

# Processamento Digital de Imagens

Profa. Flávia Magalhães F. Ferreira

PUC Minas

Unidade Vd - Segmentação por Limiarização

# Agenda

- 1 Limiarização
- 2 Limiarização Global
- 3 O papel do ruído na limiarização da imagem
- 4 O papel da iluminação na limiarização da imagem
- 5 Limiarização Local
- 6 Limiarização baseada em diversas variáveis

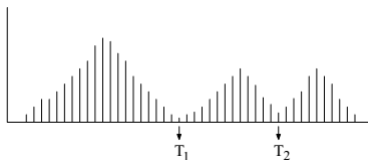
- A limiarização é uma das técnicas mais simples de segmentação e consiste na classificação dos pixels de uma imagem de acordo com a especificação de um ou mais limiares.
- Seja o histograma de níveis de cinza correspondente a uma imagem  $f(x, y)$ , de maneira que os pixels dos objetos e os do fundo tenham seus níveis de cinza separados em dois grupos dominantes.
- Uma maneira óbvia de extrair os objetos do fundo é por meio da seleção de um limiar  $T$  que separe os dois grupos.

- Então, cada ponto  $(x, y)$  tal que  $f(x, y) > T$  é denominado um ponto do objeto; caso contrário, o ponto é denominado um ponto do fundo.
- A imagem limiarizada  $g(x, y)$  pode ser definida como

$$\begin{aligned} 0, & \text{ se } f(x, y) \leq T \\ 1, & \text{ se } f(x, y) > T \end{aligned}$$

- Portanto, os *pixels* rotulados como 1 (ou qualquer outro nível de cinza conveniente) correspondem aos objetos, enquanto que aqueles rotulados como 0 correspondem ao fundo.
- Nesse caso, a limiarização é denominada binarização, pois a imagem resultante possui apenas dois valores de intensidade, 0 (preto) ou 1 (branco).

- A figura a seguir mostra um caso ligeiramente mais geral dessa abordagem, em que três grupos dominantes caracterizam o histograma da imagem (por exemplo, dois tipos de objetos iluminados sobre um fundo escuro).



- A limiarização da imagem pode ser definida como

$$\begin{aligned} l_1, & \quad \text{se } f(x, y) \leq T_1 \\ l_2, & \quad \text{se } T_1 < f(x, y) \leq T_2 \\ l_3, & \quad \text{se } f(x, y) > T_2 \end{aligned}$$

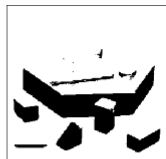
tal que, para cada intervalo, é especificado um nível de cinza correspondente.

# Limiarização

- A seleção correta do valor de limiar é crucial para que o processo de segmentação baseada na limiarização produza bons resultados.
- A figura a seguir ilustra a segmentação de uma imagem com o uso de diferentes valores de limiar.



(a) imagem original



(b)  $T = 108$



(c)  $T = 179$

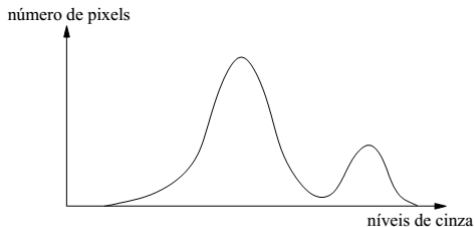


(d)  $T = 213$

- Limiarização global: a utilização de um único valor de limiar para segmentar toda a imagem não é em geral adequada, pois as imagens podem conter variações nos níveis de cinza dos objetos e do fundo por causa da iluminação não-uniforme, ruído, parâmetros do dispositivo de aquisição não-uniformes, ou outros fatores.
- Limiarização local: valores de limiar podem variar sobre a imagem como uma função de suas características locais.

