

Processo de deteção automática do papel hidrosensível em imagens tiradas através de sistemas SmartPhone

Nuno Costa

Orientador: Prof. André Marçal

Sessão pública do Seminário de Modelação

Mestrado em Engenharia Matemática

FCUP

Janeiro 2013

Estrutura da apresentação

Introdução

- Exemplos de papéis hidrosensíveis
- Exemplo de processo de fertilização agrícola
- O que é uma imagem digital?
- O que é uma imagem RGB?

Segmentação de uma imagem RGB

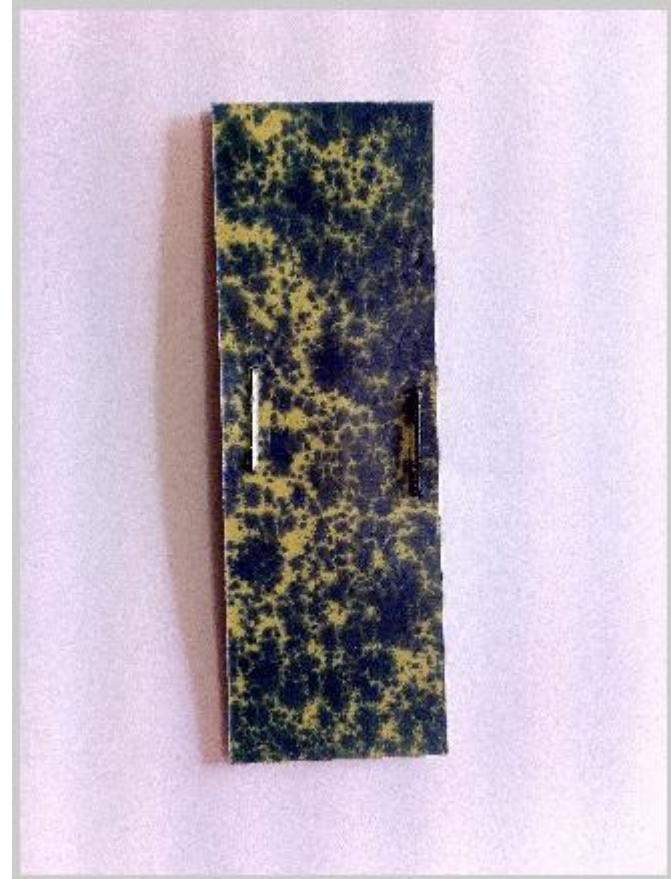
Método dos k-médias

Transformada de Hough

- Transformada de Hough para linhas
- Transformada de Hough para círculos

Bibliografia

Exemplos de papéis hidrosensíveis:



Exemplo de processo de fertilização agrícola:



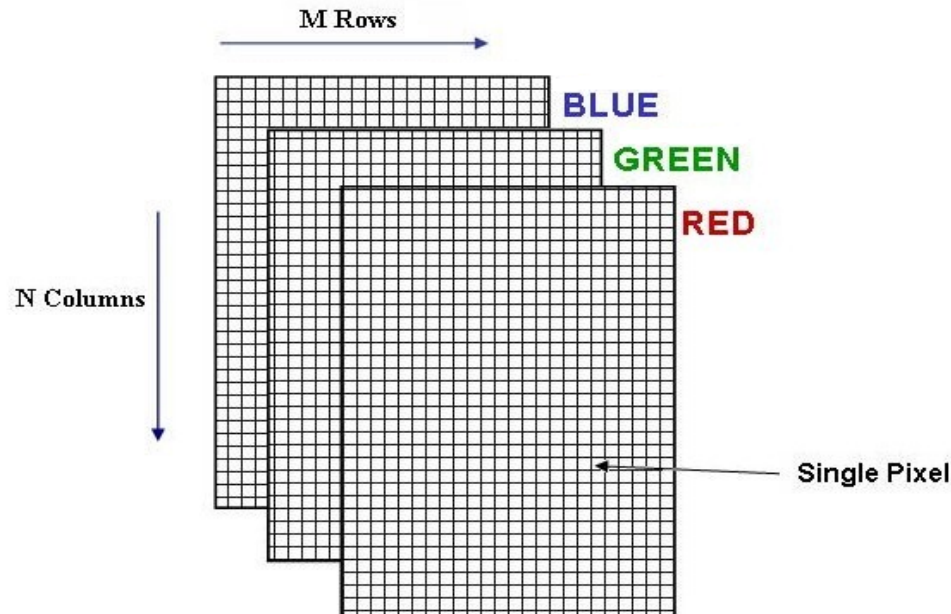
O que é uma imagem digital?

