



Tratamiento de Señales

Version 2024-I

Segmentación por Umbralización Método de Otsu

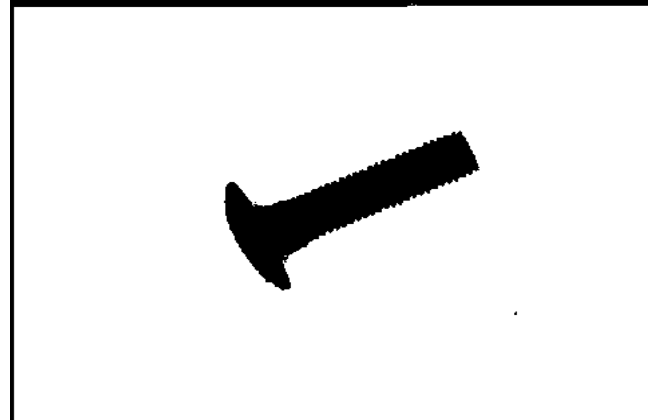
[Capítulo 8]

Dr. José Ramón Iglesias

DSP-ASIC BUILDER GROUP
Director Semillero TRIAC
Ingeniería Electronica
Universidad Popular del Cesar

Introducción

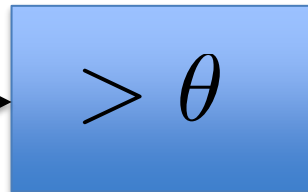
- **Umbralización:** técnica de segmentación empleada cuando hay una clara diferencia entre los objetos a extraer y el fondo.



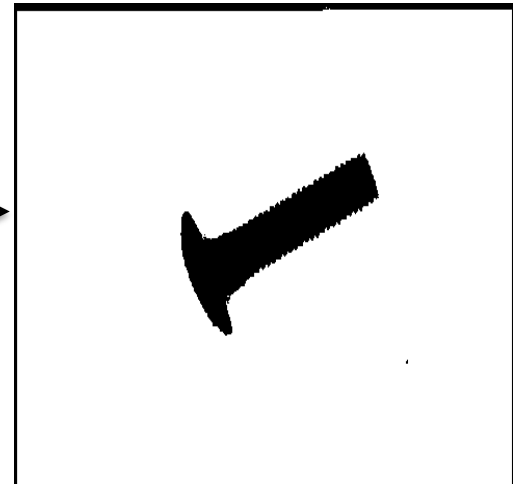
Segmentación por Umbral

La imagen es segmentada a partir de un umbral: los tonos de gris mayores que un umbral pertenecen a la región segmentada, mientras que el resto pertenece al fondo.

X INPUT



Y OUTPUT



Imágenes binarias

Umbralización (Binariaización)

Es el proceso de convertir una imagen de niveles de gris o una imagen a color, en una imagen a blanco y negro, donde

- Los píxeles *negros* corresponde al **fondo** (no interesan)
- Los píxeles *blancos* corresponde al **objeto** (región de interés)

El proceso de umbralización puede ser global o local.

Umbralización global

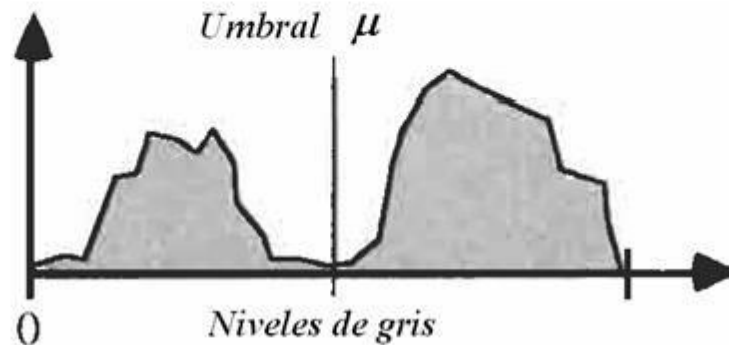
Un único umbral μ es asignado a toda la imagen.

$$I_{Bin}(m,n) = \begin{cases} 0 & \text{si } I(m,n) < \mu \\ 1 & \text{si } I(m,n) > \mu \end{cases}$$

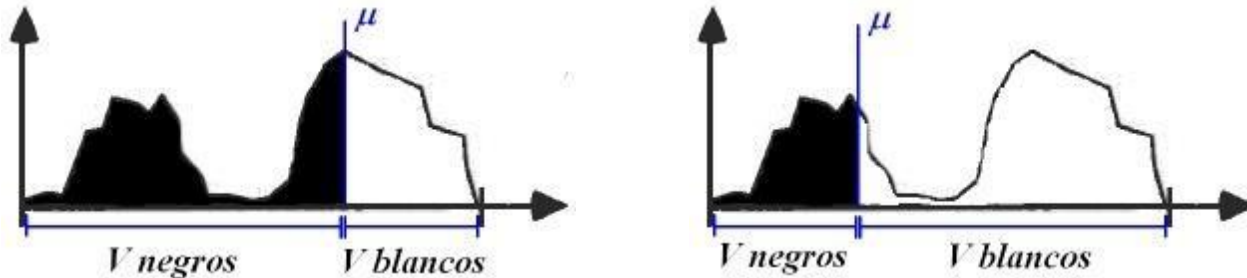


Como determinar el umbral?

El histograma puede utilizarse para diferenciar el fondo y el objeto por sus niveles de gris.



Ejemplo. El método de **Otsu** calcula el umbral μ que minimiza la varianza (**V**) **intraclase** de los píxeles blancos y negros.



ret,thresh = cv2.threshold(img, umbral, valorMax , tipo)

Los parámetros son los siguientes:

- 1.**img** es la imagen gris que va a ser analizada
- 2.**umbral** es el valor indicado a analizar en cada píxel
- 3.**valorMax** Valor que se coloca a un píxel si sobrepasa el umbral
- 4.**tipo** se elige un tipo de umbralización

La función devuelve:

- 1.**thresh** imagen binarizada
- 2.**ret** valor del umbral



Umbralización local

Se determina un umbral para cada píxel **dependiendo de su vecindario**.

Por ejemplo, el **método de Bernsen** sobre el vecindario $N_{(2w+1) \times (2w+1)}$ centrado en el píxel $I(m,n)$:

$$N_{(2w+1) \times (2w+1)} = \begin{bmatrix} I(m-w, n-w) & \cdots & I(m-w, n) & \cdots & I(m-w, n+w) \\ \vdots & & \vdots & & \vdots \\ I(m, n-w) & \cdots & I(m, n) & \cdots & I(m, n+w) \\ \vdots & & \vdots & & \vdots \\ I(m+w, n-w) & \cdots & I(m+w, n) & \cdots & I(m+w, n+w) \end{bmatrix}$$

Calcula

- Promedio

$$p(m,n) = \frac{F_{\max}(m,n) + F_{\min}(m,n)}{2}$$

- Varianza

$$v(m,n) = F_{\max}(m,n) - F_{\min}(m,n)$$

donde $F_{\max} = \max\{N_{(2w+1) \times (2w+1)}\}$ y $F_{\min} = \min\{N_{(2w+1) \times (2w+1)}\}$

La umbralización local se realiza como

$$I_{Bin}(m,n) = \begin{cases} 1, & \text{si } I(m,n) < p(m,n) \text{ y } v(m,n) > v^* \\ 0, & \text{en otro caso} \end{cases}$$

donde v^* es la varianza óptima intra-clase de Otsu.

Ejemplo.