# Limiarização Global

- Como mencionado anteriormente, a maneira mais direta de selecionar um valor global de limiar é a partir da distribuição das intensidades dos *pixels* na imagem.
- Para uma imagem consistindo de objetos com níveis de cinza aproximadamente iguais, diferindo do nível de cinza do fundo, o histograma possuirá dois picos distintos, um formado pelos *pixels* dos objetos e outro pelos *pixels* do fundo.
- Um histograma com tal distribuição é chamado bimodal.

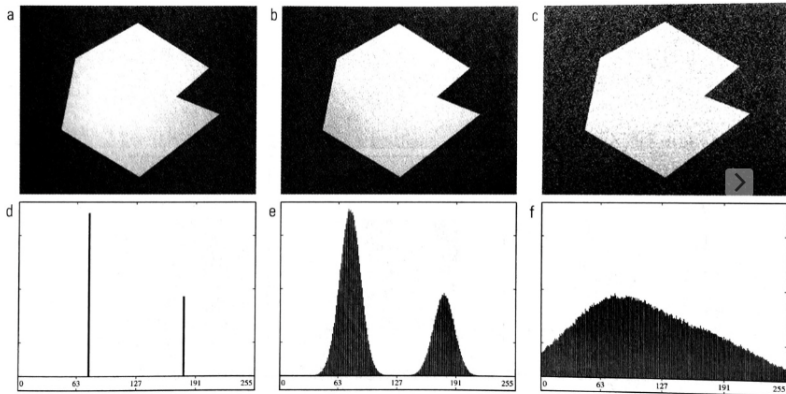


(a) histograma bimodal



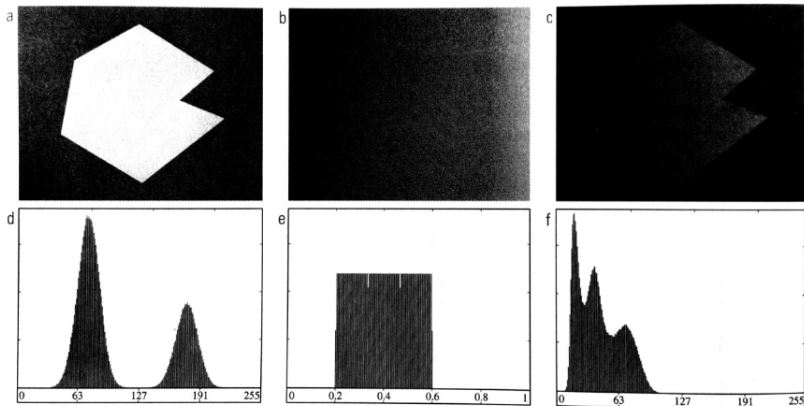
- Intuitivamente, o valor do limiar corresponde ao nível de cinza que melhor separa os dois picos.
- Vários métodos têm sido propostos para encontrar os pontos de máximo dos histogramas (picos) e, então, selecionar o valor de limiar como o ponto de mínimo (vale) entre eles.
- Há várias desvantagens em selecionar o limiar global por meio da distribuição das intensidades dos *pixels*.
  - Nem sempre as intensidades dos objetos e do fundo da imagem são bem distintas, em virtude, por exemplo, da ocorrência de baixo contraste ou ruído.
  - Além disso, a determinação de picos e vales não é um problema trivial. Por exemplo, muitos pontos de mínimo ou máximo locais podem existir na imagem.

# O papel do ruído na limiarização da imagem



(a) Imagem com 8 bits, livre de ruído; (b) Imagem com ruído gaussiano aditivo com média 0 e desvio padrão de 10 níveis de intensidade; (c) Imagem com ruído gaussiano aditivo com média 0 e desvio padrão de 50 níveis de intensidade; (d) a (f) Histogramas correspondentes.

# O papel da iluminação na limiarização da imagem



(a) Imagem ruidosa; (b) Rampa de intensidade no intervalo  $[0, 255]$ ; (c) Produto de (a) e (b); (d) a (f) Histogramas correspondentes.