Uma imagem pode ser definida como sendo uma função de duas dimensões $f(x,y)$, onde x e y são as coordenadas no espaço e a amplitude de f associado a cada ponto (x,y) é chamada de intensidade. Se os valores de x , y e f forem quantidades finitas e discretas, então a imagem é considerada uma imagem digital. Cada ponto (x,y) de uma imagem digital é normalmente conhecido como sendo um píxel.



O que é uma imagem RGB?

Uma imagem RGB (Red, Green, Blue) é uma matriz de intensidades do tipo $M \times N \times 3$, i. e. uma matriz “normal” de intensidades $M \times N$ dividida em três componentes, sendo estas as três cores primárias: vermelho, verde e azul. Ou seja, uma imagem RGB é constituída pela componente vermelha, pela componente verde e pela componente azul. Cada uma destas componentes é uma imagem do tipo *gray-scale* (escala cinzenta), i. e. uma imagem com vários tons de cinzento.



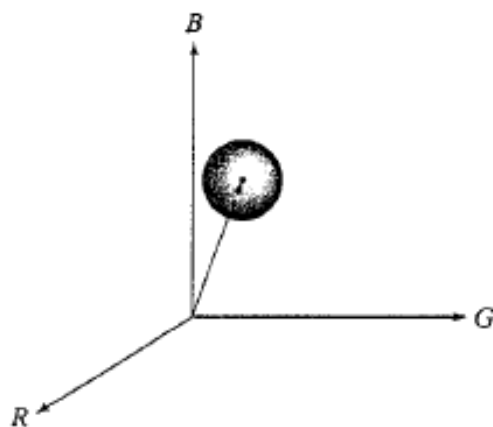
Segmentação de uma imagem RGB

Considera-se uma imagem RGB. Suponha-se que o objetivo é de segmentar os objetos, da imagem RGB, com um determinado intervalo de intensidades ou de cor. Para tal seleciona-se uma ou várias amostras do objeto de interesse que queremos segmentar. Esta amostra vai permitir obter uma estimativa da média (vetor com 3 componentes - m) das intensidades correspondentes ao objeto em questão. O objetivo da segmentação é de classificar cada píxel da imagem RGB como fazendo parte ou não desse tal intervalo de intensidades. Para poder classificar cada elemento da imagem RGB, precisa-se de um método de comparação: a distância Euclidiana (o método mais simples). Seja $z = (z_R, z_G, z_B)$ um ponto arbitrário do espaço RGB. Diz-se que z é “semelhante” à m se a distância entre os dois for menor ou igual a um determinado valor T (threshold). A distância Euclidiana entre z e m é dada por:

$$\begin{aligned}
 D(\mathbf{z}, \mathbf{m}) &= \|\mathbf{z} - \mathbf{m}\| \\
 &= [(\mathbf{z} - \mathbf{m})^T (\mathbf{z} - \mathbf{m})]^{1/2} \\
 &= [(z_R - m_R)^2 + (z_G - m_G)^2 + (z_B - m_B)^2]^{1/2}
 \end{aligned}$$

onde $\|\cdot\|$ é a norma no espaço Euclidiano.

Portanto, o lugar geométrico dos pontos cujo $D(\mathbf{z}, \mathbf{m}) \leq T$ é uma esfera sólida de raio T . Ou seja, os pontos contidos no interior ou na superfície da esfera satisfazem o critério previamente considerado, ao contrário dos pontos no exterior da esfera. Codificando com o valor 1, os pontos que satisfazem o critério e com o valor 0, os restantes pontos, obtém-se uma imagem segmentada binária.

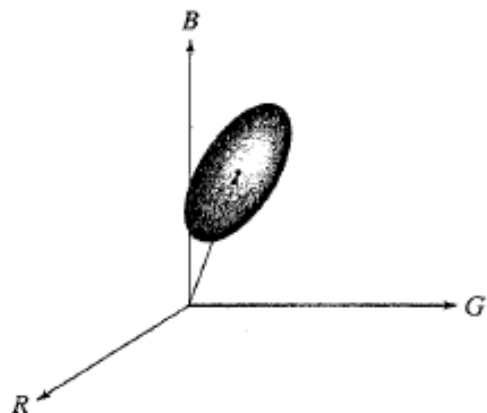


Uma generalização da equação anterior é obtida da seguinte forma:

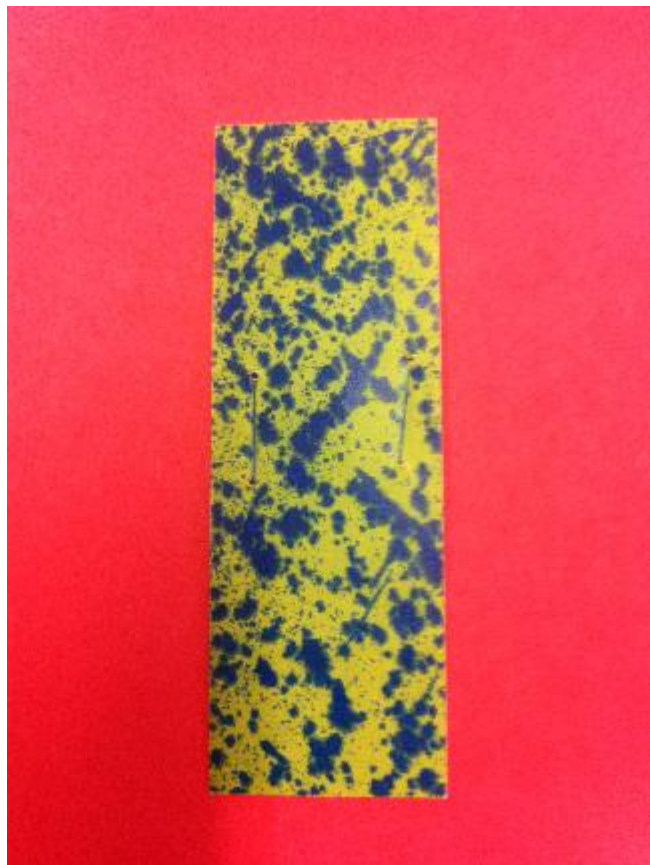
$$D(\mathbf{z}, \mathbf{m}) = [(\mathbf{z} - \mathbf{m})^T \mathbf{C}^{-1}(\mathbf{z} - \mathbf{m})]^{1/2}$$

onde \mathbf{C} é a matriz de covariância da amostra do objeto que queremos segmentar.

Esta distância é conhecida como sendo a distância de Mahalanobis. Neste caso, o lugar geométrico dos pontos cujo $D(\mathbf{z}, \mathbf{m}) \leq T$ descreve um elipsoide. Repara-se que se $\mathbf{C} = \mathbf{I}$, ou seja se a matriz de covariância for igual a matriz identidade, a distância de Mahalanobis reduz-se a distância Euclidiana. A segmentação, utilizando a distância de Mahalanobis, é feita de modo análogo a distância Euclidiana. Isto é, os pontos contidos no interior ou na superfície do elipsoide são codificados com o valor 1 e os restantes com o valor 0.



Exemplo da segmentação de um papel hidrosensível relativamente ao fundo



Método dos k-médias

O método dos k-médias é baseado em métodos de *clustering* e consiste numa técnica iterativa que agrupa todos os pixels de uma imagem em diferentes *clusters* (grupos) com determinadas características.

O objetivo deste método é de minimizar a distância (normalmente é utilizada a distância Euclidiana) entre cada um dos pixels da imagem e os centros dos *clusters*.

Passos do método/ algoritmo k-médias:

1. Escolher o número de *clusters* (**K**).
2. Escolher os **K** centros dos *clusters*, $z_1(1), z_2(1), \dots, z_K(1)$.
3. Atribuir cada um dos pixels x da imagem ao *cluster* que minimiza a distância entre o pixel e o centro do *cluster*, ou seja

$$x \in C_j(l) \text{ se } \|x - z_j(l)\| < \|x - z_i(l)\|$$

para todo $i = 1, 2, \dots, K$ e $i \neq j$. $C_j(l)$ denota o *cluster* cujo centro é $z_j(l)$ na iteração l .

4. Determinar os novos centros dos *clusters* $z_j(l+1)$ para $j = 1, 2, \dots, K$, calculando a média dos pixels pertencentes ao *cluster* $C_j(l)$,

$$z_j(l+1) = \frac{1}{N_j} \sum_{x \in C_j(l)} x, \text{ para } j = 1, 2, \dots, K$$

