j) - f(i+1,j+1) \\ g_2(i,j) = f(i,j+1) - f(i+1,j) \\ g(i,j) = |g_1(i,j)| + |g_2(i,j)| \end{array}$$

Sobel (exemplo):

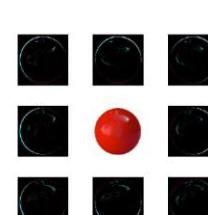
$$\begin{array}{c} \begin{bmatrix} -1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & -1 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ S_x \\ \begin{bmatrix} -1 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & 2 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ S_y \\ |S_x| + |S_y| \end{array}$$

Prewitt (completar):

$$\begin{array}{c} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \\ S_x \\ \begin{bmatrix} -1 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & 2 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ S_y \\ |S_x| + |S_y| \end{array}$$

Os filtros passa-alta podem ser "desenhados" em função da direcção. Neste caso o kernel contém coeficientes que variam em função da orientação que apresentam na imagem as fronteiras que se pretende realçar.

$$\begin{array}{c} \text{NW} \quad \text{N} \quad \text{NE} \\ \begin{bmatrix} 1 & 1 & 1 \\ -2 & -2 & -1 \\ -1 & -1 & -1 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & -1 \\ -1 & -1 & -1 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 \\ -1 & -2 & -1 \\ -1 & -1 & -1 \end{bmatrix} \\ \text{W} \quad \text{E} \\ \begin{bmatrix} 1 & 2 & -1 \\ -1 & -2 & -1 \\ 1 & 1 & -1 \end{bmatrix} \quad \begin{bmatrix} -1 & -2 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} \\ \text{SW} \quad \text{S} \quad \text{SE} \\ \begin{bmatrix} -1 & -2 & -1 \\ 1 & 2 & -1 \\ 1 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} -1 & -1 & -1 \\ -1 & -2 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & -2 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \end{array}$$





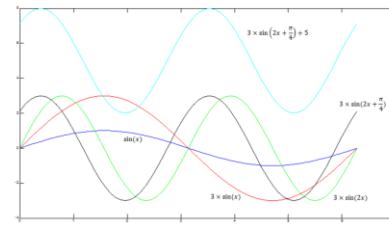
**Sinal e série amostral:**

Uma onda periódica elementar como o seno fica definida por uma das seguintes expressões (em função do período T, ou da frequência f):

$$P(t) = A \times \sin\left(\frac{2\pi}{T}t + \varphi\right) + C$$

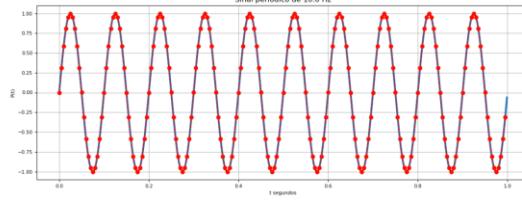
$$P(t) = A \times \sin(2\pi ft + \varphi) + C$$

A: amplitude  
f: frequência linear  
T: período  
 $\varphi$ : fase  
C: translação  
 $\omega = 2\pi/f$ : frequência angular

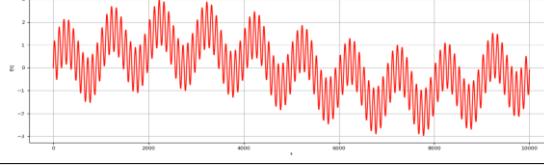


Num sinal periódico P, o valor da frequência f, quando expressa em Hertz (Hz), corresponde ao número de ciclos m que ocorre no intervalo de tempo de um segundo.

Num processo de digitalização de um sinal analógico, a amostragem de P consiste em gerar uma sequência discreta S de impulsos (série amostral) com uma frequência de amostragem  $f_s$ .

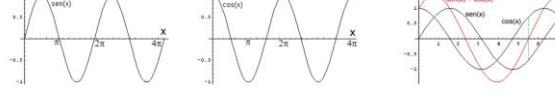


Considere-se um segundo exemplo sintético, produzido a partir da soma de três senoidais com diferentes frequências não conhecidas e que, apesar de igualmente periódica, tem um comportamento mais difícil de interpretar. O objetivo será o de saber as características das ondas elementares. Para tal recorre-se à análise espectral com a transformada de Fourier.

**Transformada de Fourier:**

A teoria de Fourier afirma que qualquer sinal pode ser expresso como uma soma de várias senoidais, ou seja, uma soma de funções seno e cosseno.

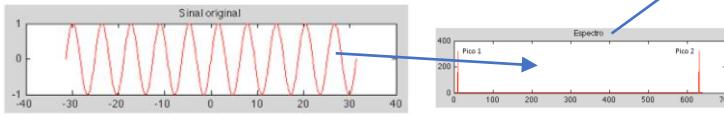
$$f(x) = a_0 + a_1 \sin(x) + a_2 \sin(2x) + a_3 \sin(3x) + \dots + b_1 \cos(x) + b_2 \cos(2x) + b_3 \cos(3x) + \dots$$



A transformada de Fourier (TF) permite realizar a decomposição de um sinal (definido no domínio espacial) nas suas componentes seno e cosseno, representando-o no chamado "domínio das frequências", ou espaço de Fourier.

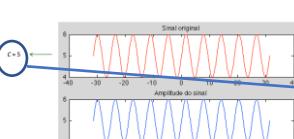
O espectro de frequências de uma função resulta da aplicação da DFT. Os valores resultantes desta transformação pertencem ao domínio dos números complexos ( $z = a+bi$ ), sendo constituídos por uma parte real (a) e uma parte imaginária (b). A magnitude do espectro é igual ao valor absoluto da DFT ( $|z|$ )

Considere-se o seguinte sinal periódico (seno).



O espectro de um sinal contém as magnitudes das frequências das ondas que o constituem. Neste exemplo, como o sinal representa um senoide (seno), há apenas uma frequência.

O espectro tem duas partes: uma parte negativa e uma parte positiva. Para sinais reais (sem parte imaginária), a parte negativa do espectro é sempre uma versão "espelhada" da parte positiva. Assim, neste exemplo, a parte positiva terá apenas um pico e a parte negativa terá um pico idêntico ao da parte positiva.



No espectro de frequências a frequência  $f=0$  corresponde à componente DC. Esta componente tem uma amplitude igual ao valor médio do sinal, para além dos picos de frequência das componentes sinusoidais. Assim, se o espectro de um sinal tiver um valor diferente de zero na origem, saber-se-á que o valor médio do sinal é diferente de zero.



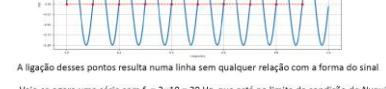
O teorema de amostragem de Nyquist-Shannon estabelece a condição que determina qual deve ser a menor frequência de amostragem que torna adequada a reconstrução do sinal:

$$f_s > 2 \times f_{max}$$

Por outras palavras, a frequência de amostragem deve ser maior que duas vezes a máxima frequência do sinal.

Consideremos, como exemplo, uma onda periódica P de um seno com  $f = 10$  Hz, ou seja, um sinal com uma frequência de dez ciclos por segundo [simplificada com  $A = 1$  e  $\varphi = 0$ ]. Uma série com  $f_s = 10$  Hz (que não respeita a condição referida anteriormente) tem uma amostra por período.

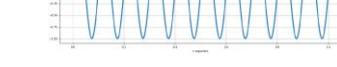
=0



A ligação destes pontos resulta numa linha sem qualquer relação com a forma do sinal.

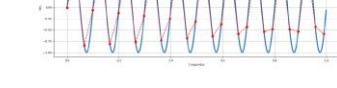
Veja-se agora uma série com  $f_s = 2 \times 10 = 20$  Hz, que está no limite da condição de Nyquist (mas ainda sem a respeitar); neste caso têm-se duas amostras por período que, conectadas, formam uma linha semelhante à anterior.

=2



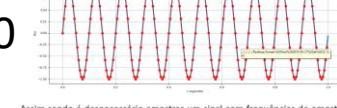
Aumentando-se a frequência de amostragem para um valor que respeite a condição, por exemplo,  $f_s = 3 \times 10 = 30$  Hz, pode-se verificar que a série adquire um comportamento que possibilita já reconstituir o sinal de uma forma clara, podendo ser quase suficiente para os objetivos a cumprir.

=3



Com um valor superior de, por exemplo,  $f_s = 20 \times 10 = 200$  Hz, tem-se uma série amostrada já bastante fiel ao sinal original.

=20



Assim sendo é desnecessário amostrar um sinal com frequências de amostragem acima das que permitem reconstituí-lo sem ambiguidade.

**Transformada Directa (DFT):**

$$F(u) = \sum_{x=0}^{N-1} [f(x) \times e^{-j2\pi ux/N}] = \sum_{x=0}^{N-1} [f(x) \times (\cos(-\frac{2\pi ux}{N}) + j \sin(-\frac{2\pi ux}{N}))]$$

**Transformada Inversa (IFT):**

$$f(x) = \frac{1}{N} \times \sum_{u=0}^{N-1} [F(u) \times e^{j2\pi ux/N}] = \frac{1}{N} \times \sum_{x=0}^{N-1} [f(x) \times (\cos(\frac{2\pi ux}{N}) + j \sin(\frac{2\pi ux}{N}))]$$

N = número de amostras que se tem

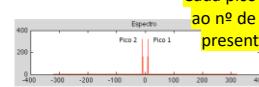
x = amostra corrente (0...N-1)

f(x) = valor do sinal em x

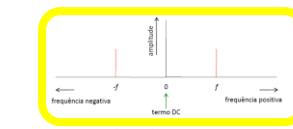
u = frequência corrente (0 Hertz até N-1 Hertz)

F(u) = quantidade de frequência u presente no sinal

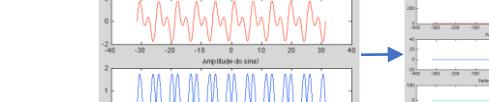
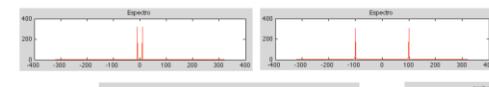
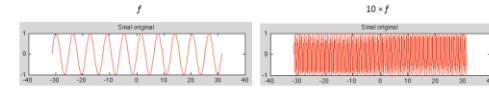
Representação centrada do espectro



Cada pico corresponde ao nº de frequências presentes no sinal



Quanto mais afastado para a esquerda (e para a direita) estiver um pico, maior é a frequência que o mesmo representa. Por outras palavras, um pico bastante afastado para a direita (e esquerda) significa que o sinal contém uma componente periódica de alta frequência.