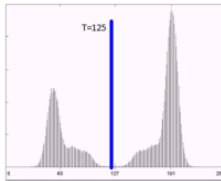
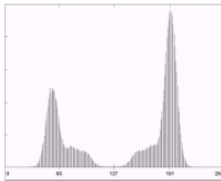
# Limiarização global simples

- 1 Selecionar uma estimativa inicial para o limiar global  $T$ , tal que  $T$  seja maior que o mínimo valor de intensidade da imagem e menor que o máximo valor de intensidade da imagem.
- 2 Segmentar a imagem usando  $T$ . Isso dará origem a dois grupos de pixels:  $G_1$ , com valores  $> T$ , e  $G_2$ , com valores  $\leq T$ .
- 3 Calcular os valores de intensidade média  $m_1$  e  $m_2$ , para os grupos  $G_1$  e  $G_2$ , respectivamente.
- 4 Calcular um novo valor de limiar:

$$T = \frac{1}{2}(m_1 + m_2)$$

- 5 Repetir as etapas 2 a 4 até que a diferença entre os valores de  $T$ , em iterações sucessivas, seja menor que  $\Delta T$ .

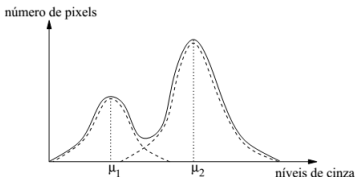
# Limiarização



A aplicação do algoritmo anterior encontrou  $T = 125,4$  após 3 iterações, começando com  $T = m$ .

# Limiarização Global: problema teórico de Decisão Estatística

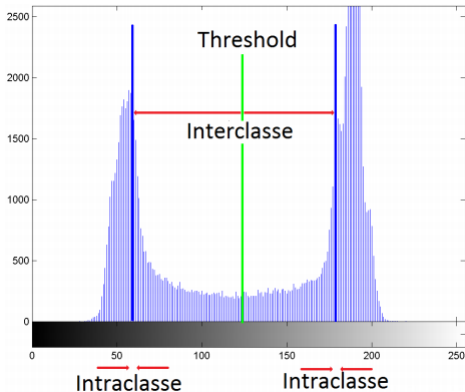
- Algumas técnicas determinam o valor de limiar por meio da otimização de certas medidas de separação entre as classes de objetos na imagem.
- As distribuições de intensidade dos *pixels* podem ser modeladas como funções de densidade de probabilidades, por exemplo, a função Gaussiana.



- Para cada distribuição, a média e a variância são calculadas, e o limiar é escolhido tal que a probabilidade de erro de classificação, definido por essas funções de probabilidade, seja minimizado.

# Limiarização Global: Método de Otsu

- Otsu (1979) propôs uma técnica não paramétrica (não estima parâmetros do modelo) e não supervisionada para a seleção automática do limiar visando à segmentação da imagem.
- Otimizar a separabilidade das classes (objeto e fundo):
  - maximizar variabilidade interclasse (variância entre as classes)
  - minimizar variabilidade intraclasse



# Limiarização Global: Método de Otsu

- Calcular o histograma normalizado da imagem de entrada. Designar os componentes do histograma como:

$$p_i = n_i / MN$$

para  $i = 0, 1, \dots, L - 1$ , em que  $n_i$  é o número de pixels com intensidade  $i$  e  $\{0, 1, \dots, L - 1\}$  são os  $L$  distintos níveis de intensidade,  $M$  e  $N$  são as resoluções espaciais da imagem.