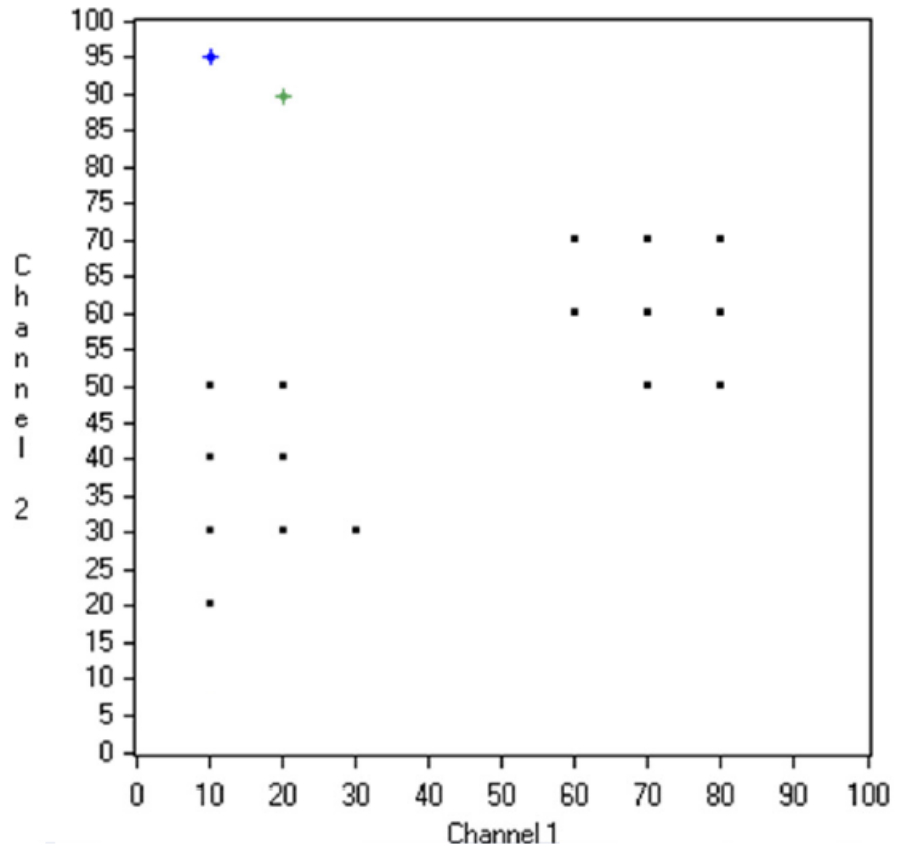
onde N_j é o número de pixels pertencentes ao *cluster* $C_j(l)$.

5. Repetir os passos 3 e 4 enquanto houver mudanças de pixels entre os *clusters*. Caso contrário, o procedimento está terminado.

Este algoritmo garante a convergência dos centros para determinados valores, mas não garante a solução ideal. Ou seja, a solução final irá depender dos centros inicialmente escolhidos.

Exemplos

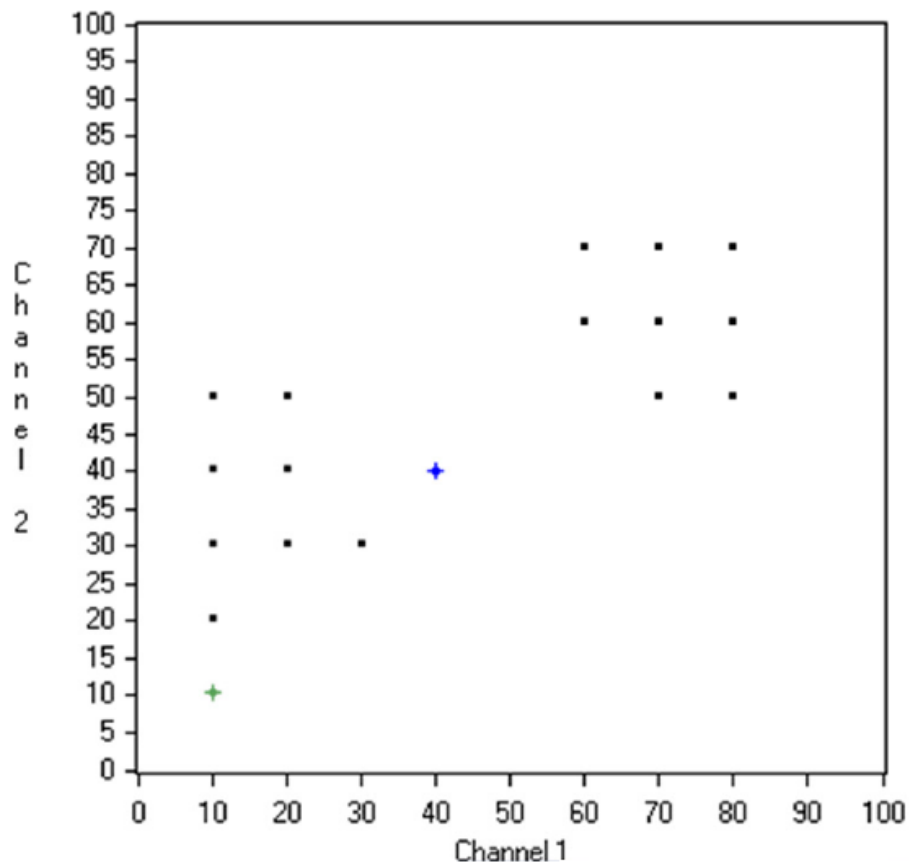
- Escolha incorreta dos centros iniciais $z_1(1)$ e $z_2(1)$