**Transformada de Fourier:**

Numa imagem as variações são representadas pelas variações espaciais dos tons de cinzento dos pixels.

A transformação do domínio espacial para o domínio das frequências espaciais resulta numa função de valores complexos, ou seja, do tipo  $z = a+bi$ .

$$F(u) = re(F(u)) + im(F(u))j$$

*re*: parte real do número complexo  
*im*: parte imaginária do número complexo

$$\text{Magnitude} = |F(u)| = \sqrt{(re(F(u)))^2 + (im(F(u)))^2}$$

$$\text{fase} = \text{atan}\left(\frac{im(F(u))}{re(F(u))}\right)$$

A visualização do espectro pode fazer-se com a determinação da função da Magnitude.

**Transformada Discreta de Fourier (2D):** Para uma imagem de dimensões  $M \times N$ , a Transformada Directa Discreta de Fourier bidimensional é dada por:

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \times e^{-j2\pi \left( \frac{ux}{M} + \frac{vy}{N} \right)}$$

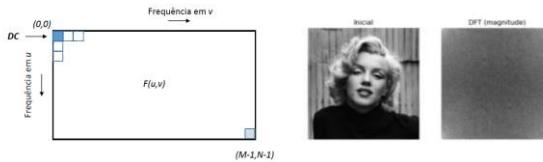
A função  $f(x,y)$  é a imagem no domínio espacial e o termo exponencial é a função-base que corresponde à representação de cada ponto  $F(u,v)$  do espaço de Fourier.

A magnitude do espectro calcula-se com a seguinte expressão:

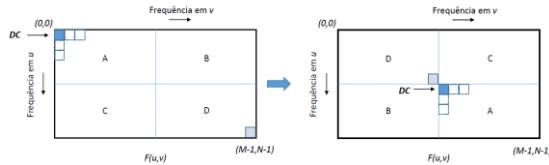
$$Mag(u, v) = \frac{1}{(N \times M)} |F(u, v)|$$

A Transformada Directa Discreta de Fourier (DDFT) corresponde à TF amostrada e, como tal, não contém todas as frequências que formam uma imagem, mas apenas um conjunto de amostras que é suficientemente grande para descrever o domínio espacial da imagem.

As funções de base são ondas seno e co-seno com frequências progressivas, isto é,  $F(0, 0)$  representa a componente DC da imagem (que tem frequência zero), correspondente à intensidade média, e  $F(M-1, N-1)$  representa a frequência mais elevada em xx e em yy.



Por motivos de facilidade de interpretação/análise do espectro, a sua disposição é geometricamente alterada por forma a que o valor de  $F(0, 0)$  se localize ao centro da janela.

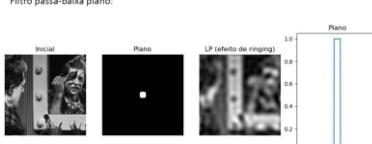
**Filtragem no domínio das frequências:**

A TF está directamente relacionada com muitas operações, entre as quais as operações de filtragem.

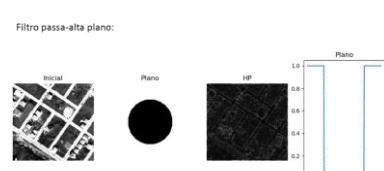
Uma operação de filtragem, que no domínio espacial resulta de uma operação de convolução entre um filtro (kernel) e a imagem de cintzento, executa-se com uma simples multiplicação no domínio das frequências (TF da imagem inicial).

Assim podem-se definir filtros passa-baixa, passa-alta e passa-banda, eliminando frequências do espectro de forma seleccionada.

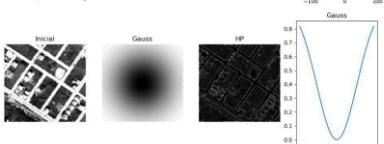
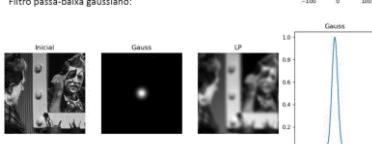
Filtro passa-baixa plano:



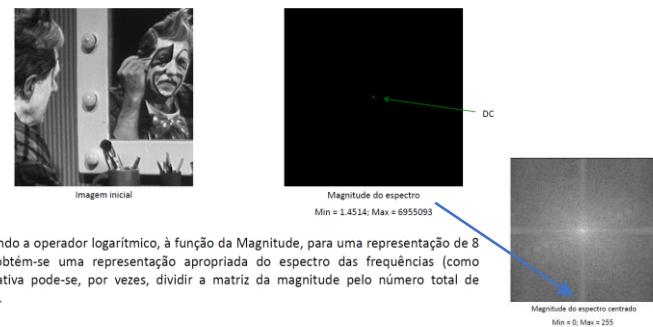
Filtro passa-alta plano:



Filtro passa-banda plano:

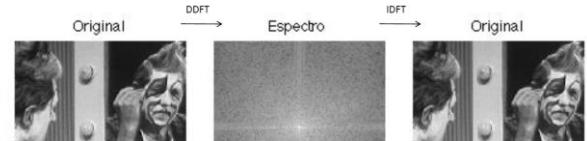


Quase sempre o valor de DC é, de longe, a maior componente do espectro de frequências. O intervalo numérico dos valores do espectro é bastante grande para ser visualizado no ecrã, o que faz com que a sua representação imediata não seja frequentemente viável.

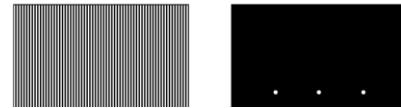


A Transformada Inversa Discreta de Fourier (IDFT) bidimensional é dada por:

$$f(x, y) = \left( \frac{1}{M \times N} \right) \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} [F(u, v) \times e^{j2\pi \left( \frac{ux}{M} + \frac{vy}{N} \right)}]$$

**Interpretação:**

O respetivo espectro (à direita) tem três frequências: o valor de DC ao centro (frequência 0) e os dois picos correspondentes à frequência das linhas na imagem original. Como, no domínio espacial, a função se altera segundo a direcção horizontal, os pontos do espectro apresentam-se alinhados ao longo de uma recta horizontal que passa pelo centro.



A distância dos pontos ao centro pode ser explicada da seguinte forma: a frequência máxima ( $k_{\max}$ ), que pode ser representada na imagem original, corresponde a 2 pixels de distância (um preto e um branco):

$$k_{\max} = \frac{1}{2}$$

Sendo o sinal constituído por faixas verticais com espessura de 2 pixels (ciclo de 4 pixels) tem-se:

$$k = \frac{1}{4} = \frac{k_{\max}}{2}$$

Portanto, os pontos do espectro situam-se a meio caminho entre o centro e os limites da imagem, ou seja, a frequência representada é igual a metade da frequência máxima.

O filtro Butterworth de ordem n e frequência de corte  $D_0$  é definido como:

$$B(u, v) = \frac{1}{1 + \left( \frac{\sqrt{u^2 + v^2}}{D_0} \right)^{2n}}$$

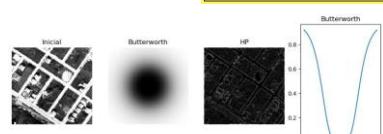
Este filtro tem a vantagem de se poder controlar a nitidez com a ordem (n).

Um filtro passa-baixa Butterworth mantém frequências dentro do raio  $D_0$  e descarta as de fora. Introduz ainda uma transição gradual de 1 para 0 para reduzir o efeito de ringing que se verifica com o filtro plano.

Filtro passa-baixa Butterworth:



Filtro passa-alta Butterworth:



**Filtro Passa-Alta = Imagem original - Filtro Passa-Baixa**

**Teorema da convolução:**

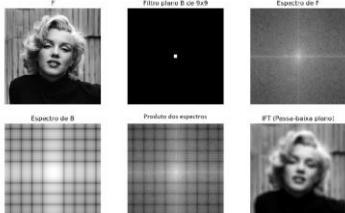
Em matemática, o **teorema da convolução** afirma que, em condições adequadas, a TF de uma convolução de dois sinais obtém-se multiplicando as suas transformadas individuais de *Fourier* e executando a inversa da TF desse produto.

$$f(x, y) * h(x, y) = idft(F(u, v) \times H(u, v))$$

Este teorema é importante porque estabelece a ligação entre operações no domínio das frequências e a ação de filtros espaciais lineares.

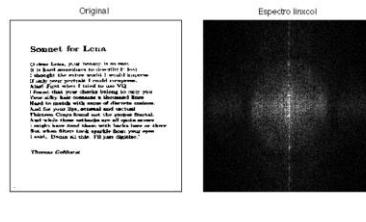
As operações de filtragem linear, feitas por convolução no domínio espacial, podem ser realizadas por multiplicações simples no domínio de *Fourier*, tornando mais rápido o processo de filtragem.