Método de Niblack

Para cada píxel $I(m, n)$ de la imagen el umbral es determinado como

$$\mu(m, n) = \mathbf{med}(m, n) + w \cdot \mathbf{std}(m, n)$$

donde

$\mathbf{med}(m, n)$: es el promedio de los píxeles de la ventana de $(2M + 1) \times (2N + 1)$, centrada en (m, n)

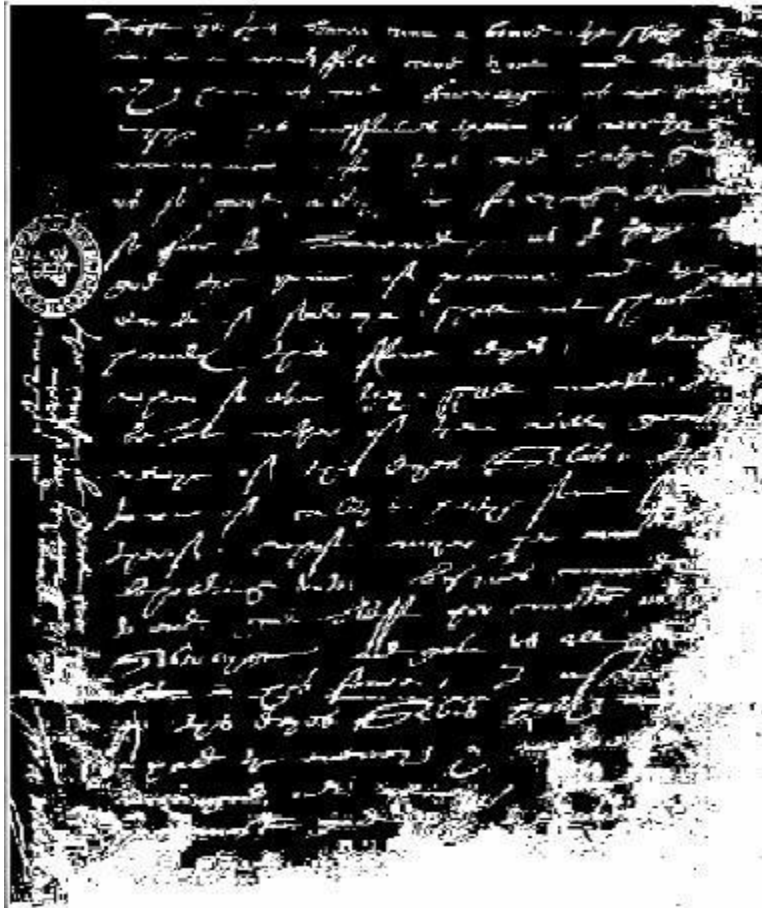
$$\mathbf{med}(m, n) = \frac{1}{MN} \sum_{i=m-M}^{m+M} \sum_{j=n-N}^{n+M} I(i, j)$$

$\mathbf{std}(m, n)$: es la desviación estándar de los píxeles de la ventana de $(2M + 1) \times (2N + 1)$, centrada en (m, n)

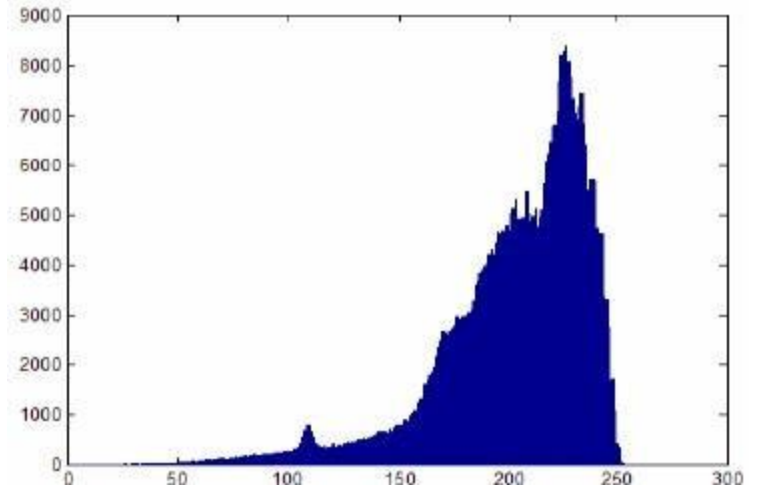
$$\mathbf{std}(m, n) = \sqrt{\frac{1}{MN} \sum_{i=m-M}^{m+M} \sum_{j=n-N}^{n+M} [I(i, j) - \mathbf{med}(i, j)]^2}$$

w : es el peso que puede configurarse según sea el caso.

Ejemplo. Umbralización difícil



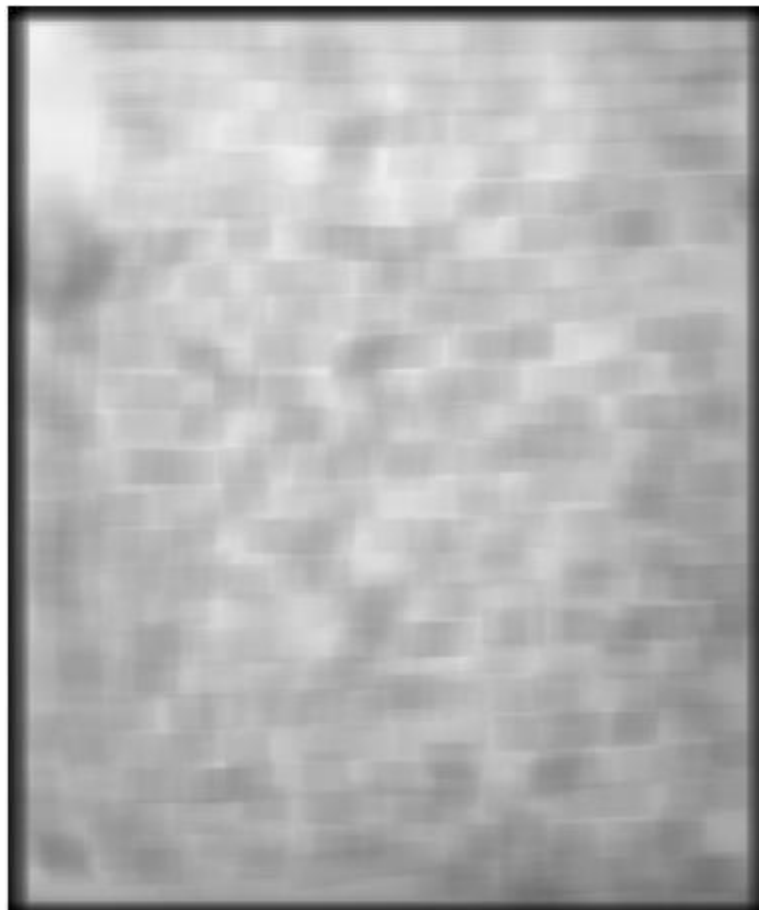
Histograma



No se sabe donde colocar
el umbral!