- Calcular as probabilidades acumuladas até a intensidade  $k$ :

$$P_a(k) = \sum_{i=0}^k p_i$$

para  $k = \{0, 1, \dots, L - 1\}$

- Calcular as médias de intensidade até cada uma das intensidades  $k$  (para  $k = \{0, 1, \dots, L - 1\}$ ):

$$m(k) = \sum_{i=0}^k i p_i$$



# Limiarização Global: Método de Otsu

- Calcular a intensidade média global  $m_G$ :

$$m_G = \sum_{i=0}^{L-1} i p_i$$

- Calcular a variância global  $\sigma_G^2$ :

$$\sigma_G^2 = \sum_{i=0}^{L-1} (i - m_G)^2 p_i$$

- Calcular a variância entre as classes,  $\sigma_B^2(k)$ :

$$\sigma_B^2(k) = \frac{[m_G P_a(k) - m(k)]^2}{P_a(k)[1 - P_a(k)]}$$

para  $k = \{0, 1, \dots, L - 1\}$

- Obter o Limiar de Otsu  $k^*$ , que é o valor de  $k$  que maximiza  $\sigma_B^2(k)$ :

$$\sigma_B^2(k^*) = \max_{0 \leq k \leq L-1} \sigma_B^2(k)$$

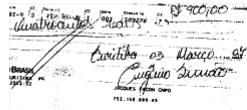
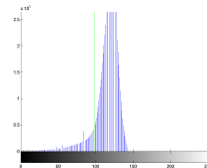
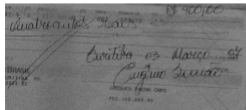
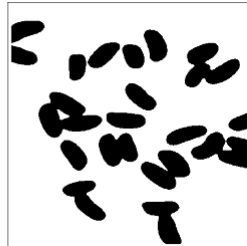
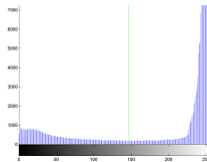
para  $k = \{0, 1, \dots, L-1\}$

- Obter a medida de separabilidade normalizada  $\eta^* = \eta(k^*)$  avaliada para  $k^*$ :

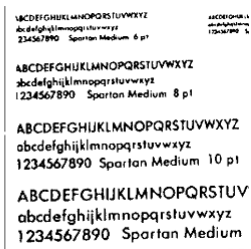
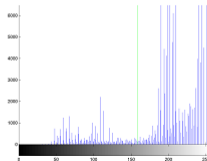
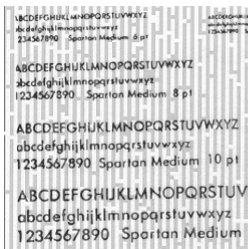
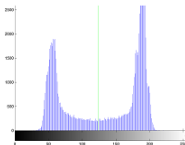
$$\eta^* = \frac{\sigma_B^2(k^*)}{\sigma_G^2}$$

com  $\sigma_G^2$  sendo a variância global e  $\sigma_B^2(k^*)$  sendo a variância entre as classes usando o limiar  $k^*$ .

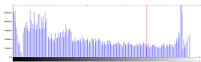
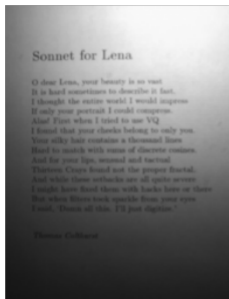
# Limiarização Global - Otsu



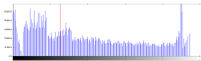
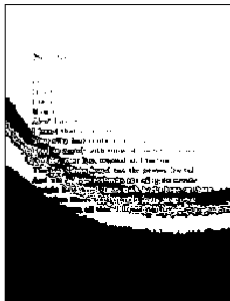
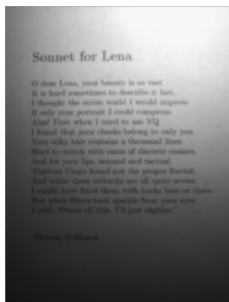
# Limiarização Global - Otsu