Multiplica-se matriz do filtro pretendido pela imagem resultante da DFT (espectro com ambas as partes real e imaginária) e aplica-se de seguida a IFT. Na figura pode-se ver a ilustração do processo de filtragem de uma imagem, por um filtro rectangular da média aritmética.

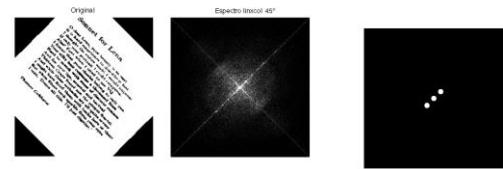
**Exemplos de Aplicação:**

**Detectação de orientação de texto:** A TF é usada para adquirir informação acerca da estrutura geométrica do domínio espacial de uma imagem. O reconhecimento de texto, usando técnicas de processamento de imagem fica simplificado se for assumido que as linhas de um dado texto estão dispostas numa dada direção.

A limiarização dos picos da imagem do espectro permite identificar a orientação da disposição das linhas de texto.

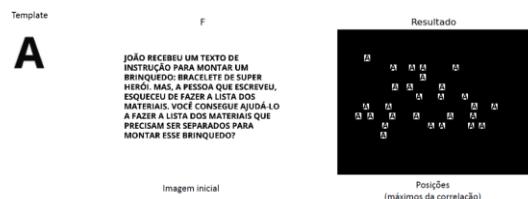


No caso de uma imagem rodada de 45º, após a identificação dos picos de frequência sobre a linha principal, a imagem pode ser rodada usando para tal o conhecimento do ângulo que se obtém nesse procedimento.



A linha perpendicular à linha principal resulta da existência de cantos escuros na imagem rodada.

A Transformada de *Fourier* pode ser usada para executar operações de localização de objectos numa imagem, por determinação do espectro de potência cruzada (que se verá no capítulo de segmentação de imagem).



O **objetivo** da morfologia matemática passa pela quantificação de estruturas do ponto de vista geométrico (com base na teoria dos conjuntos).

### Metodologia (conjuntos → transformação → medida):

- Utiliza a noção de conjunto para representar estruturas;
- Transformação dos conjuntos por forma a torná-los mensuráveis:
  - Interação do conjunto de objetos em estudo, com outro objeto com forma conhecida (**elemento estruturante**);
  - A transformação do conjunto inicial, ao longo de sucessivas operações, evidencia as suas características estruturais, que são registadas ao longo dos novos conjuntos gerados, o que implica que o **conjunto transformado é mais simples que o conjunto original**.
- Realização de medidas sobre os conjuntos transformados (superfície, perímetro, forma, etc...)

O processamento morfológico de imagem pode ser aplicado nos seguintes contextos:

- Morfologia binária** as imagens são binárias (mais frequente);
- Morfologia numérica** as imagens podem ser, ou de níveis de cinzento (monocromáticas) ou coloridas (policromáticas).

Os termos «Morfologia» e «Matemática», associados (que constituem a designação desta teoria), referem-se à utilização de conceitos de lógica de conjuntos e operações numéricas.

### Propriedades das Transformações Morfológicas

Se  $\phi$  é uma transformação morfológica então obedece a uma ou mais do que uma das seguintes propriedades:

- $\phi$  é **extensiva** se o conjunto resultante contém conjunto inicial X.

$$X \subset \phi(X)$$

- $\phi$  é **anti-extensiva** se o conjunto resultante está contido no conjunto inicial X.

$$\phi(X) \subset X$$

- $\phi$  é **crescente** se a relação de inclusão entre os conjuntos inicial e resultante se mantém.

$$Y \subset X \Rightarrow \phi(Y) \subset \phi(X)$$

- $\phi$  é **idempotente** se a sua aplicação sucessiva a X não o altera.

$$\phi(\phi(X)) \equiv X$$

- Duas transformações morfológicas  $\phi_1$  e  $\phi_2$ , são **duals** se:

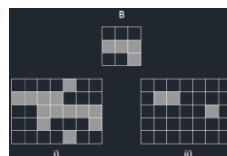
$$\phi_1(X) \equiv \phi_2(X^c)$$

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

$$E[\phi(X)] = E[X]$$

**Transformação de Vizinhança:** Uma transformação de vizinhança consiste na identificação/alteração de um pixel de uma imagem no caso de se verificar uma dada configuração de vizinhança V em redor desse pixel

- Na figura, tem-se um quadrado elementar B que percorrerá todos os pixels da imagem i).
- Em ii) estão marcadas as posições em que a configuração de vizinhança  $V(B_i)$  é idêntica à de B.
- Como transformações de vizinhança mais comuns têm-se o adelgaçamento e o espessamento.



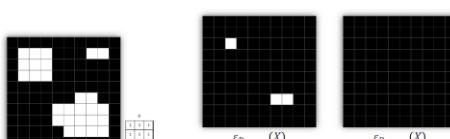
### Síntese das propriedades das transformações morfológicas elementares:

	Extensiva	Anti-extensiva	Crescente	Idempotente	Homotópica	Transformação dual
<b>E</b> Erosão	-	X	X	-	-	Dilatação
<b>D</b> Dilatação	X	-	X	-	-	Erosão
<b>Y</b> Abertura	-	X	X	X	-	Fecho
<b>Φ</b> Fecho	X	-	X	X	-	Abertura

### Transformações morfológicas elementares:

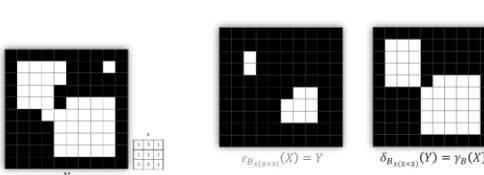
A transformação morfológica de **erosão** ( $\epsilon$ ) de um dado conjunto X, por um elemento estruturante com a sua origem em  $x$  ( $B_x$ ), define-se pela expressão seguinte:

$$\epsilon_B(X) = \{x: B_x \subset X\}$$



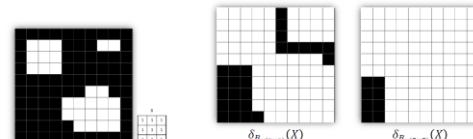
A **abertura** ( $\gamma$ ) de X consiste em executar a dilatação do resultado da erosão do conjunto X.

$$\gamma_B(X) = \delta_B(\epsilon_B(X))$$



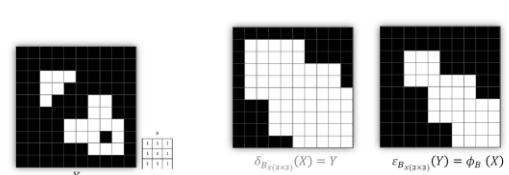
A transformação morfológica de **dilatação** ( $\delta$ ) de um dado conjunto X, por um elemento estruturante com a sua origem em  $x$  ( $B_x$ ), define-se pela expressão seguinte:

$$\delta_B(X) = \{x: B_x \cap X \neq \emptyset\}$$



O **fecho** ( $\phi$ ) de X consiste em executar a erosão do resultado da dilatação do conjunto X.

$$\phi_B(X) = \epsilon_B(\delta_B(X))$$



## Transformação tudo-ou-nada ( Hit-or-Miss transformation HMT):

A transformação "Tudo ou Nada" aplicada a X consiste numa transformação de vizinhança que recorre a um EE composto  $B = (B_1 \cdot B_2)$  com  $B_1 \cap B_2 \neq \emptyset$  e que resulta da verificação simultânea das seguintes condições  $B_1$  coincide com X e  $B_2$  coincide com  $X^c$ .

$$HMT_B(X) = \{x : (B_1)_x \subseteq X \wedge (B_2)_x \subseteq X^c\}$$

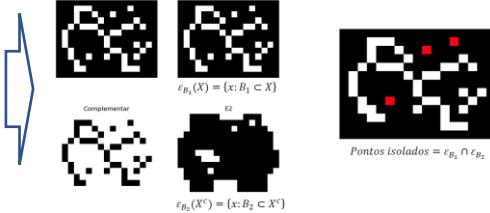
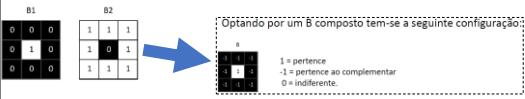
$$HMT_B(X) = \varepsilon_{B_1}(X) \cap \varepsilon_{B_2}(X^c)$$

Os índices do elemento estruturante composto são geralmente três: "1" (é domínio de  $B_1$ ), "-1" (é domínio de  $B_2$ ) e "0" (é indiferente).

A HMT é geralmente usada para encontrar configurações específicas em grupos de pixels ou objetos e determina-se pela intersecção entre as erosões de X por  $B_1$  e de  $X^c$  por  $B_2$  (como se verá em alguns exemplos a seguir):

**Pontos isolados** pixels sem quaisquer outros pixels na sua vizinhança.

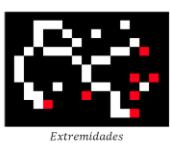
Desenham-se duas configurações (EE)  $B_1$  e  $B_2$ , tal que,  $B_1 \subseteq X$  e  $B_2 \subseteq X^c$ .