No Clusters:

2

- Escolha mais adequada dos centros iniciais $z_1(1)$ e $z_2(1)$



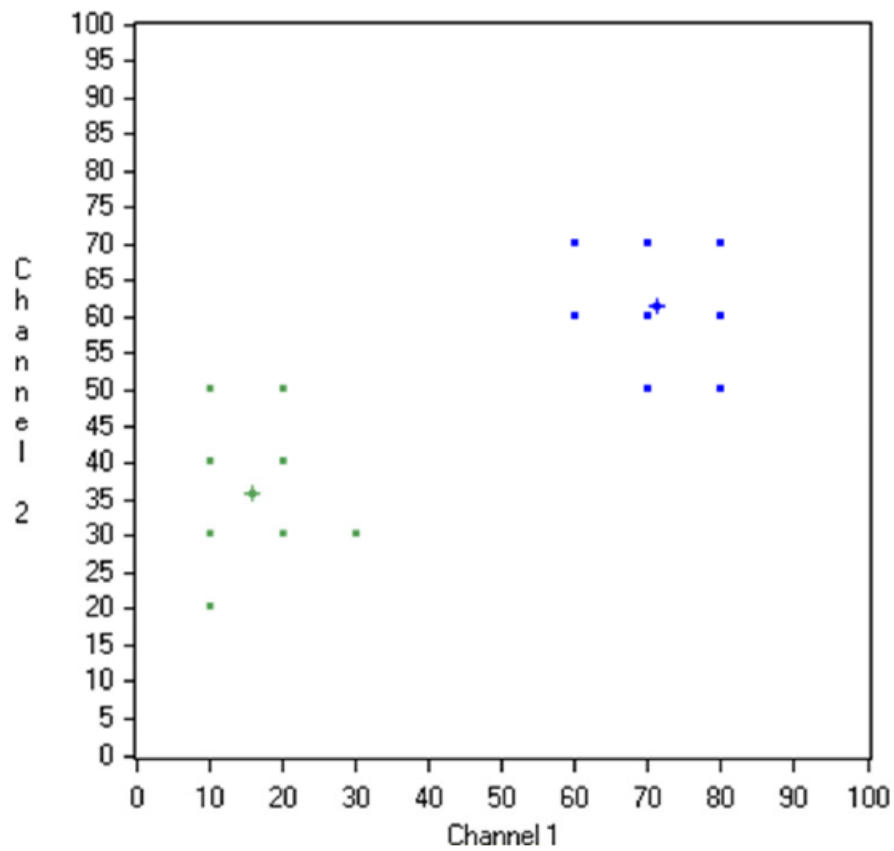
No Clusters:

2

Seed:

Cluster 1:
10,10

Cluster 2:
40,40



Iteration : 3

FIM

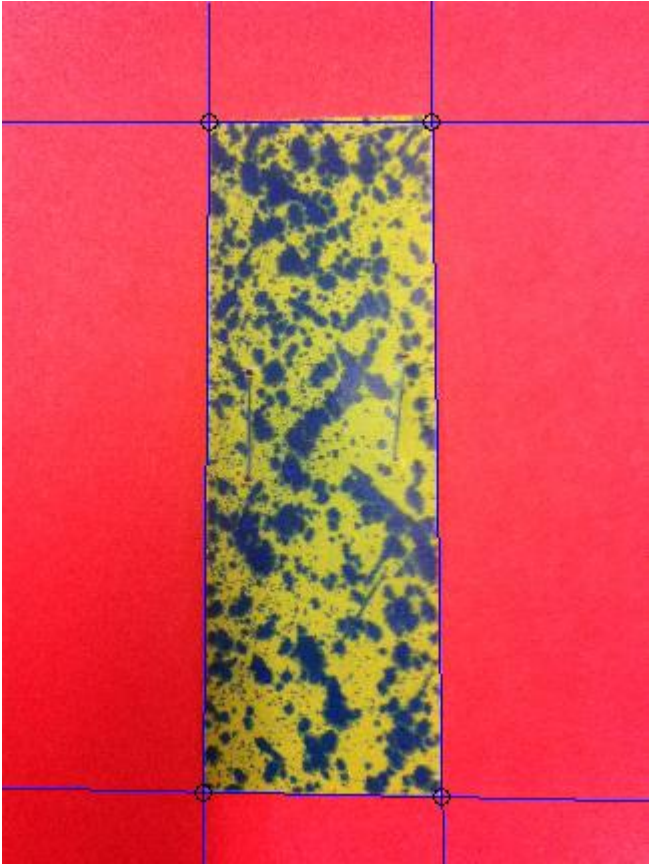
Cluster 1

Mean Position:
16.25000, 36.25000

Cluster 2

Mean Position:
71.25000, 61.25000

Exemplo da segmentação das gotas relativamente ao papel hidrosensível baseado no método dos k-médias