- Limitação de Limiarização Global

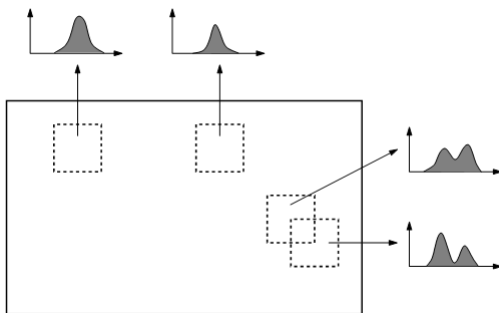


- Limitação de Limiarização Global



# Limiarização Local

- Em casos nos quais a utilização de um único limiar para segmentar toda a imagem não produza bons resultados, a limiarização local pode ser uma alternativa mais apropriada.
- Uma forma comum de realizar a limiarização local é analisar as intensidades dos *pixels* em uma região da imagem para determinar limiares locais. Esse processo é ilustrado na figura a seguir.



# Limiarização Local

- Algumas medidas estatísticas simples para calcular um limiar local na vizinhança  $v$  do ponto  $p$  são:  
a média dos valores

$$T = \text{media}_v(p)$$

a mediana dos valores

$$T = \text{mediana}_v(p)$$

ou a média dos valores mínimo e máximo

$$T = \frac{\min_v(p) + \max_v(p)}{2}$$

- Um problema crítico encontrado nas técnicas de limiarização local é a escolha do tamanho da janela da vizinhança  $v$ .
- O tamanho selecionado deve ser suficientemente grande para garantir uma boa estimativa do valor da média, mas sem sofrer distorções por causa da não-uniformidade do fundo.



# Limiarização Local - Niblack

- A ideia de Niblack (1986) foi a estabelecer uma função de limiar que levasse em consideração o contraste local de uma janela  $M \times N$  centralizada no pixel  $p$  usando medidas estatísticas:

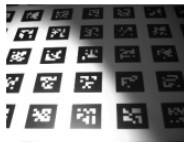
$$T(p) = \mu(p) + k\sigma(p)$$

onde  $\mu(p)$  e  $\sigma(p)$  são a média e o desvio padrão dos *pixels* dentro da janela,  $k$  é um parâmetro definido pelo usuário (valor *default* é 0,5) e  $T(p)$  representa o valor do limiar a ser aplicado no ponto  $p$ .

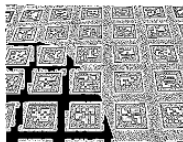
- Realiza-se uma varredura com a janela por toda a imagem, sendo avaliados somente os pixels dentro da janela. Portanto, pode ser classificado como um algoritmo local e adaptativo, pois calcula um limiar para cada ponto da imagem.
- É um bom método quando se deseja minimizar a influência de sombras e iluminação e priorizar o contraste presente em regiões da imagem. Porém, sua utilização em imagem com grande variação de tons de cinza, utilizando uma janela pequena, pode acarretar na inclusão de ruídos.

# Limiarização Local - Niblack

- Resultado



$310 \times 240$



$N = M = 3$



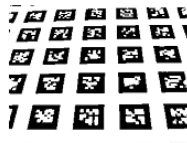
$N = M = 15$



Otsu



$N = M = 21$



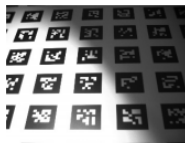
$N = M = 31$

- Sauvola & Pietikainen (2000) propuseram uma nova formulação (pequena modificação em relação ao algoritmo de Niblack), em que o desvio padrão  $\sigma(p)$  pondera diretamente o valor médio da janela,  $\mu(p)$ :

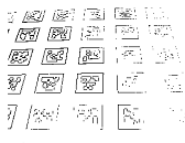
$$T(p) = \text{round} \left\{ \mu(p) \left[ 1 + k \left( \frac{\sigma(p)}{R} - 1 \right) \right] \right\}$$

onde  $R$  está relacionado com a faixa dinâmica do desvio padrão (valor *default* é 128).