Método de Otsu

Umbrales de Niblack



Umbralización de Niblack



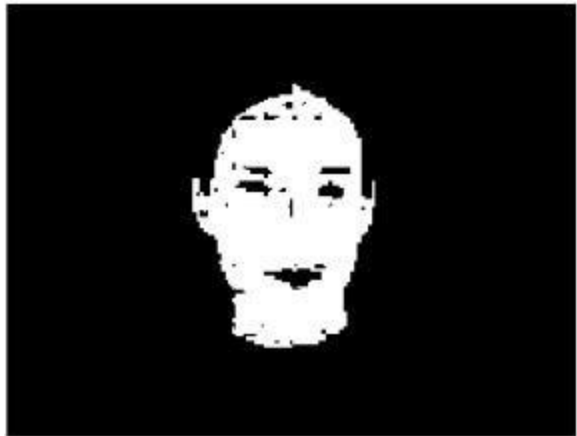
Modelo de color



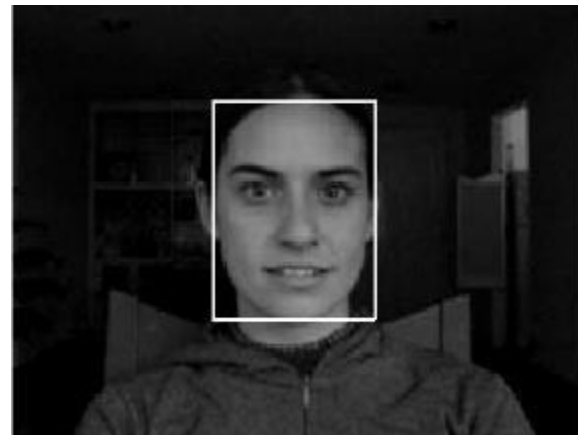
Imagen de entrada



Probabilidad de piel



Piel detectada



Rostro

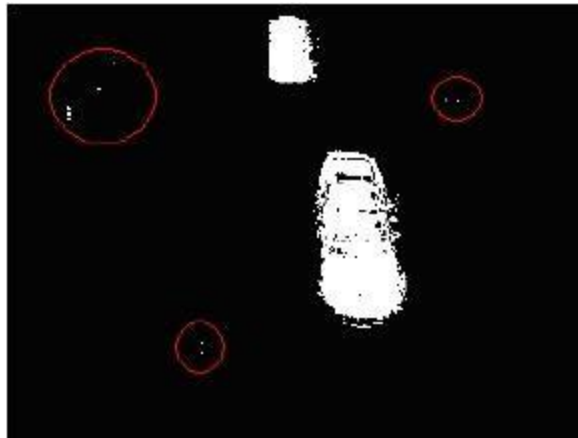
Modelo de fondo



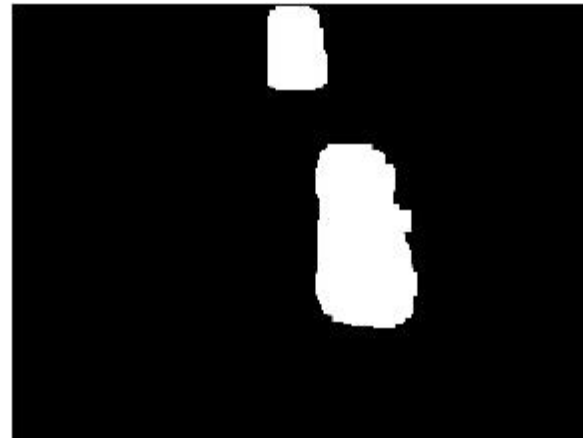
Video de entrada



Modelo de fondo



Primer plano



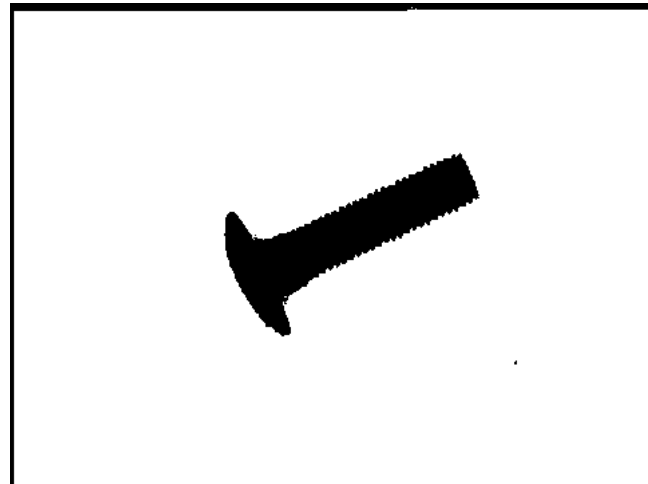
Objetos

Introducción

- **Dependencia del valor umbral:**

$$T = T(f(x, y), p(x, y), x, y)$$

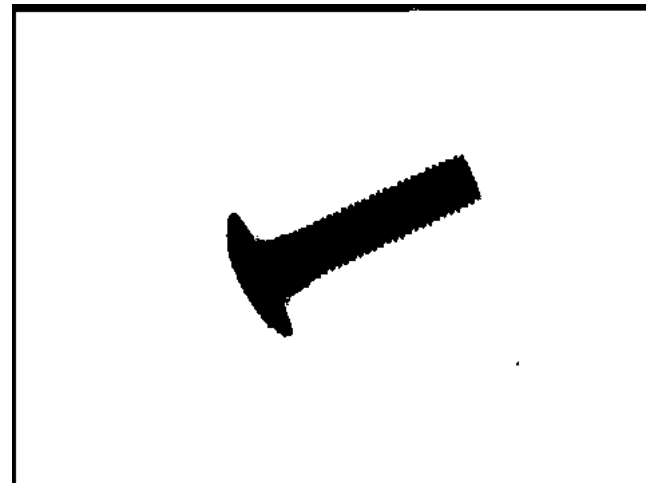
– *Global, local o dinámico.*



Introducción

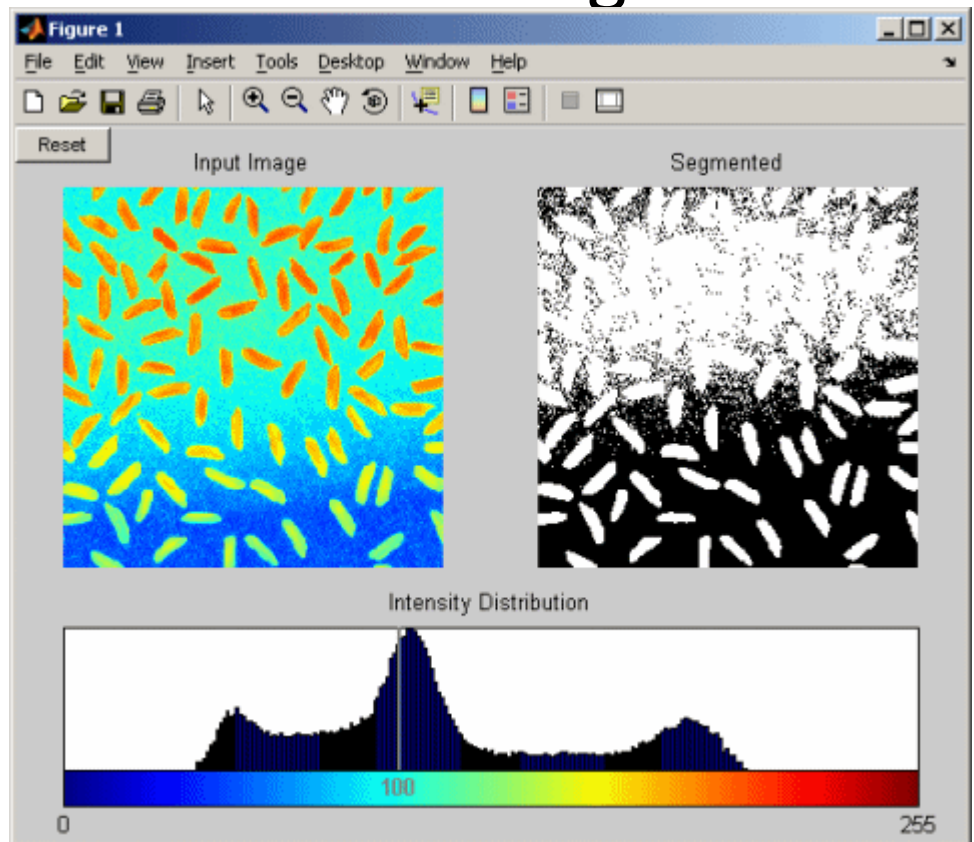
- Necesidad de definir un valor umbral T .

$$g(x, y) = \begin{cases} 1 \Leftrightarrow f(x, y) > T \\ 0 \Leftrightarrow f(x, y) \leq T \end{cases} \quad g(x, y) = \begin{cases} 1 \Leftrightarrow f(x, y) < T \\ 0 \Leftrightarrow f(x, y) \geq T \end{cases}$$



Introducción

- La mayoría de las técnicas de umbralización se basan en estadísticas sobre el histograma unidimensional.
- Para localizar umbrales es posible también usar otro tipo de procedimientos.