**Pontos extremos** pixels com um pixel no máximo na sua vizinhança próxima).

Com  $B_1$  e  $B_2$ :

B1[90]	B2[90]	B1[45]	B2[45]
1 1 1 1 1 1 1 1 1			
-1 -1 -1 -1 -1 -1 -1 -1 -1			
0 0 0 0 0 0 0 0 0			

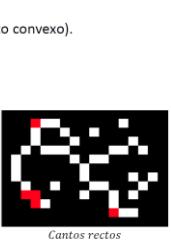


Com  $B_1$  e  $B_2$ :

B1[90]	B2[90]	B1[45]	B2[45]
1 1 1 1 1 1 1 1 1			
-1 -1 -1 -1 -1 -1 -1 -1 -1			
0 0 0 0 0 0 0 0 0			

Optando por B composto:

B1[90]	B2[90]	B1[45]	B2[45]
1 1 0 1 1 0 1 1 0			
-1 1 1 -1 1 1 -1 1 1			
0 0 0 0 0 0 0 0 0			



Optando por B composto:

Com  $B_1$  e  $B_2$ :

B1	B2
1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1
-1 -1 -1 -1 -1 -1 -1 -1 -1	-1 -1 -1 -1 -1 -1 -1 -1 -1

Optando por B composto:

B
0 1 0 1 0 1 0 1 0
1 1 1 1 1 1 1 1 1



Optando por B composto:

## Transformação de espessamento (THICK):

Espessamento de um conjunto X consiste numa transformação de vizinhança que adiciona a X todos pontos que correspondam a uma determinada configuração de vizinhança  $V(Bx)$

$$THICK(X, B) = X \cup HMT(X, B)$$

É aplicada apenas a imagens binárias e produz uma imagem binária como resultado.

De forma geral, a operação de adelgaçamento é determinada por translação do elemento estruturante B, por todos os pixels da imagem, comparando em cada um a sua configuração de vizinhança com a configuração dos correspondentes pixels na imagem

Havendo uma coincidência entre ambas as configurações, então no pixel correspondente à posição do centro de B é atribuído o valor 0 caso contrário mantém se inalterável.

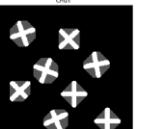
A dilatação e o fecho morfológicos são exemplos de transformações de espessamento.

## Envelope (convex hull):

B1[90]	B2[90]
1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1
-1 -1 -1 -1 -1 -1 -1 -1 -1	-1 -1 -1 -1 -1 -1 -1 -1 -1



B1[90]	B2[90]
1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1
-1 -1 -1 -1 -1 -1 -1 -1 -1	-1 -1 -1 -1 -1 -1 -1 -1 -1



- Envelope convexo determina-se por execução da transformação HMT, para determinar concavidades nos objetos e consequente preenchimento.
- A operação é iterativa e continuará até atingir a idempotência.

## Esqueleto por zonas de influência (SKIZZ - Skeleton by Influence Zones):

- SKIZZ é uma estrutura de esqueleto que divide uma imagem em regiões, cada qual contendo um objeto distinto da imagem.
- As fronteiras são definidas por forma a que todos os pontos interiores a cada área estejam mais próximos do correspondente objeto interior a essa área É por vezes designada por Diagrama de Voronoi. A operação é iterativa e continuará até atingir a idempotência
- O SKIZZ pode ser obtido por um processo métrico, calculando distâncias euclidianas, ou por processos morfológicos, envolvendo dilatações com elementos estruturantes de diferentes tamanhos.



## Transformação de adelgaçamento (THIN):

Adelgaçamento de um conjunto X consiste numa transformação de vizinhança que retira a X todos pontos que correspondam a uma dada configuração de vizinhança  $V(Bx)$ .

$$THIN(X, B) = X \cap NOT[HMT(X, B)]$$

É aplicada apenas a imagens binárias e produz uma imagem binária como resultado.

De forma geral, a operação de adelgaçamento é determinada por translação do elemento estruturante B, por todos os pixels da imagem, comparando em cada um a sua configuração de vizinhança com a configuração dos correspondentes pixels na imagem.

Havendo uma coincidência entre ambas as configurações, então no pixel correspondente à posição do centro de B é atribuído o valor 0 caso contrário mantém se inalterável.

A erosão e a abertura morfológicas são exemplos de transformações de adelgaçamento.

## Esqueletização

é um exemplo de adelgaçamento morfológico que reduz as regiões de uma imagem binária a uma estrutura mínima que preserve a extensão e conectividade das regiões originais.

É uma transformação usada frequentemente para "estreitar" resultados de detecção de fronteiras, reduzindo a espessura das linhas a outras com apenas um pixel de espessura. Note-se que o esqueleto resultante de um conjunto conexo é igualmente um conjunto conexo.

O esqueleto de um conjunto pode ser determinado de diversas formas:

- Localização dos centros de circunferências máximas bi-tangentes aos limites da região considerada



- Fórmula de Lantuejoul: para uma imagem binária discreta X, o esqueleto  $S(X)$  é a união de todos os subconjuntos  $S_k(X)$  com o elemento estruturante B de dimensão k.

$$S(X) = \bigcup S_k(X)$$

$$S_k(X) = r_{k+1}(X) \cap NOT[r_k(r_k(X))]$$



## Escanhoamento

(prune) é também uma operação de adelgaçamento que visa suprimir sucessivamente os pontos extremos de um conjunto binário, até se verificar a condição de idempotência.



**Transformações geodésicas binárias:**

As transformações geodésicas binárias são transformações morfológicas sobre uma imagem binária Y, condicionadas por uma determinada geodesia binária X. Destacam-se as seguintes:

1 - Dilatação geodésica;

2 - Erosão geodésica;

3 - Reconstrução geodésica binária por dilatações geodésicas sucessivas;

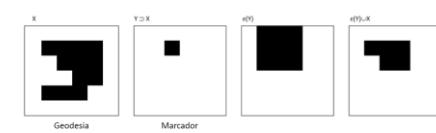
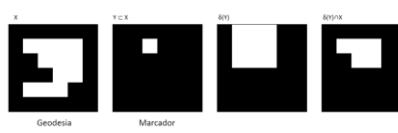
4 - Reconstrução geodésica binária por erosões geodésicas sucessivas.

**Dilatação geodésica:** dilatação morfológica de um conjunto Y condicionada à geodesia X.

$$\delta_X(Y) = \delta(Y) \cap X$$

**Erosão geodésica:** erosão morfológica de um conjunto Y condicionada à geodesia X.

$$\epsilon_X(Y) = \epsilon(Y) \cup X$$



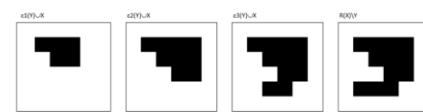
**Reconstrução geodésica binária por dilatações geodésicas sucessivas:**

$$R_X(Y) = \delta_X^n(Y) = \lim_{n \rightarrow \infty} (\delta_X \circ \dots \circ \delta_X)(Y) = \delta_X^{\lfloor n \rfloor - 1} \left( \delta_X^{n-2} \left( \dots \left( \delta_X^1(Y) \right) \right) \right)$$



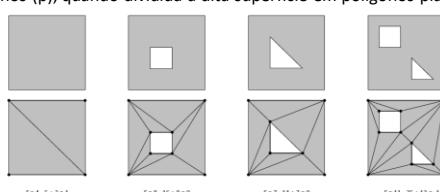
**Reconstrução geodésica binária por erosões geodésicas sucessivas:**

$$R_X(Y) = \epsilon_X^n(Y) = \lim_{n \rightarrow \infty} (\epsilon_X \circ \dots \circ \epsilon_X)(Y) = \epsilon_X^{\lfloor n \rfloor - 1} \left( \epsilon_X^{n-2} \left( \dots \left( \epsilon_X^1(Y) \right) \right) \right)$$

**Número de conectividade (número de Euler):**

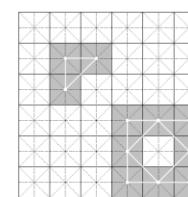
De forma geral, o número de conectividade de uma superfície ou conjunto, ou também chamado número de Euler (E) é igual ao número de vértices (v), menos o número de arestas (a), mais o número de polígonos (p), quando dividida a dita superfície em polígonos planos definidos pelas arestas e pelos vértices.

$$E = v - a + p$$



No domínio de representação de uma imagem digital, o valor de E determina-se a partir do grafo de representação dos pixels. A seguir exemplifica-se a determinação de E para o objecto binário sombreado.

$$E = v - a + p$$



(Pixels == vértices)

Número de pixels = 11

Número de arestas do grafo = 15

Número de triângulos = 5

$$E = 11 - 15 + 5 = -1$$

Número de objectos = 2

Número de buracos = 1

$$E = 2 - 1 = 1$$

**As operações elementares da morfologia matemática binária podem ser estendidas para as imagens de escala de cinzentos através do uso das operações de mínimo e máximo que estabelecem um paralelo com as operações de erosão e dilatação binárias.**

Estas operações atribuem a cada pixel da imagem novos valores correspondentes ao mínimo ou ao máximo valor de uma dada vizinhança em torno desse pixel. A vizinhança fica definida de acordo com a forma do elemento estruturante. As intensidades dos pixels das imagens numéricas são consideradas como elevações topográficas.

A morfologia matemática numérica tem aplicação em processos de contraste de imagem, descrição de texturas, detecção de fronteiras e limiarização entre outras.