Transformada Hough

- Normalmente, o método de Hough é precedido pela utilização de métodos de deteção de bordas ou arestas (*edge detection*) em imagens, por exemplo o detetor de Sobel ou de Canny ou operações morfológicas. Ou seja, métodos que permitem detetar descontinuidade no que diz respeito as intensidades dos pixéis numa imagem.
- O que acontece ao utilizar estes métodos, de modo geral, é que não se consegue obter um conjunto de pixéis que caracterizam completamente uma borda ou aresta devido a presença de ruído, iluminação não uniforme ou outros efeitos que introduzem descontinuidades nas intensidades de uma imagem.
- A transformada de Hough permite fazer a ligação entre esses pixéis descontínuos para formar objetos completos.

Transformada Hough para linhas

Considera-se um ponto (x_i, y_i) pertencente a uma imagem. Sabe-se que existe uma infinidade de retas que passam por este ponto satisfazendo a equação,

$$y_i = ax_i + b$$

com determinados valores de a e b . Reescrevendo a equação anterior da seguinte maneira,

$$b = -x_i a + y_i$$

e considerando o plano ab (igualmente conhecido como o espaço de parâmetros), obtém-se uma única reta para valores de (x_i, y_i) fixos.

Considerando agora um segundo ponto (x_j, y_j) com, igualmente, uma reta associada no espaço de parâmetros que intersecta a reta anterior nas coordenadas (a', b') . No plano xy , a' representa o declive e b' representa a ordenada na origem da reta que passa por ambos os pontos. Ou seja, todos os pontos pertencentes a esta reta no plano xy têm retas que intersectam (a', b') no plano ab .

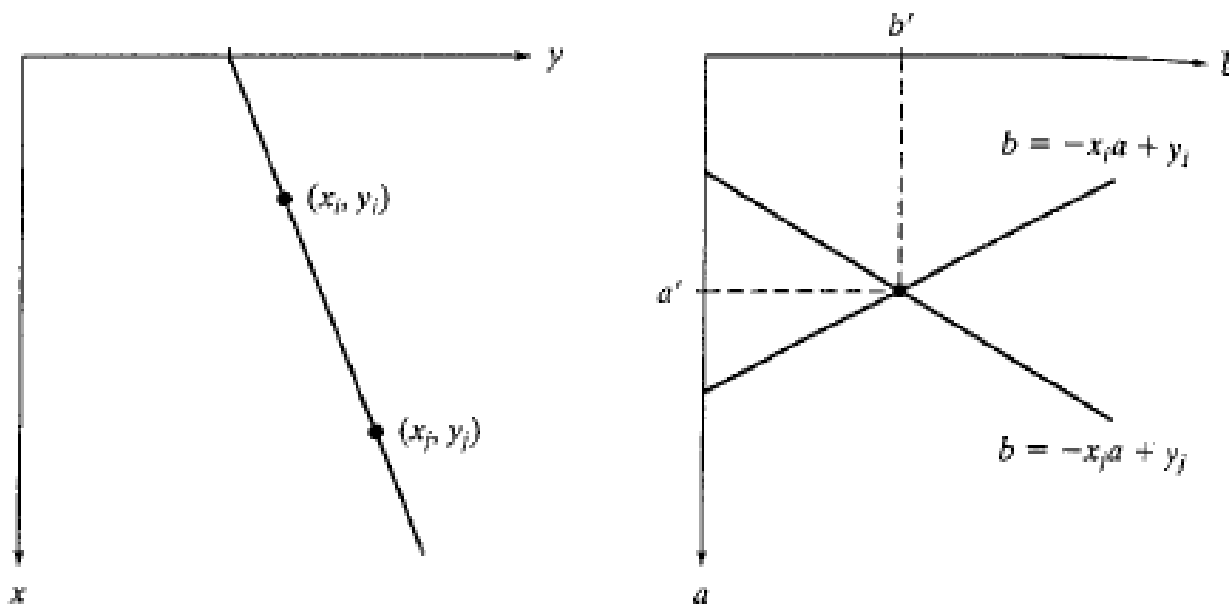


Fig. - Esquerda: Plano xy . Direita: Plano ab (espaço de parâmetros)

Existe no entanto, um problema com esta abordagem:

- o parâmetro a tende para infinito para retas que tendem para a direção vertical no plano xy .