Introducción

- **Procedimientos paramétricos:** la distribución de los niveles de gris de una clase de objeto lleva a encontrar los umbrales.
- **Procedimientos no paramétricos:** los umbrales se obtienen de forma óptima de acuerdo a algún criterio.

Método de Otsu

- **Ventajas:**

- Buena respuesta del método frente a la mayoría en situaciones del mundo real (imágenes ruidosas, con histogramas planos, mal iluminadas...).
- Automatismo: no precisa de supervisión humana, preprocesamiento de la imagen y otro tipo de información acerca de la misma.

- **Desventajas:**

- A medida que el número de clases en la imagen aumenta, el método necesita mucho más tiempo para seleccionar un umbral multinivel adecuado.

Método de Otsu

- **Descripción:**

- Partimos de una imagen en niveles de gris con N píxels y L posibles niveles diferentes.
- Probabilidad de ocurrencia del nivel de gris i en la imagen:

$$p_i = \frac{f_i}{N}$$

$f_i \rightarrow$ Frecuencia de repetición del nivel de gris i -ésimo con $i = 1, 2, \dots, L$.

Método de Otsu

Umbralización de 2 niveles

- **Descripción:**

- En el caso particular de umbralización en dos niveles (binarización), los píxels se dividen en dos clases \rightarrow **C1** y **C2**, con niveles de gris $[1,2,...,t]$ y $[t+1,t+2,...,L]$ respectivamente, donde las distribuciones de probabilidad de ambas clases son:

$$C_1 : \frac{p_1}{\omega_1(t)}, \dots, \frac{p_t}{\omega_1(t)}$$
$$C_2 : \frac{p_{t+1}}{\omega_2(t)}, \frac{p_{t+2}}{\omega_2(t)}, \dots, \frac{p_L}{\omega_2(t)}$$

$$\omega_1(t) = \sum_{i=1}^t p_i$$

$$\omega_2(t) = \sum_{i=t+1}^L p_i$$

Método de Otsu

Umbralización de 2 niveles

- **Descripción:**

- Las medias para cada una de las clases se definen como:

$$\mu_1 = \sum_{i=1}^t \frac{i \cdot p_i}{\omega_1(t)} \quad \mu_2 = \sum_{i=t+1}^L \frac{i \cdot p_i}{\omega_2(t)}$$

- La intensidad media total de la imagen se define, siendo fácil demostrar así mismo:

$$\omega_1 \cdot \mu_1 + \omega_2 \cdot \mu_2 = \mu_T$$

$$\omega_1 + \omega_2 = 1$$

Método de Otsu

Umbralización de 2 niveles

- **Descripción:**

- Haciendo uso de un análisis discriminante, Otsu definió la varianza entre clases de una imagen umbralizada como:

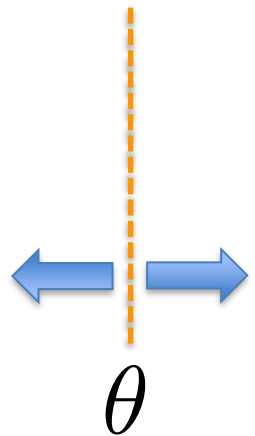
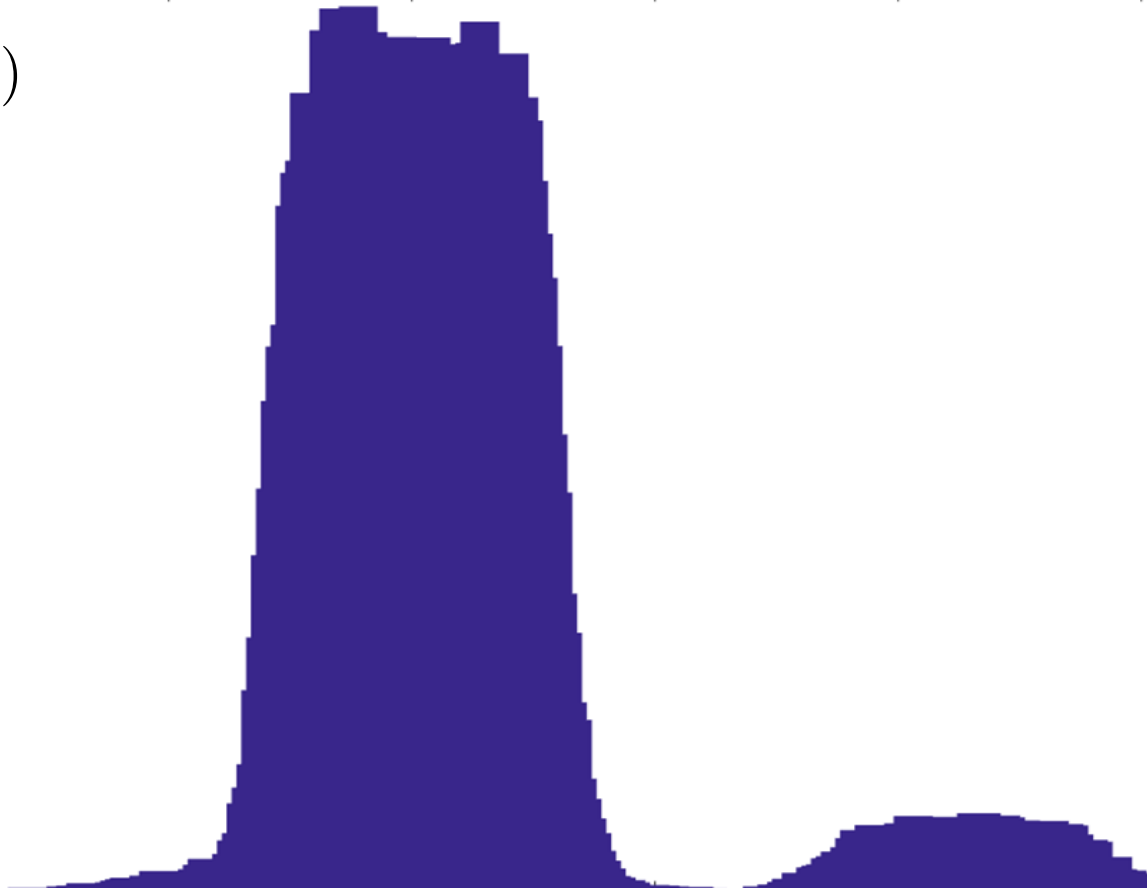
$$\sigma_B^2 = \omega_1 \cdot (\mu_1 - \mu_T)^2 + \omega_2 \cdot (\mu_2 - \mu_T)^2$$

- La idea es ahora encontrar el umbral, t , que maximice la varianza (Otsu demostró que este era el umbral óptimo):

$$t^* = \underset{t}{\text{Max}} \{ \sigma_B^2(t) \} \quad \text{Donde:} \quad 1 \leq t \leq L$$

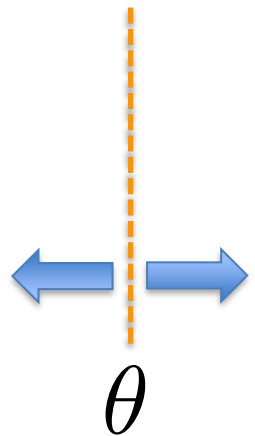
Método para estimar θ de manera automática
(Método de Otsu)