### Transformações morfológicas elementares:

A operação de **erosão**  $\varepsilon$  de uma dada função  $f$ , por um elemento estruturante  $B$ , posicionado com a sua origem em  $x$  ( $B_x$ ), é dada pela expressão:

$$\varepsilon_B(f(x)) = \min_{B \in B} f(x + b)$$



B

imagem

f

epsilon\_B(f)

A operação de **dilatação**  $\delta$  de uma dada função  $f$ , por um elemento estruturante  $B$ , posicionado com a sua origem em  $x$  ( $B_x$ ), é dada pela expressão:

$$\delta_B(f(x)) = \max_{B \in B} f(x + b)$$



B

imagem

f

delta\_B(f)

A operação de **abertura**  $\gamma$  de uma dada função  $f$ , por um elemento estruturante  $B$ , posicionado com a sua origem em  $x$  ( $B_x$ ), é dada pela expressão:

$$\gamma_B(f(x)) = \delta_B(\varepsilon_B(f(x)))$$



B

imagem

f

epsilon\_B(f)

gamma\_B(f)

A operação de **fecho**  $\phi$  de uma dada função  $f$ , por um elemento estruturante  $B$ , posicionado com a sua origem em  $x$  ( $B_x$ ), é dada pela expressão:

$$\phi_B(f(x)) = \varepsilon_B(\delta_B(f(x)))$$



B

imagem

f

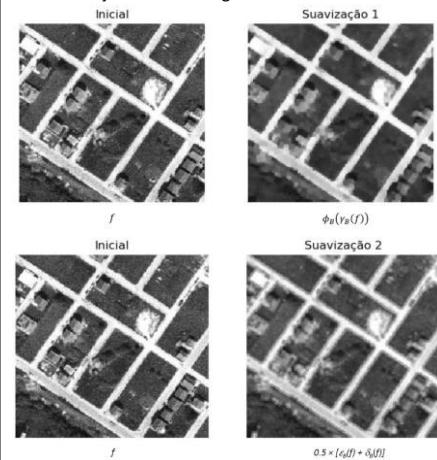
delta\_B(f)

phi\_B(f)

### Suavização morfológica:

A suavização de uma imagem por métodos morfológicos pode ser obtida com várias abordagens:

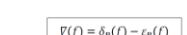
- Uma consiste em executar uma operação de abertura seguida de um fecho. Desta forma removem-se artefactos claros e escuros com tamanho igual ou abaixo do tamanho do EE.
- Uma segunda abordagem de suavização consiste em executar a média entre as operações de erosão e dilatação de uma imagem.



### Gradiente morfológico:

O gradiente morfológico  $\nabla$  (também designado por gradiente de Beucher) determina, em cada pixel, pela diferença algébrica entre a dilatação  $\delta(f)$  e a erosão  $\varepsilon(f)$ :

$$\nabla(f) = \delta_B(f) - \varepsilon_B(f)$$



B

imagem

f

epsilon\_B(f)

delta\_B(f)

n(f)

- A assunção comum relativa ao gradiente é a de que as fronteiras dos objetos ou arestas estão localizadas onde se verificam diferenças elevadas entre os valores de pixels vizinhos.
- Os operadores de gradiente são usados para evidenciar essas variações.
- A espessura de uma aresta detetada por um gradiente morfológico é igual a dois pixels (considerando um elemento estruturante elementar, ou seja,  $3 \times 3$ ): um pixel em cada lado da fronteira).
- Havendo ruído aleatório, a imagem deve ser filtrada antes de aplicar o operador de gradiente, para evitar realçar também o ruído.
- Ao contrário dos operadores de gradiente lineares de Sobel Prewitt ou Roberts os gradientes morfológicos obtidos com elementos estruturantes simétricos tendem a depender menos da direcionalidade das arestas dos objetos.

- Os Semi-gradientes podem ser usados para detetar os limites interno ou externo de uma fronteira. Os semi-gradientes têm só um pixel de espessura.

O semi-gradiente por erosão, ou gradiente interno  $\nabla^+$ , define-se como sendo a diferença entre a imagem original e a sua erosão numérica:

$$\nabla^+(f) = f - \varepsilon_B(f)$$

O gradiente interno evidencia:

- as fronteiras internas dos objectos que são mais claros que o fundo.
- as fronteiras externas dos objectos que são mais escuros que o fundo.

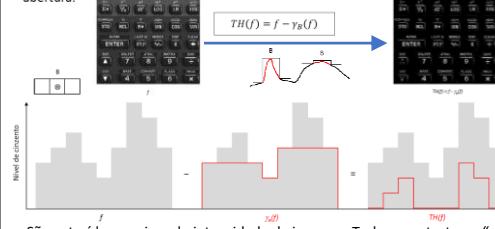
O gradiente externo evidencia:

- as fronteiras internas dos objectos que são mais escuros que o fundo.
- as fronteiras externas dos objectos que são mais claros que o fundo.

$$\nabla^+(f) = \delta_B(f) - f$$

### Transformação Chapéu-Alto (TH Top-Hat):

Chapéu-Alto (Top-Hat): consiste na diferença algébrica entre a função (imagem) e a sua abertura.



### Transformação Chapéu-Alto Auto complementar:

A soma do "Top Hat" com o "Bottom Hat" extrai todos os objectos da imagem que não estão contidas no elemento estruturante dado, qualquer que sejam os seus contrastes relativos (ou seja, picos ou vales).

A partir das respectivas expressões deduz-se facilmente que,

$$TH(f) + BH(f) = f - \gamma_B(f) + \phi_B(f) - f = \phi_B(f) - \gamma_B(f)$$

Ou seja, o Top-Hat auto-complementar é dado pela **diferença entre o fecho e a abertura**.

### Transformação Chapéu-Alto e contraste de imagem:

Um simples operador morfológico de contraste pode ser obtido com a adição, à imagem inicial, da diferença de ambos os operadores de Top-Hat e Bottom-Hat. Este procedimento consiste em começar por realçar os objectos mais claros (adição do Top-Hat), seguido do realce dos objectos mais escuros (subtração do Bottom-Hat).

$$f + TH(f) - BH(f) = f + f - \gamma_B(f) - \phi_B(f) + f$$

Chapéu-Baixo (Bottom-Hat): consiste na diferença algébrica entre o fecho da imagem e a imagem inicial.



B

imagem

f

phi\_B(f)

BH(f)

São extraídos os "vales" de intensidade da imagem. Todas as estruturas "em vale" que não estão contidas no elemento estruturante são removidas da imagem.

Os valores resultantes que ficam de fora do intervalo dinâmico da imagem inicial ( $z_{\min} < z_{\max}$ ) terão os valores  $z_{\min}$  ou  $z_{\max}$ , consoante ficarem abaixo ou acima dos extremos do intervalo ( $z_{\min} = 0$  e  $z_{\max} = 255$ , para imagens de 8 bits).

**Filtragem sequencial alternada:**

A filtragem de uma imagem que contenha ruídos escuro e claro pode ser obtida através da aplicação de uma sequência de operações de fecho-abertura ou abertura-fecho.

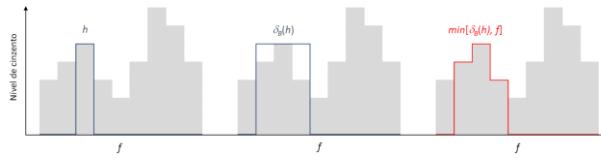
Uma só operação sequencial de fecho-abertura, ou abertura-fecho, com um elemento estruturante grande não produz resultados aceitáveis.

Uma solução para este problema consiste em aplicar fechos e aberturas alternados, começando com elementos estruturantes de baixas dimensões e ir aumentando progressivamente o seu tamanho até um dado tamanho final.

A este processo de filtragem por aplicação sequencial de fecho-abertura, ou abertura-fecho, dá-se o nome de **filtragem sequencial alternada**.

**Dilatação geodésica:** consiste em determinar o valor mínimo entre a dilatação da imagem marcadora  $h$  ( $\leq f$ ) e a função  $f$ .

$$\delta_f(h) = \min[\delta_B(h(x,y)); f(x,y)]$$

**Reconstrução geodésica por dilatações geodésicas sucessivas.**

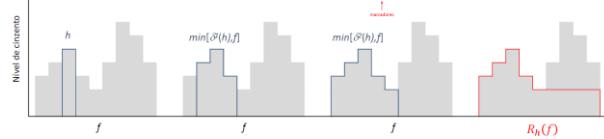
$$R_h(f) = \delta_f^{(0)}(h)$$

, até se verificar a condição

$$\delta_f^{(i+1)}(h) = \delta_f^{(i)}(h)$$

Início

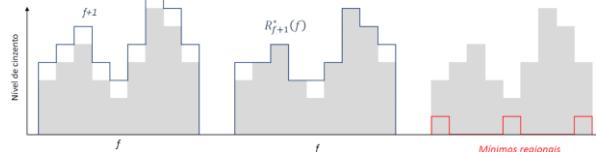
Reconstrução

**Extremos regionais:**

Uma imagem de mínimos/máximos regionais é, geralmente, uma imagem binária.

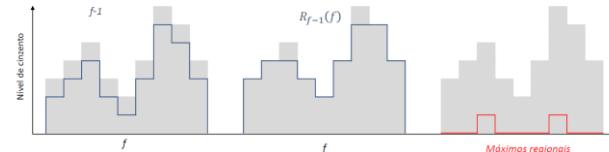
Um conjunto binário  $M_h$  de uma imagem numérica  $f$  é um **mínimo regional** de elevação  $h$  se e só se  $M_h$  for uma superfície conexa de igual altitude  $h$ , a partir da qual seja impossível alcançar um ponto de elevação inferior sem ter que antes ascender na função.

$$k_{\min(f)} = R_{f+1}^*(f) - f$$



O mesmo se aplica ao conceito de **máximo regional** de elevação  $h$ , sendo que se tem que descer na função antes de alcançar outro máximo regional.

$$k_{\max(f)} = f - R_{f-1}(f)$$

**Transformação watershed:**

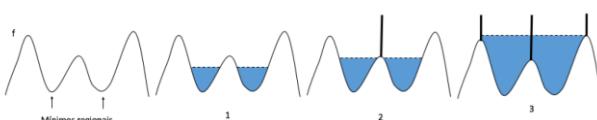
Transformação que visa segmentar uma **imagem de cintos** em regiões distintas, a partir de uma **imagem binária de marcadores**:

- Imagen 1 de entrada:** imagem de cintos (superfície topográfica definida pelos valores de cintos dos pixels).
- Imagen 2 de entrada:** imagem de marcadores (pixels isolados ou conjuntos de pixels).
- Imagen de saída:** Imagem binária, ou do tipo *label*, com as bacias de escoamento resultantes da transformação (ou, em alternativa, as linhas de *watershed*).