- Resultado



$310 \times 240$



$N = M = 3$



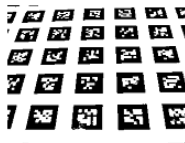
$N = M = 15$



Otsu

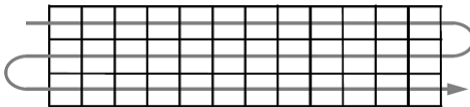


$N = M = 21$



$N = M = 31$

- Um outra abordagem de limiarização local, **menos cara computacionalmente**, é aquela baseada em média móvel (*moving average*) ao longo das linhas de digitalização da imagem, como proposto por Wellner (1993).
- Método muito útil no processamento de documentos, em que a velocidade é um requisito fundamental.
- A forma em que a janela para o cálculo da média móvel é deslocada na imagem tem um papel importante no resultado. Normalmente, obedece um padrão de linha por linha, em zigue-zague, para reduzir o viés de iluminação.



- Digamos que  $f_{k+1}$  denota a intensidade  $f$  do ponto encontrado na etapa  $(k + 1)$ . A média móvel (intensidade média) com este novo ponto é dada por:

$$\begin{aligned}\mu(k + 1) &= \frac{1}{n} \sum_{i=k+2-n}^{k+1} f_i \\ &= \mu(k) + \frac{1}{n}(f_{k+1} - f_{k-n})\end{aligned}\tag{1}$$

onde  $n$  define o tamanho da janela (número de pontos usados no cálculo da média) e  $m(1) = f_1/n = f(0,0)/n$  (valor que seria obtido para a média no primeiro ponto da primeira linha, caso, no tratamento de bordas, inseríssemos  $n - 1$  zeros).

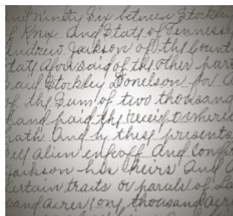
- O algoritmo é inicializado uma única vez, e não em todas as linhas.

- A média móvel é calculada para cada ponto da imagem e a segmentação é calculada da seguinte forma:

$$g(x, y) = \begin{cases} 1, & \text{se } f(x, y) > T_{xy} \\ 0, & \text{se } f(x, y) \leq T_{xy} \end{cases}$$

com  $T_{xy} = k m_{xy}$ , em que  $k$  é constante e  $m_{xy}$  é a média móvel calculada na vizinhança do ponto  $(x, y)$  da imagem.

- Resultado

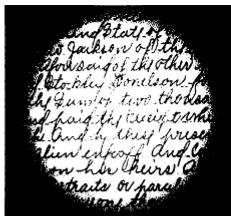


811 × 757

and Ninety Six between Stocken  
of Knox. And Stat of Tennessee  
between Jackson of the bound  
tary of the said of the other part  
aid Stocken Donelson for a  
of the sum of two thousand  
and paid the receipt where  
with and by these presents  
of alien excoff. And Confe  
arison hbr theirs And C  
ertain traits or parals of de  
and aere / ony thousand

Niblack

$N = 40$  e  $k = 0.5$



Otsu

and Ninety Six between Stocken  
of Knox. And Stat of Tennessee  
between Jackson of the bound  
tary of the said of the other part  
aid Stocken Donelson for a  
of the sum of two thousand  
and paid the receipt where  
with and by these presents  
of alien excoff. And Confe  
arison hbr theirs And C  
ertain traits or parals of de  
and aere / ony thousand