$h(\mathbf{X})$



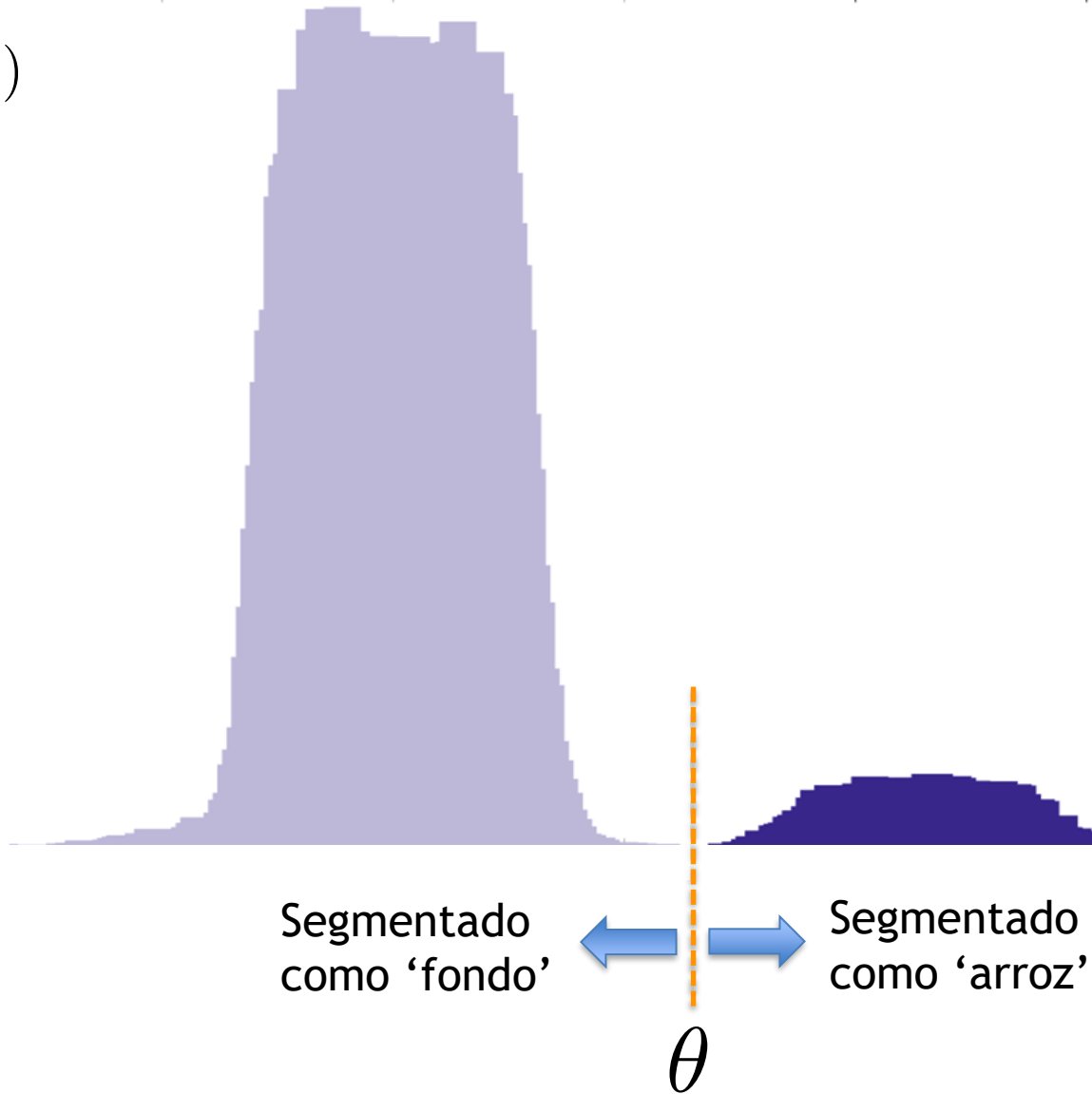
Segmentado
como 'arroz'

$h(\mathbf{X})$

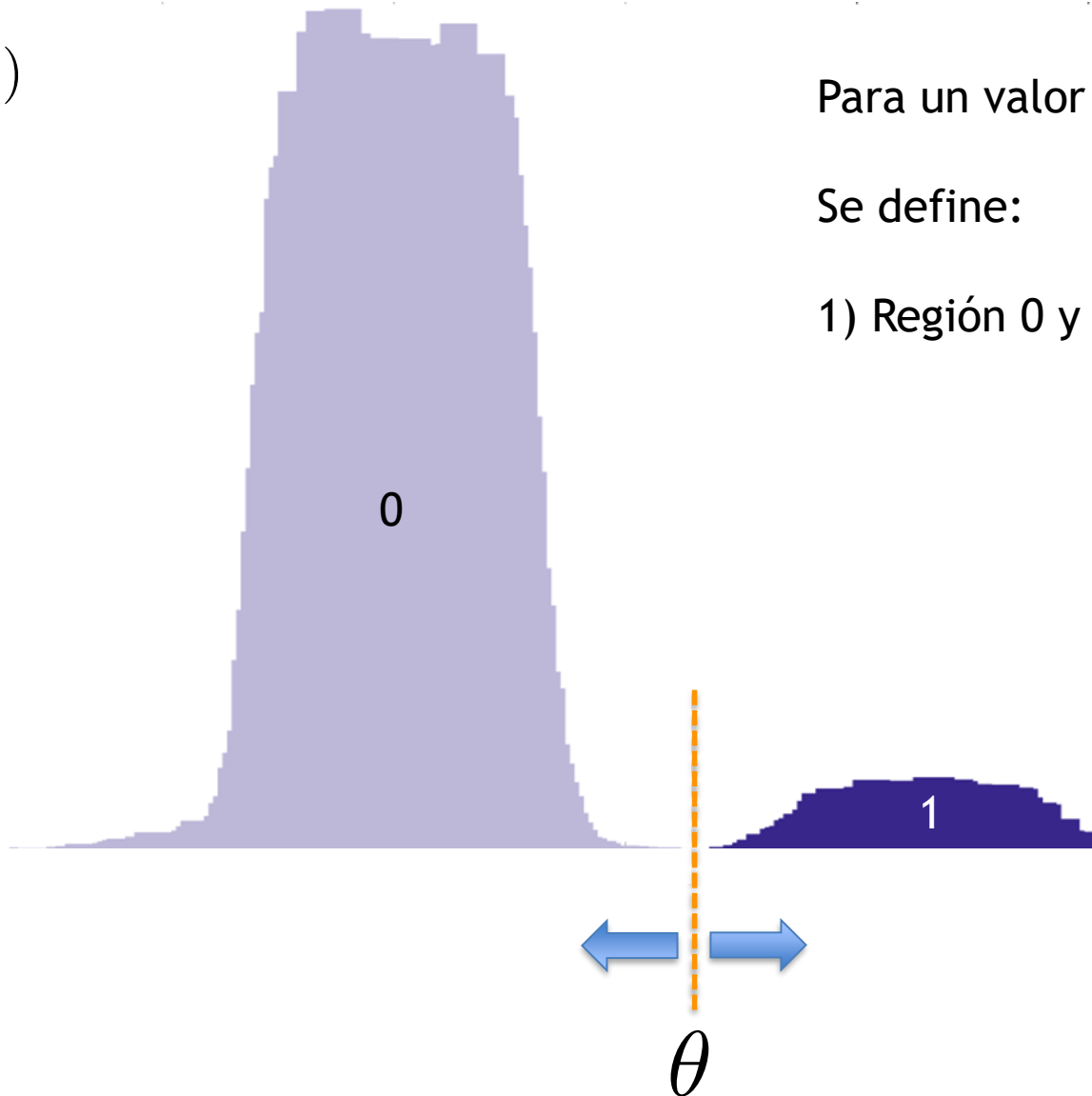


Segmentado
como 'arroz'