A transformação resulta da execução de um processo análogo ao processo físico de imersão vertical da superfície em água, com velocidade constante, a partir de "furos" feitos nessa superfície (imagem binária de marcadores).

Por analogia com um processo de imersão, em água, de uma superfície com furos nos seus mínimos regionais, em certa altura dois ou mais vales inundados acabarão por se fundir. Para evitar que tal suceda, erguem-se barreiras em todos os pontos da superfície onde a fusão acontece.

No final do processo, apenas as barreiras são representadas. Estas barreiras, as linhas de *watershed* da superfície, separam as bacias de escoamento umas das outras.

**Transformações geodésicas numéricas:**

**Transformações geodésicas numéricas** são transformações morfológicas aplicadas a uma imagem numérica  $h$ , mas condicionadas pela morfologia de uma outra imagem numérica  $f$ . Destacam-se as seguintes:

1 - Dilatação geodésica

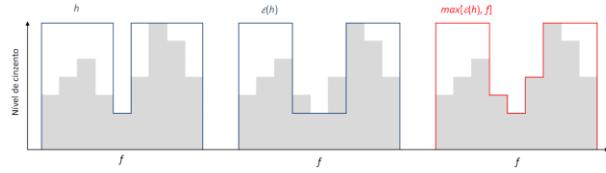
2 - Erosão geodésica

3 - Reconstrução geodésica numérica por dilatações geodésicas sucessivas.

4 - Reconstrução geodésica numérica por erosões geodésicas sucessivas (reconstrução dual)

**Erosão geodésica:** consiste em determinar o máximo entre a erosão da imagem marcadora  $h$  ( $\geq f$ ) e a função  $f$ .

$$\epsilon_f(h) = \max[\epsilon_B(h(x,y)); f(x,y)]$$



**Reconstrução geodésica numérica por erosões geodésicas sucessivas (Reconstrução Dual).**

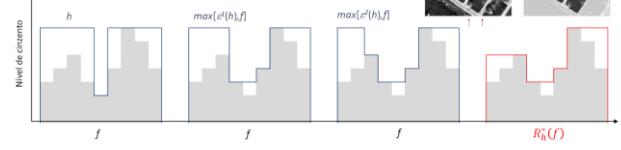
$$R_h^*(f) = \epsilon_f^{(0)}(h)$$

, até se verificar a condição

$$\epsilon_f^{(i+1)}(h) = \epsilon_f^{(i)}(h)$$

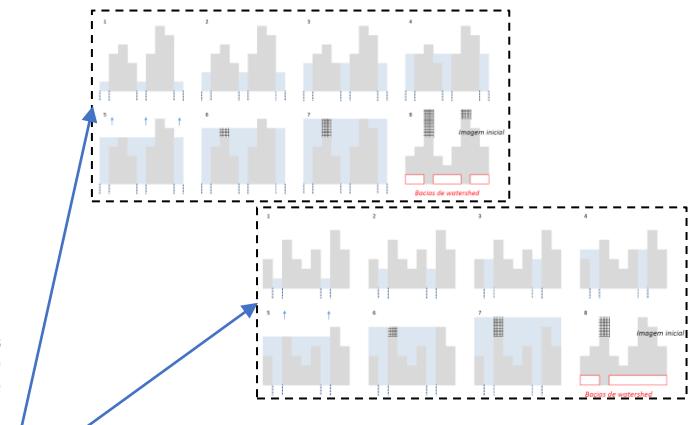
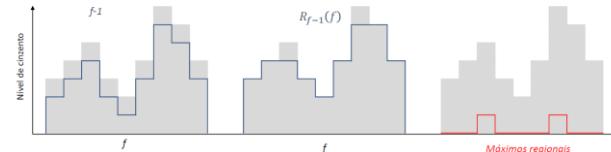
Início

Reconstrução Dual

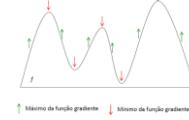


O mesmo se aplica ao conceito de **máximo regional** de elevação  $h$ , sendo que se tem que descer na função antes de alcançar outro máximo regional.

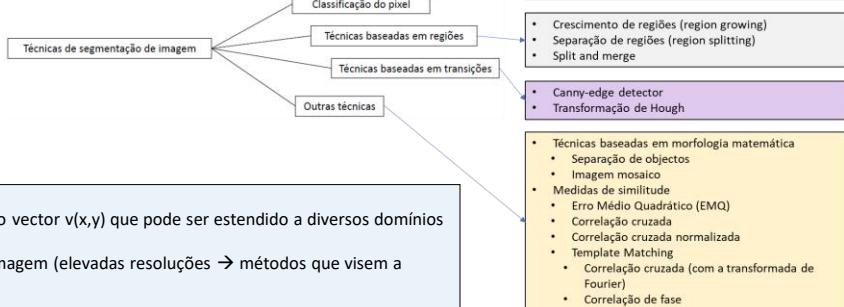
$$k_{\max(f)} = f - R_{f-1}(f)$$



**Watershed da imagem do gradiente:** em vez da imagem original  $f$ , aplica-se sobre a imagem de cintos do gradiente morfológico  $\nabla$  (ou outro), a partir dos respectivos mínimos regionais.



- A segmentação de imagem (seg)** refere-se à decomposição de uma imagem digital em uma série de regiões distintas (conjuntos de pixels).
- O objetivo da SEG é o de simplificar, ou alterar, a representação inicial de uma imagem, para outra representação que tenha maior significado e, consequentemente, que seja mais simples de analisar.
- A SEG é tipicamente usada para localizar objetos e limites de objetos nas imagens.
- Todos os pixels de uma região segmentada são semelhantes no que respeita a determinadas características, ou propriedades computacionais, tais como a cor, a intensidade, ou a textura.
- Regiões adjacentes são significativamente distintas quanto às mesmas características.
- A SEG reduz a informação dos pixels a informação contextualizada a regiões.



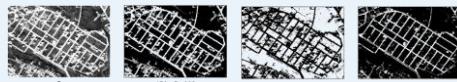
### Classificação do pixel

- A atribuição de uma classe a um pixel é guiada pela definição de um certo vector  $v(x,y)$  que pode ser estendido a diversos domínios (espectral, temporal), melhorando-se assim a segmentação.
- O volume de cálculo é proporcional ao número de pixels que contém a imagem (elevadas resoluções → métodos que visem a extração de objetos por inteiro)

### Threshold

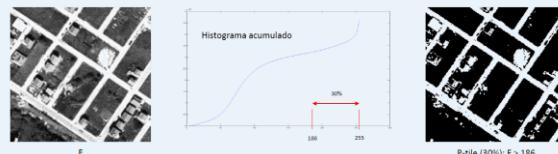
**Limiarização (threshold) simples do histograma:** é uma das formas de segmentação mais usuais, associada à técnica de classificação do pixel.

- É útil na discriminação entre objectos de primeiro plano (*foreground*) e regiões de contexto (*background*).
- Mediante a seleção de um certo valor de *threshold*  $t$ , a imagem de intensidade é convertida em imagem binária.
- Estimam-se valores para  $t$  por observação e segmentação directa do histograma, ou por metodologias de segmentação do histograma com diferentes graus de automatização.

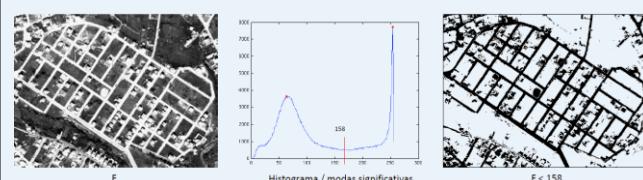


**Threshold pelo método "p-tile":** Aplica-se quando se conhece a fração aproximada que um certo objecto ocupa na imagem inteira.

Determina-se o valor de  $t$ , encontrando o valor de intensidade que corresponda à fração do número de pixels pretendida. O valor de  $t$  é facilmente extraído a partir do histograma acumulado  $A$ , estimando  $A(t) = 1/p$ .



**Threshold por pesquisa das modas:** determinação dos picos de frequência e dos mínimos locais entre os picos (histogramas multimodais).



**Threshold pelo método de Otsu:** Este método é usado para executar a limiarização automática de uma imagem de cintezos em duas classes, por intermédio de uma análise estatística optimizada.

O valor óptimo de limiarização que separa as duas classes resulta da minimização de ambas as variâncias ponderadas de entre todos os *thresholds* do histograma possíveis.

$$\sigma_w^2 = w_r(t_1) \times \sigma_r^2(t_1) + w_g(t_1) \times \sigma_g^2(t_1)$$

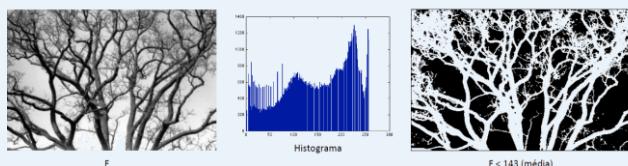
$$\sigma^2 = \min(\sigma_w^2)$$

$$w_r(t) = \sum_{i=1}^{t-1} p_i \quad w_g(t) = \sum_{i=t}^{255} p_i$$

**Threshold multi-níveis (quantização):** Algumas imagens podem conter objectos de diferentes intensidades. Para tal podem-se usar limiarizações em vários níveis de intensidade para separar esses objectos.

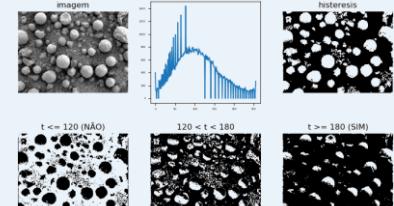


**Threshold pela média:** calcula-se o valor médio da imagem e usa-se como valor para  $t$ . Aplica-se este método quando a imagem tem aproximadamente a mesma quantidade de pixels nos motivos de *foreground* e *background*.