Solução possível:

- utilizar novos parâmetros, ou seja as coordenadas polares (ρ, θ) .

Portanto, a representação de uma reta passa a ser:

$$x\cos\theta + y\sin\theta = \rho$$

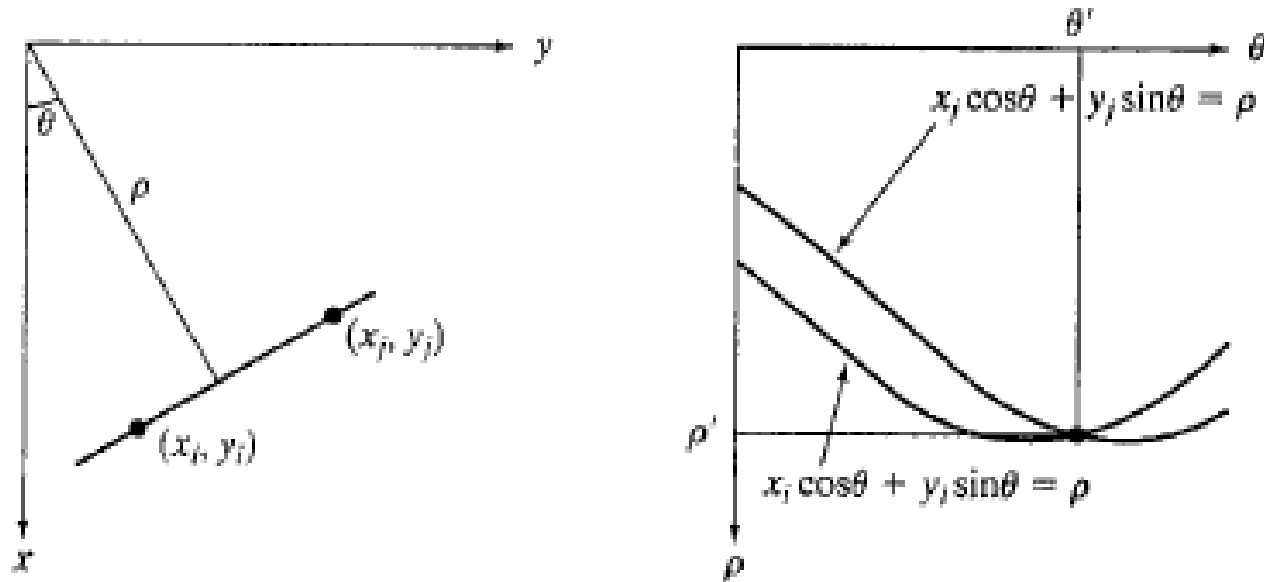


Fig. - Esquerda: Representação geométrica dos parâmetros ρ e θ no plano xy . Direita: Cada curva sinusoidal representa o conjunto de todas as retas que passam por um determinado ponto (x_i, y_i) . O ponto de interseção (ρ', θ') corresponde aos parâmetros $(a = -\frac{\cos \theta}{\sin \theta}$ e $b = \frac{\rho}{\sin \theta})$ da reta que passa por (x_i, y_i) e (x_j, y_j) .

Transformada Hough para círculos

No caso da detecção de formas circulares, os círculos são geralmente mais simples de representar no espaço de parâmetros comparando com as linhas, no sentido em que os parâmetros de um círculo são diretamente transferidos para o espaço de parâmetros. Dado a equação de um círculo,

$$r^2 = (x - a)^2 + (y - b)^2$$

onde a e b são as coordenadas do centro do círculo de raio r , repara-se que um círculo possui três parâmetros (r, a e b).

Portanto, o espaço de parâmetros para um círculo pertence a \mathbb{R}^3 , enquanto que para as linhas pertencia a \mathbb{R}^2 . A medida que o número de parâmetros de uma forma aumenta assim como a dimensão do espaço de parâmetros, aumenta-se a complexidade da transformada de Hough.

Computacionalmente, a transformada de Hough para a detecção de círculos funciona com base no mesmo conceito utilizado na detecção de linhas:

- utilização de métodos de detecção de descontinuidades (*edge points*)
- no espaço de parâmetros, cada *edge points* do plano xy define novos círculos centrados nas coordenadas do *edge point* e com valores de raio de 0 até um determinado valor máximo, representando cones de “votos”
- ou seja, os parâmetros $(r', a' e b')$ com maior número de “votos”, no espaço de acumulação (\mathbb{R}^3), indicam a posição do centro do círculo inicial (a', b') de raio r' .