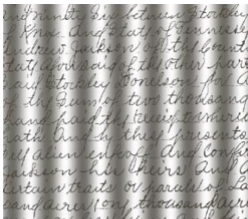
Wellner

$n = 40$  e  $k = 0.95$

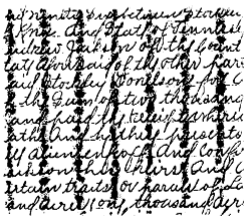


# Limiarização Local

- Resultado



806 × 710

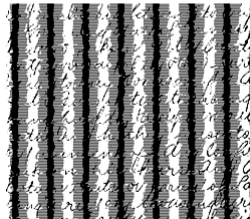


Niblack

$N = 40$  e  $k = 0.5$



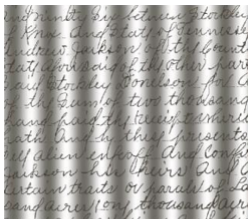
Otsu



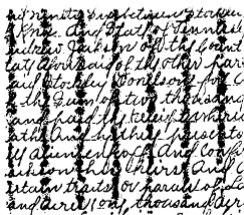
Wellner

$n = 40$  e  $k = 0.95$

- Resultado

A handwritten document in cursive script, likely from the 18th or 19th century. The text is written in dark ink on a light background. The handwriting is somewhat slanted and cursive, with some words being difficult to decipher. The text appears to be a legal or official document, mentioning names like 'Andrew Jackson' and 'Donelson'.

806 × 710

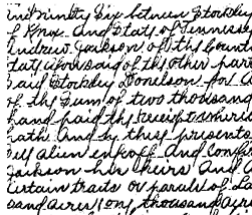
The same handwritten document as above, but with a high-contrast threshold applied. The background is now black, and the text is white. This makes the text stand out more clearly, but some details of the original script are lost.

Niblack

$N = 40$  e  $k = 0.5$

The same handwritten document as above, but with Otsu's thresholding method applied. The background is black, and the text is white. The thresholding is more precise than Niblack's, resulting in a cleaner image with less noise.

Otsu

The same handwritten document as above, but with Otsu's thresholding method applied. The background is black, and the text is white. The thresholding is more precise than Niblack's, resulting in a cleaner image with less noise.

Wellner

$n = 40$  e  $k = 0.95$

# Limiarização baseada em diversas variáveis

- Em alguns casos, o sensor pode disponibilizar mais de uma variável (e não apenas um tom de cinza) para identificar os pixels e, assim, permitir uma limiarização multivariada.
- Exemplo: imagem em cores (RGB). Cada pixel pode ser representado por um vetor  $\underline{z} = (R, G, B)^T$ .
- a Limiarização variada pode ser vista como um cálculo de distância.
- Denotemos por  $\underline{a}$  uma cor média na qual estamos interessados para a segmentação. Então segmentamos a imagem de entrada da seguinte forma:

$$g = \begin{cases} 1, & \text{se } D(\underline{z}, \underline{a}) \leq T, \\ 0, & \text{caso contrário} \end{cases}$$

que  $T$  é um limiar e entende-se que o cálculo da distância é realizado para todos os pixels da imagem de entrada. Aqui,  $D(\underline{z}, \underline{a})$ , que é a distância de cor, define um volume no espaço de cor.

- A **distância euclidiana** n-dimensional é definida como:

$$\begin{aligned} D(\underline{z}, \underline{a}) &= ||\underline{z} - \underline{a}|| \\ &= [(\underline{z} - \underline{a})^T (\underline{z} - \underline{a})]^{\frac{1}{2}} \end{aligned}$$

em que  $D(\underline{z}, \underline{a}) = T$  descreve uma esfera no plano 3D, ou hiperesfera no plano n-dimensional.

- Uma medida de distância mais poderosa é a **distância de Mahalanobis**, definida como:

$$D(\underline{z}, \underline{a}) = [(\underline{z} - \underline{a})^T C^{-1} (\underline{z} - \underline{a})]^{\frac{1}{2}}$$

na qual  $C$  é a matriz de covariância das amostras dos vetores  $\underline{z}$ .

# Distância de Mahalanobis: Explicação Intuitiva

- Considere-se o problema de estimar a probabilidade de um ponto de teste no espaço euclidiano N-dimensional pertencer ao conjunto, onde se tem pontos amostrais que certamente pertencem a esse conjunto. Um primeiro passo poderia ser a determinação da média  $\mu$  do centro de massa dos pontos amostrais  $z$ . Intuitivamente, quanto mais próximo estiver o ponto em questão deste centro de massa, mais provável é que pertença ao conjunto.