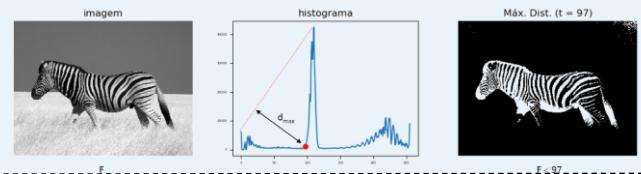


**Threshold por histerese:** se não existe um "vale" bem definido no histograma, significa que há pixels do *background* que têm valores de cinzento semelhantes aos pixels do *foreground* e vice-versa. Como tal efectuam-se duas limiarizações  $t_1$  e  $t_2$  ( $t_2 > t_1$ ) seguindo os seguintes princípios:

- Pixels com valores  $< t_1$  pertencem ao *background*;
- Pixels com valores  $> t_2$  pertencem ao objecto;
- Pixels com valores dentro do intervalo  $[t_1, t_2]$  pertencem ao objecto se forem adjacentes aos pixels do objecto.



**Threshold pela máxima distância (d):** determinação do nível de cinzento que se encontra à maior distância da recta que une o primeiro índice de frequência diferente de zero e o pico significativo do histograma.



**Threshold adaptado (adaptive threshold):** É uma técnica de segmentação do pixel um pouco mais refinada, aplicada em circunstâncias de variação de luminosidade ao longo da imagem. Assim, procede-se da seguinte forma:

- Aplicar filtro passa-baixa (média) com filtro de dimensões suficientemente grandes (por exemplo 11x11) para "esbater" a informação a segmentar;
- Determinar o valor absoluto da subtração da imagem filtrada à imagem original, seguido da subtração de uma constante K (i.e. 10) ao resultado;
- Limiarização simples do resultado anterior.



As operações de limiarização apresentam geralmente as seguintes:

#### Vantagens:

- Facilidade de implementação.
- Rapidez (especialmente se repetido em imagens similares).
- Adequado para certos tipos de imagens (por exemplo, documentos de texto).

#### Desvantagens:

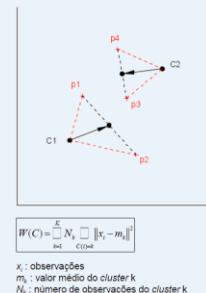
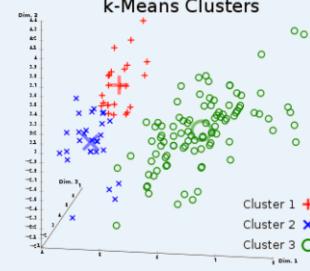
- A segmentação é por vezes pouco eficaz, pois não garante a coerência dos objectos (por exemplo, a existência de vazios, pixels "estranhos", etc.).

**Pixel Clustering:**

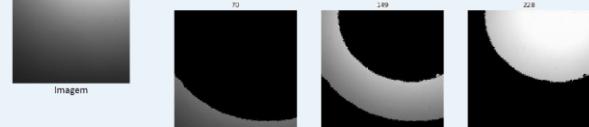
**Clustering:** Consiste numa técnica de segmentação que visa organizar/agrupar os pixels em grupos (*clusters*) com base nos seus atributos de intensidade radiométrica.

- Definir centroides iniciais para os clusters.
- Determinar os pixels que estão mais próximos de cada centroide dos *clusters*.
- Cada centroide é re-determinado a partir de uma medida da respectivo *cluster* (por exemplo, a distância euclidiana).
- O processo é iterativo e termina quando os centroides não sofrerem alterações.

**K-Means clustering:** técnica que visa agrupar os pixels em *clusters*, pela sua média.

**K-Means Clusters**

**Gray K-Means clustering:** agregação dos pixels de uma imagem de cinzentos em *clusters*.



**Color K-Means clustering:** agregação dos pixels de uma imagem de cor.

**Histograma multivariado (HMV)**

Fernando Soares

O objetivo da análise multivariada é estabelecer e utilizar as correlações espaciais das intensidades dos pixels para reduzir a informação dos dados a um nível mais refinado, com vista a uma interpretação mais fácil.

Como a análise multivariada investiga a correlação espacial de intensidades entre imagens, se houver poucas ou nenhuma informações complementares entre elas, é improvável que a técnica produza informação relevante adicional.

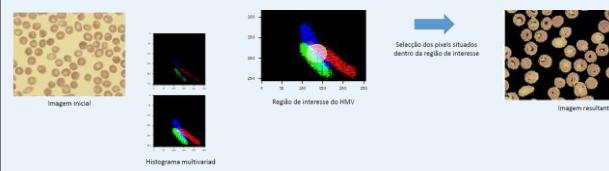
Neste contexto, a análise multivariada do histograma deve ser vista como uma ferramenta complementar dos meios convencionais de análise de imagem.

As dimensões do HMV de uma imagem RGB ( $2^n \times 2^n \times 3$  bits) são de ( $n \times n \times 3$ ). O cálculo do HMV de uma imagem RGB faz-se da seguinte forma ( $p, q = 1 \dots n$ ):

$$HMV[p, q, 0] = HMV[p, q, 0] + 1 \quad \begin{cases} p = \text{rgb}[i, j, 0] \\ q = \text{rgb}[i, j, 1] \end{cases} \quad \text{Bandas 1 e 2}$$

$$HMV[p, q, 1] = HMV[p, q, 1] + 1 \quad \begin{cases} p = \text{rgb}[i, j, 0] \\ q = \text{rgb}[i, j, 2] \end{cases} \quad \text{Bandas 1 e 3}$$

$$HMV[p, q, 2] = HMV[p, q, 2] + 1 \quad \begin{cases} p = \text{rgb}[i, j, 1] \\ q = \text{rgb}[i, j, 2] \end{cases} \quad \text{Bandas 2 e 3}$$

**Técnicas baseadas em regiões**

- Estas técnicas de SEG visam a partição, ou a agregação de regiões, de acordo com características associadas aos pixels da imagem (intensidade; textura, ou padrão dos objetos; perfis espectrais que fornecem dados multidimensionais)
- Com base nessas características, e de acordo com critérios de similaridade entre os pixels da imagem, constrói-se progressivamente, com algoritmos iterativos, as regiões homogéneas e termina-se o processo iterativo assim que todos os pixels da imagem forem classificados segundo aquelas as características.
- A fronteira entre cada duas diferentes regiões define-se onde ambas se encontram.
- Na caracterização final das regiões  $R_i$  de uma imagem / dever-se-á respeitar a condição de independência espacial para todas elas.

$$I = \bigcup_j^S R_j \quad \wedge \quad R_i \cap R_j = \emptyset \quad , \quad i \neq j$$

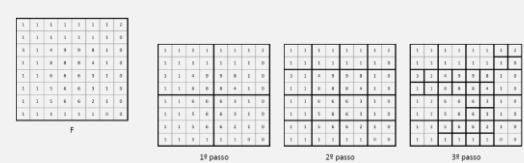
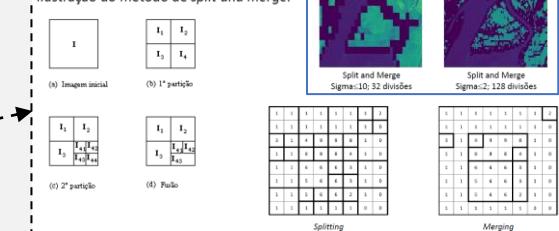
**Separação de regiões (region splitting):** separação da imagem, ou de uma ou mais áreas de interesse de uma imagem, num conjunto de regiões disjuntas, por forma a que haja coerência entre os pixels de cada região, quanto a certas propriedades.

Exemplos de técnicas:

I. A transformação morfológica *watershed* é frequentemente referenciada como técnica usada na definição preliminar das regiões.

II. Área total dividida em regiões disjuntas se os correspondentes pixels não respeitarem um certo constrangimento de similaridade. Por norma, a partição inicial faz-se em quatro quadrantes. O processo continua em cada nova área de interesse, até que não aconteça mais nenhuma partição.

III. Sucessiva divisão em quatro quadrados iguais; uma região mantém-se se cumprir o critério de homogeneidade (diferença  $\leq 1$ ); termina quando não houver mais divisões.

**Split and merge:** Ilustração do método de split and merge.

**Split and merge:** Combinção entre as técnicas de *region splitting* e *region growing*.

- Após a partição por *region splitting*, executa-se um processo de fusão das sub-regiões que cumpram os critérios de homogeneidade estabelecidos.
- O método de SEG, conhecido por *quadtree decomposition*, é um exemplo da técnica *split-and-merge*, que consiste em particionar a imagem em quadrados idênticos (imagens com área de representação quadrada e com dimensão  $2^k$ ) e depois anexar as regiões que tenham propriedades radiométricas semelhantes.

## Técnicas baseadas em transições:

**Transições** são descontinuidades na intensidade de uma dada imagem, e podem ser dos tipos:

- i) "Degrau": onde a mudança se faz de forma abrupta entre dois valores distintos. 
  - ii) "Linha": onde a mudança acontece, regressando no entanto ao valor anterior após uma reduzida distância.

Contudo, estas mudanças são efectivamente raras de acontecer, verificando-se quase sempre os modelos de "rampa" e "telhado".

Estas técnicas assentam na determinação de linhas de transição numa imagem, por uso de operadores de detecção de zonas de transição (as transições correspondem a zonas de descontinuidade radiométrica, textural, etc.).