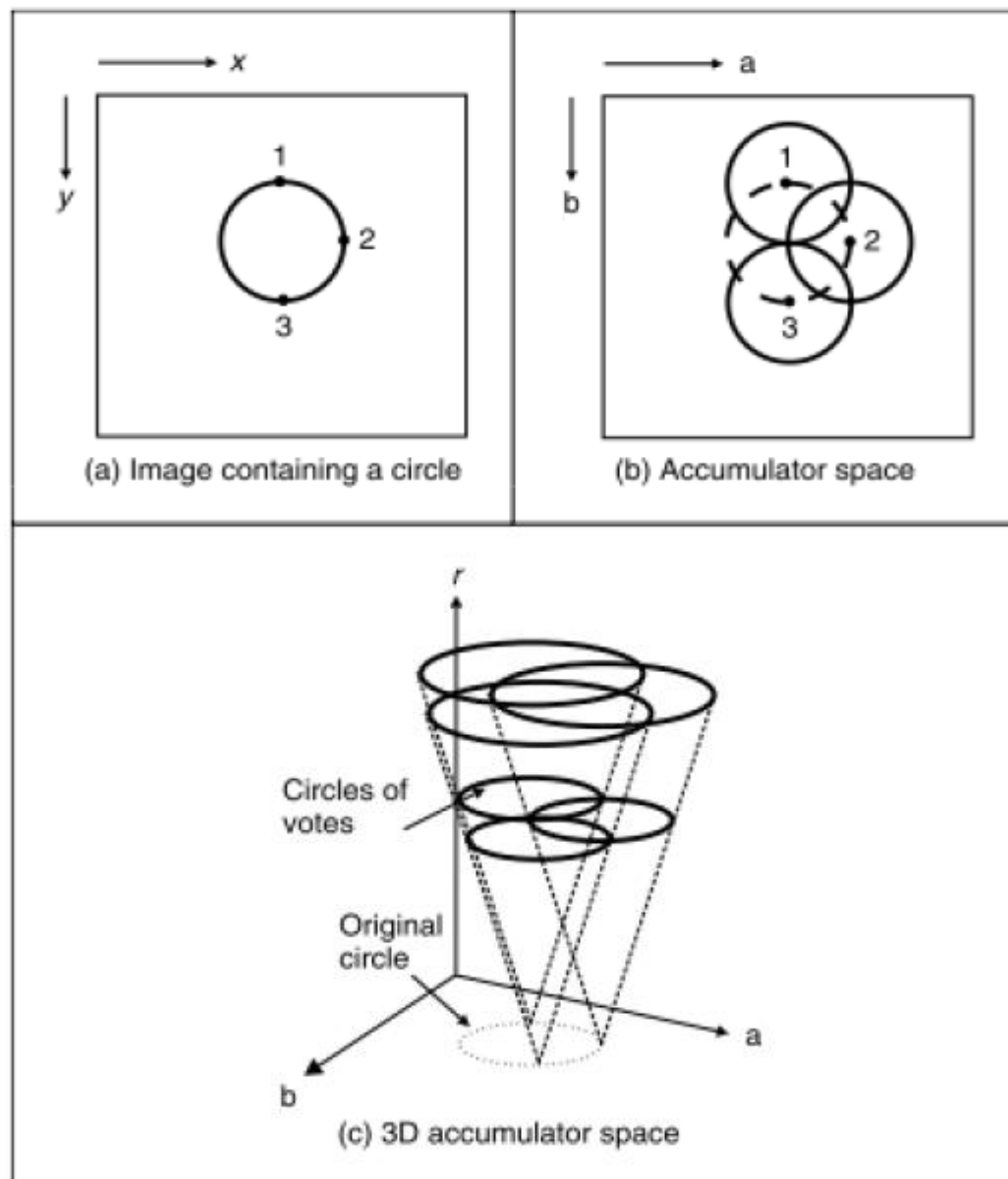


Fig. - (a) Imagem inicial com 3 *edge points* pertencentes a um padrão circular. (b) Cada *edge point*, no espaço de parâmetros, define novos círculos. (c) Representação dos cones de “votos” associados a cada *edge point*.

Bibliografia:

- Gonzalez, R. C., Woods, R. E. and Eddins, S. L. [2004]. Digital image processing using MATLAB, 7nd ed., Prentice Hall, Upper Saddle River, NJ.