# Distância de Mahalanobis: Explicação Intuitiva

- Todavia, precisa-se também de determinar a dimensão do conjunto. Uma abordagem simplista é estimar o desvio-padrão  $\sigma$  das distâncias dos pontos amostrais  $z$  até ao centro de massa  $\mu$ . Se a distância entre o ponto em teste e o centro de massa é menor que um desvio-padrão, poder-se-á concluir que é altamente provável que esse ponto pertença ao conjunto. Quanto mais distante esteja, mais provável é que o ponto não deva ser classificado como pertencente ao conjunto.
- Esta abordagem intuitiva pode tornar-se quantificada ao definir-se a distância normalizada entre o ponto de teste e o conjunto como  $\frac{z-\mu}{\sigma}$ . Relacionando isto com a distribuição normal obtém-se a probabilidade de o ponto de teste pertencer ao conjunto.

# Distância de Mahalanobis: Explicação Intuitiva

- O problema com a abordagem descrita é a suposição de que os pontos amostrais  $z$  são distribuídos à volta do centro de massa  $\mu$  de um modo esférico. Se a distribuição fosse claramente não-esférica, por exemplo elipsoide, deveríamos esperar que a probabilidade de o ponto de teste pertencer ao conjunto iria depender não apenas da distância ao centro de massa, mas também da direção. Nestas direções onde o elipsoide tem um eixo menor, o ponto de teste deverá estar mais próximo, mas se o eixo for maior, então o ponto de teste poderá estar mais distante do centro de massa.

# Distância de Mahalanobis: Explicação Intuitiva

- Distância Euclideana x Distância de Mahalanobis

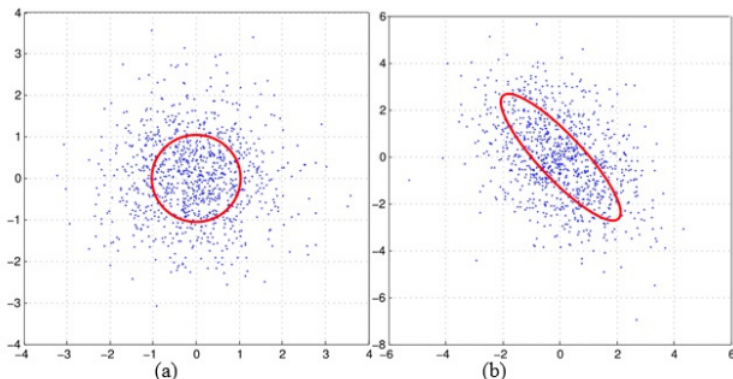


Figura 1. Comparación entre distancia Euclidiana y Mahalanobis: (a) distancia Euclidiana y (b) distancia de Mahalanobis. En ambas figuras, la línea roja denota los puntos que están a igual distancia del centro



# Distância de Mahalanobis: Explicação Intuitiva

- Colocando isto numa formulação matemática, o elipsoide que melhor representa a distribuição de probabilidade do conjunto pode ser estimado construindo a matriz de covariância da amostra. A distância de Mahalanobis é simplesmente a distância do ponto de teste ao centro de massa dividida pelo comprimento do elipsoide na direção do ponto de teste.
- A distância de Mahalanobis tem em conta a variabilidade. Em vez de tratar todos os valores de igual modo quando calcula a distância ao ponto central, pondera-os pela diferença à amplitude de variação na direção do ponto de teste. A fronteira de Mahalanobis torna-se assim clara. Esta função constrói um espaço ao longo do eixo de alongamento elíptico que for detectado.

# Distância de Mahalanobis: Explicação Intuitiva

- Em termos das medidas de Mahalanobis, uma amostra "A" terá um valor substancialmente menor de distância à média que uma amostra "B" se se distribuir ao longo do eixo do grupo com maior variabilidade. Assim, a amostra "A" é mais provavelmente classificada como relacionada com o grupo.