- A variação radiométrica nos pixels da fronteira entre duas regiões distintas adjacentes é bem mais significativa que a variação entre os pixels internos a cada uma dessas regiões.
  - No entanto os resultados de segmentação por detecção de transições devem passar por fases suplementares de processamento, no sentido de adequar as linhas encontradas aos objectos específicos que se pretendem ver delimitados.
  - Executada uma das técnicas de segmentação, crescimento de regiões, ou detecção de transições, torna-se intuitivo definir a outra, já que a construção de regiões torna-se acessível a partir do conhecimento das respectivas transições, assim como também é acessível definir as transições a partir do conhecimento das regiões.
  - Pode acontecer que as regiões de uma imagem, obtidas com metodologias de crescimento de regiões, não coincidam com as que se obtêm por metodologias de detecção de transições, ou vice-versa.

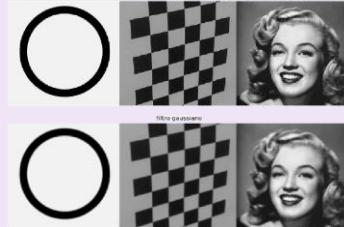
### Canny-edge detector:

**Canny-edge detector:** é um algoritmo desenhado para ser um detector de fronteiras

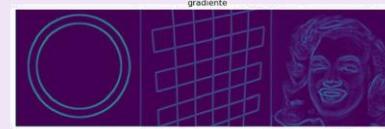
Parte-se de um imagem de cinzentos e obtém-se uma imagem binária com as posições das descontinuidades radiométricas marcantes da imagem inicial.

E executado em diversas etapas que focam conceitos diversos de processamento de imagem, tais como filtragem, quantização, e gradiente.

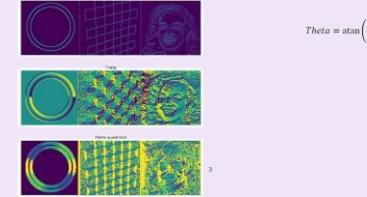
1. A imagem inicial é suavizada com um filtro gaussiano



2. Executa-se um filtro derivativo simples (por exemplo o de Sobel) sobre a imagem suavizada, para realçar as primeiras derivadas espaciais mais elevadas nas direções horizontal e vertical.



3. Determina-se, em cada pixel, o ângulo de inclinação entre as direcções horizontal e vertical do gradiente, e quantizam-se os valores para as direcções de  $kx45^\circ$  ( $k=0...3$ ).



4. Executa-se, ao longo das linhas de junção do gradiente, uma busca dos pixels máximos nas quatro direcções de (3) e coloca-se a zero todos os pixels que não pertencem à linha de topo, resultando apenas uma linha (processo conhecido por supressão de pixels não-máximos (*non-maximal suppression*)).



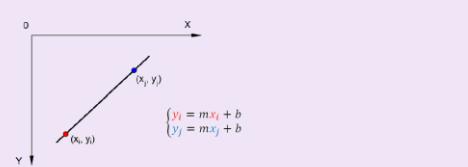
5. A segmentação das linhas é controlada por dois *thresholds*  $t_1$  e  $t_2$ , com  $t_1 > t_2$ . Todos os pixels acima de  $t_2$  (pixels fortes) pertencem às linhas de fronteira; todos os que estão abaixo de  $t_1$  (pixels fracos) não pertencem. Entre  $t_1$  e  $t_2$ , a escolha dos pixels é executada por histerese (*hysteresis*).

Transformação de Hough

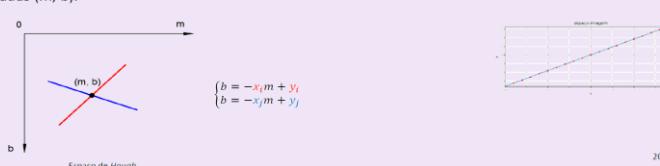
**Transformação de Hough:** é um algoritmo originalmente desenhado para detectar linhas rectas (tendo sido posteriormente alargado para linhas curvas).

A equação reduzida de uma linha, representada no espaço-imagem, escreve-se com base no conhecimento do seu declive  $m$  e do valor da ordenada na origem  $b$ . O ponto de coordenadas  $(x, y)$  pertence à linha.

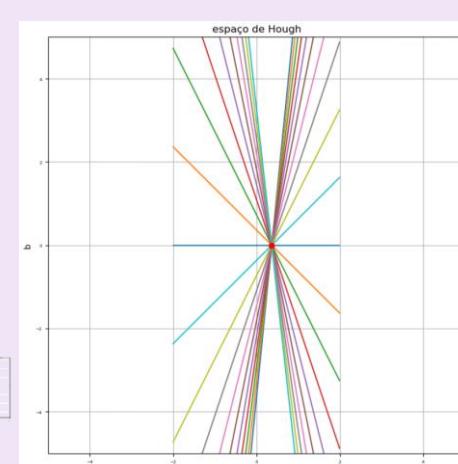
No espaço-imagem, dois pontos de uma mesma linha dão origem a duas equações reduzidas, em que se mantêm iguais os parâmetros  $m$  e  $b$ .



Reescrevendo as equações com base nos parâmetros  $(m, b)$ , define-se o espaço de Hough. Neste espaço, para um dado conjunto de pontos  $\{(x_k, y_k) : k > 1\}$ , sobre a linha do espaço-imagem, obtém-se linhas com diferentes inclinações, intersectando-se todas num ponto de coordenadas  $(m, b)$ .



O espaço de Hough é, na verdade, uma “função acumuladora”, ou seja, em cada ponto desse espaço é adicionada uma unidade sempre que por ele passe uma linha. No caso da figura, o pico da soma está na coordenada  $(m, b) = (tg(20^\circ), 0) = (0.36, 0)$ , ou seja os pontos pertencem todos à mesma recta de declive 0.36, que passa em  $(0, 0)$ .



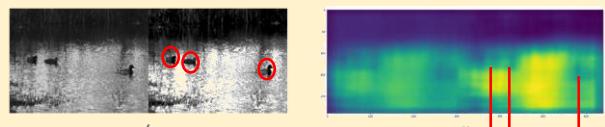


**Medidas de similitude:****3 – Correlação cruzada:**

É o processo de fazer correr uma matriz  $t$  de pequenas dimensões ( $m \times n$ ) por todas as posições de uma matriz  $f$  de maiores dimensões ( $M \times N$ ), calculando o valor da correlação em cada posição (multiplicando os pixels de  $t$  pelos pixels de  $f$  e somando todas essas parcelas).

$$CC(x, y) = \sum_{i=-a}^a \sum_{j=-b}^b f(x+i, y+j) \times t(i, j)$$

$$a = \text{int}\left(\frac{m}{2}\right); b = \left(\frac{n}{2}\right)$$



- Os valores de CC estão compreendidos entre 0 e  $\infty$ .
- A correlação é sensível às variações de luminosidade.

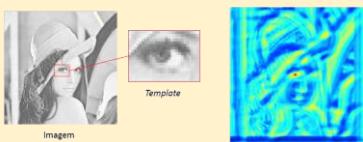
**4 – Correlação Cruzada Normalizada:**

$$CCN(x, y) = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [(f_{i,j} - \bar{f}_{i,j}) \times (t - \bar{t})]}{\sqrt{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [(f_{i,j} - \bar{f}_{i,j})^2] \times \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [(t - \bar{t})^2]}}$$

- A subtração das médias torna o processo insensível às diferenças de brilho.
- A divisão pelos quadrados das diferenças torna o processo insensível às diferenças de contraste.
- Os valores de CCN estão compreendidos entre -1 e 1.

**Template Matching**

Técnica de segmentação que visa encontrar objectos semelhantes a um dado objecto que se designa por *template*. É geralmente realizada com base no cálculo da correlação cruzada normalizada.



Fernando Soares

**5 – Correlação de Fase:**

Técnica que permite determinar uma translação de imagem, através da determinação do espectro de potência cruzado (cross power spectrum). Sejam  $f_1$  e  $f_2$  duas imagens deslocadas uma da outra:

O desfasamento entre as duas imagens pode ser calculado da seguinte forma:

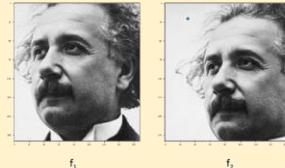
- $g_1$  = transformada de Fourier da imagem  $f_1$ ;
- $g_2$  = transformada de Fourier da imagem  $f_2$ ;
- Cálculo do espectro de potência cruzado:

$$CF = \text{real} \left[ \text{IFT} \left( \frac{g_1 \times \bar{g}_2}{|g_1 \times \bar{g}_2|} \right) \right]$$

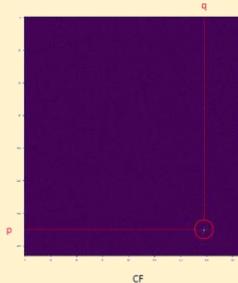
IFT: inversa da transformada de Fourier  
real(z): parte real do número complexo z  
 $\bar{g}_2$ : conjugado do número complexo  $g_2$

- Determinação do pico máximo de CF.

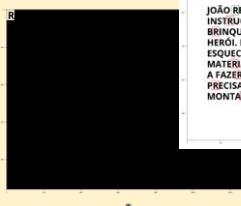
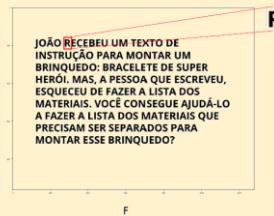
Exemplo: translação de imagem.



Pico = máximo de CF  
Deslocamento =  $(M - p, N - q)$



Exemplo: template matching.



JOÃO RECEBEU UM TEXTO DE INSTRUÇÃO PARA MONTAR UM BRINQUEDO: BRACELETE DE SUPER HERÓI. MAS, A PESSOA QUE ESCRVEU, ESQUECEU DE FAZER A LISTA DOS MATERIAIS. VOCÊ CONSEGUE AJUDÁ-LO A FAZER A LISTA DOS MATERIAIS QUE PRECISAM SER SEPARADOS PARA MONTAR ESSE BRINQUEDO?