$h(\mathbf{X})$



$h(\mathbf{X})$

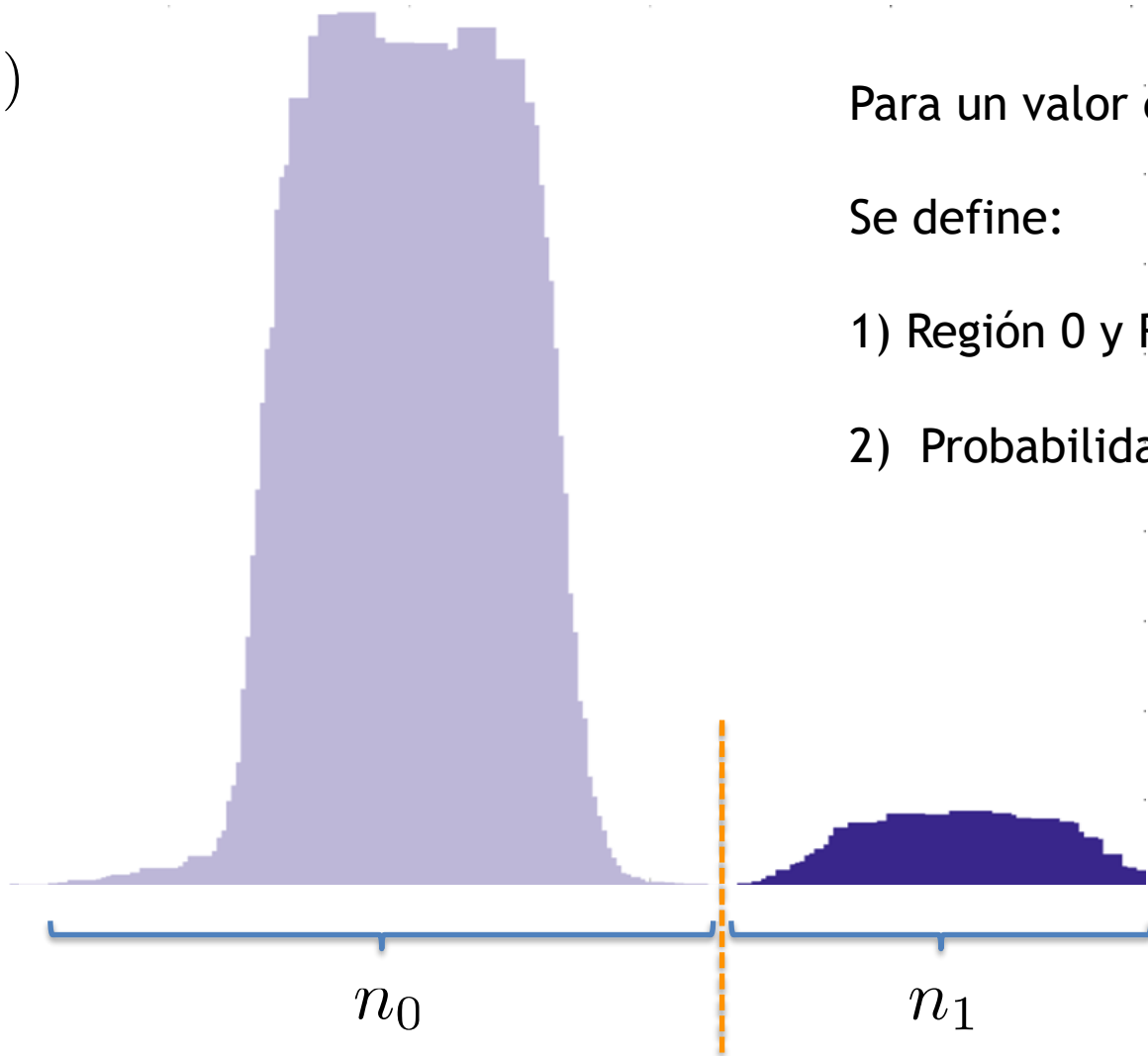


Para un valor de θ

Se define:

1) Región 0 y Región 1

$h(\mathbf{X})$



Para un valor de θ

Se define:

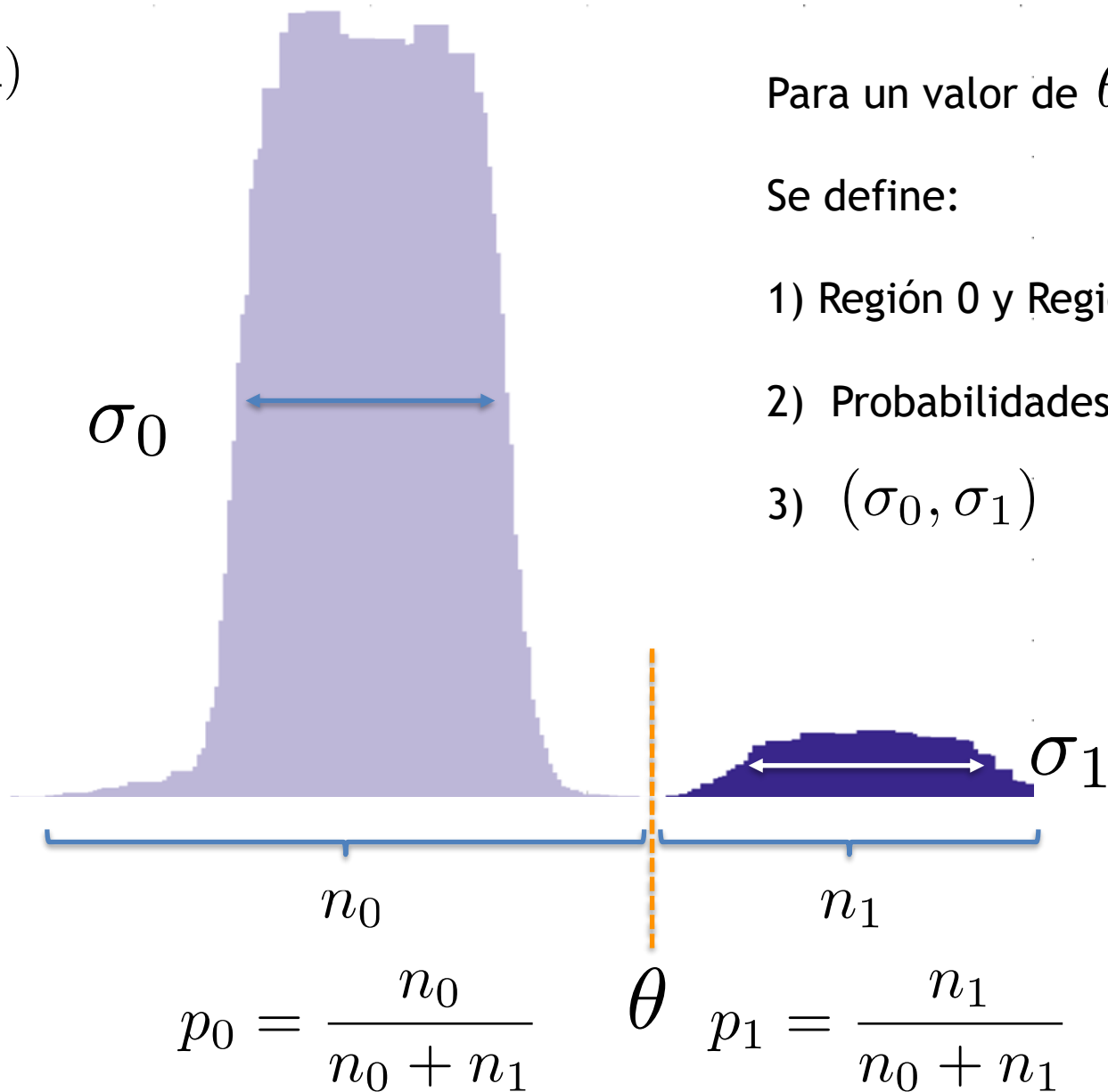
1) Región 0 y Región 1

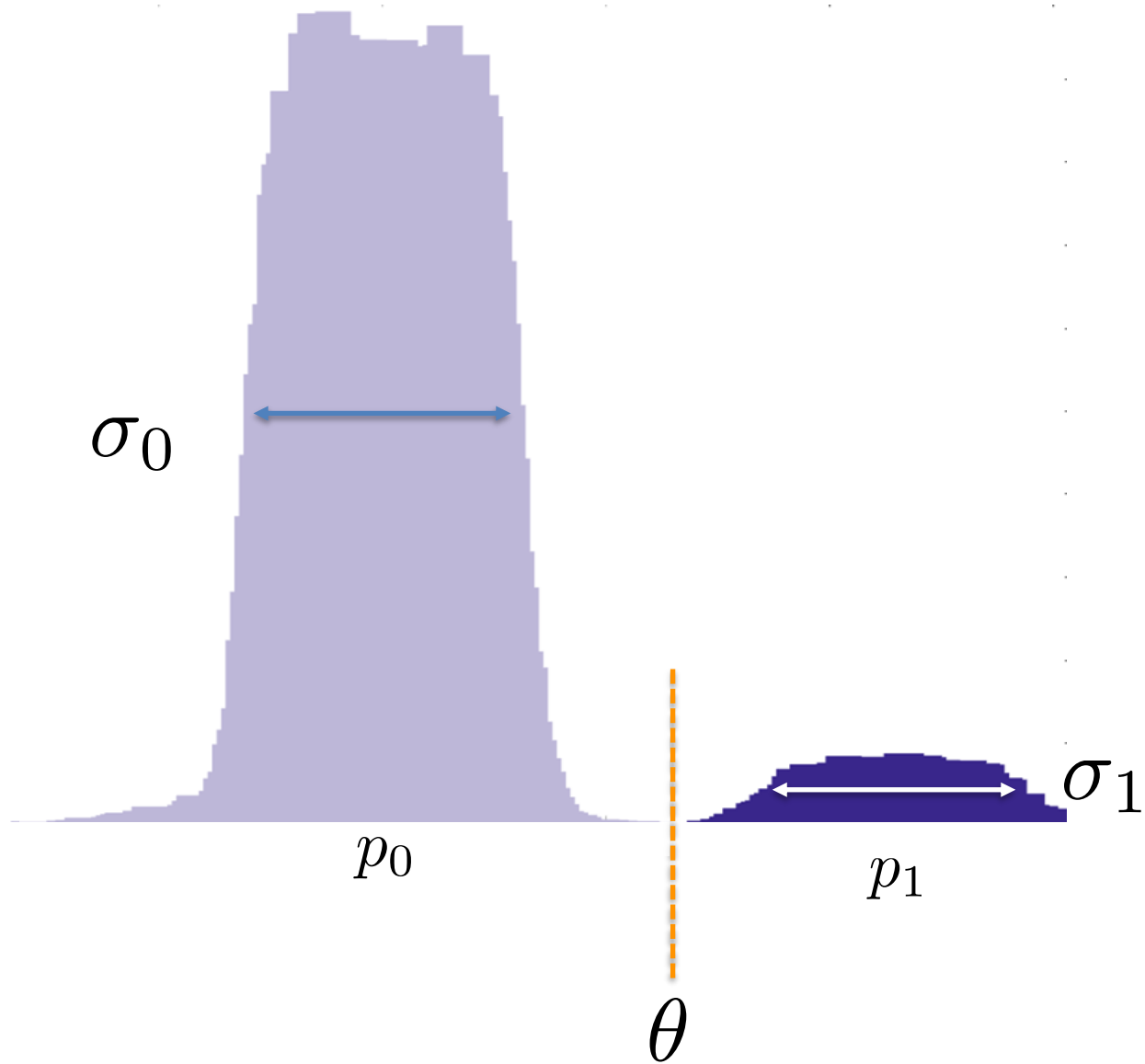
2) Probabilidades (p_0, p_1)

$$p_0 = \frac{n_0}{n_0 + n_1}$$

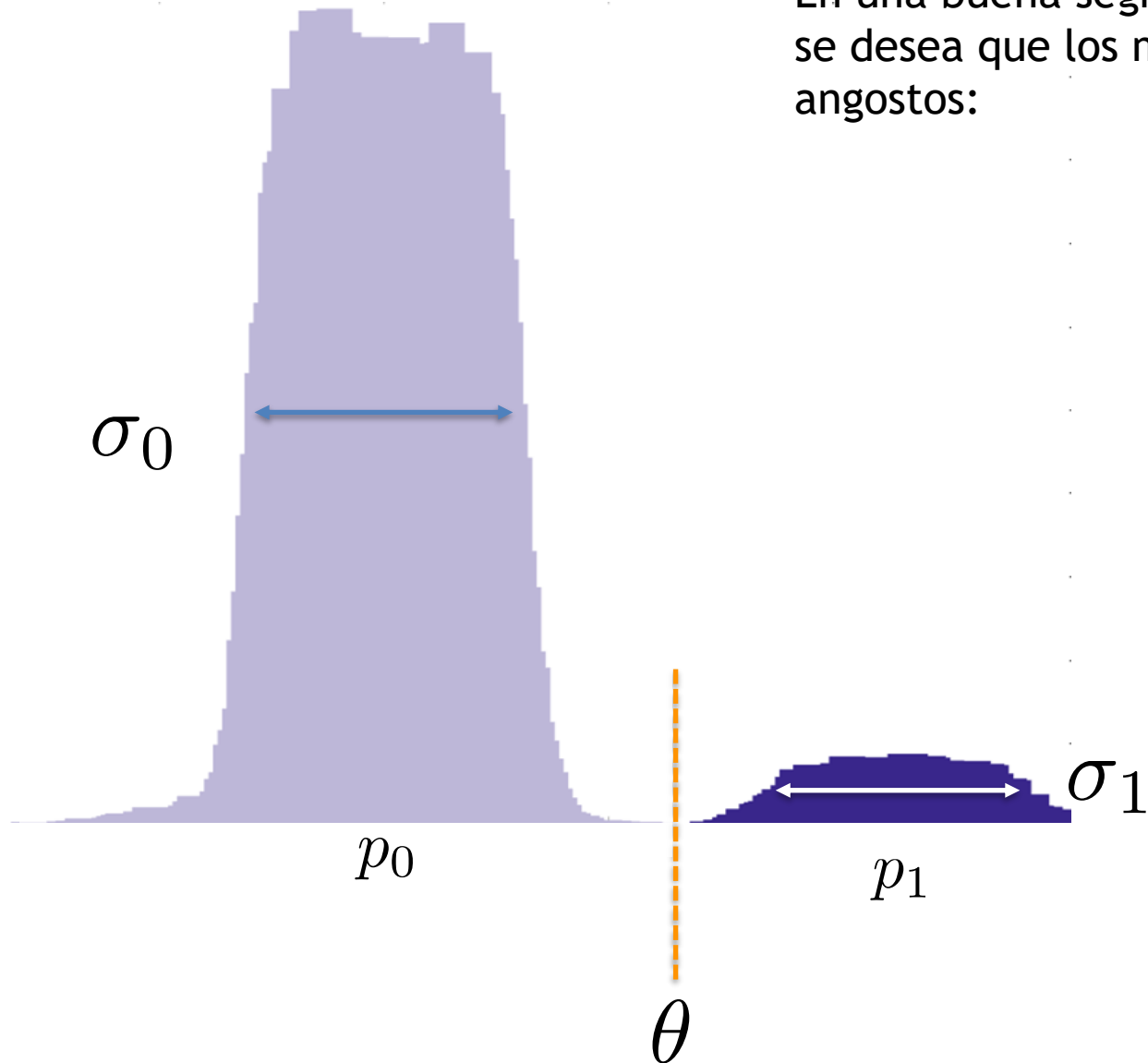
$$\theta \quad p_1 = \frac{n_1}{n_0 + n_1}$$

$h(\mathbf{X})$

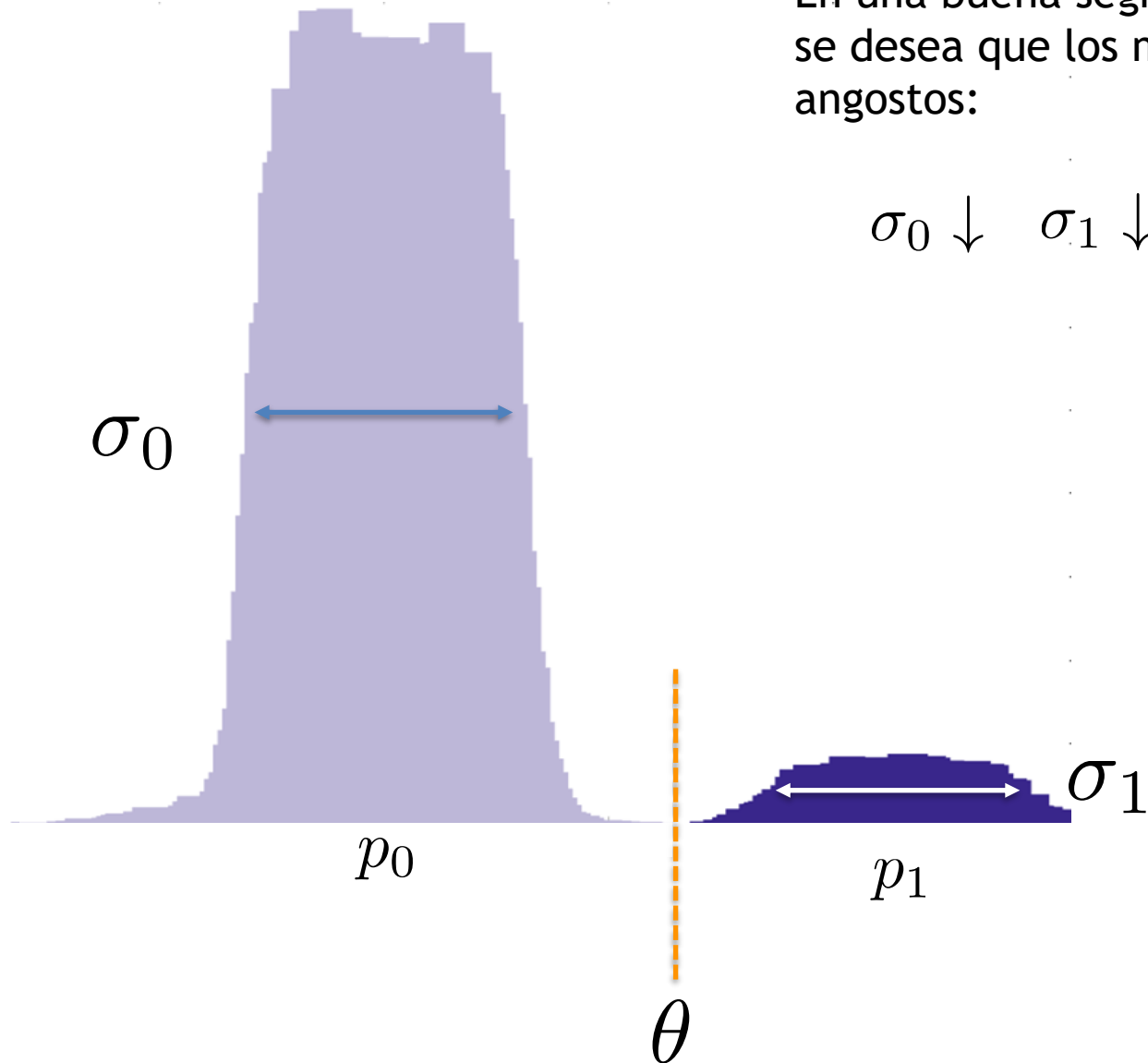




En una buena segmentación,
se desea que los modos sean
angostos:



En una buena segmentación,
se desea que los modos sean
angostos:

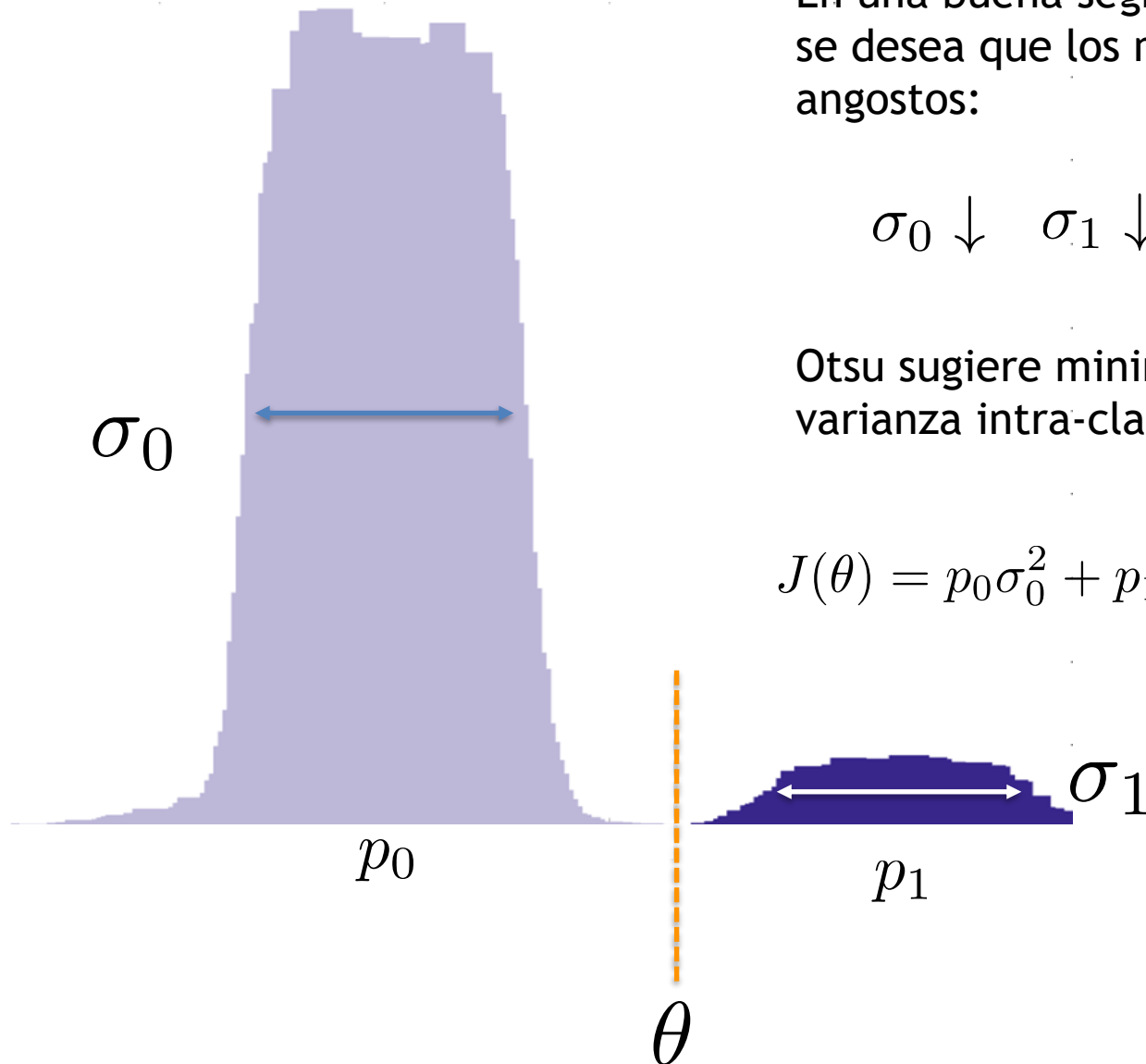


En una buena segmentación,
se desea que los modos sean
angostos:

$$\sigma_0 \downarrow \quad \sigma_1 \downarrow$$

Otsu sugiere minimizar la
varianza intra-clase:

$$J(\theta) = p_0\sigma_0^2 + p_1\sigma_1^2 \rightarrow \min$$

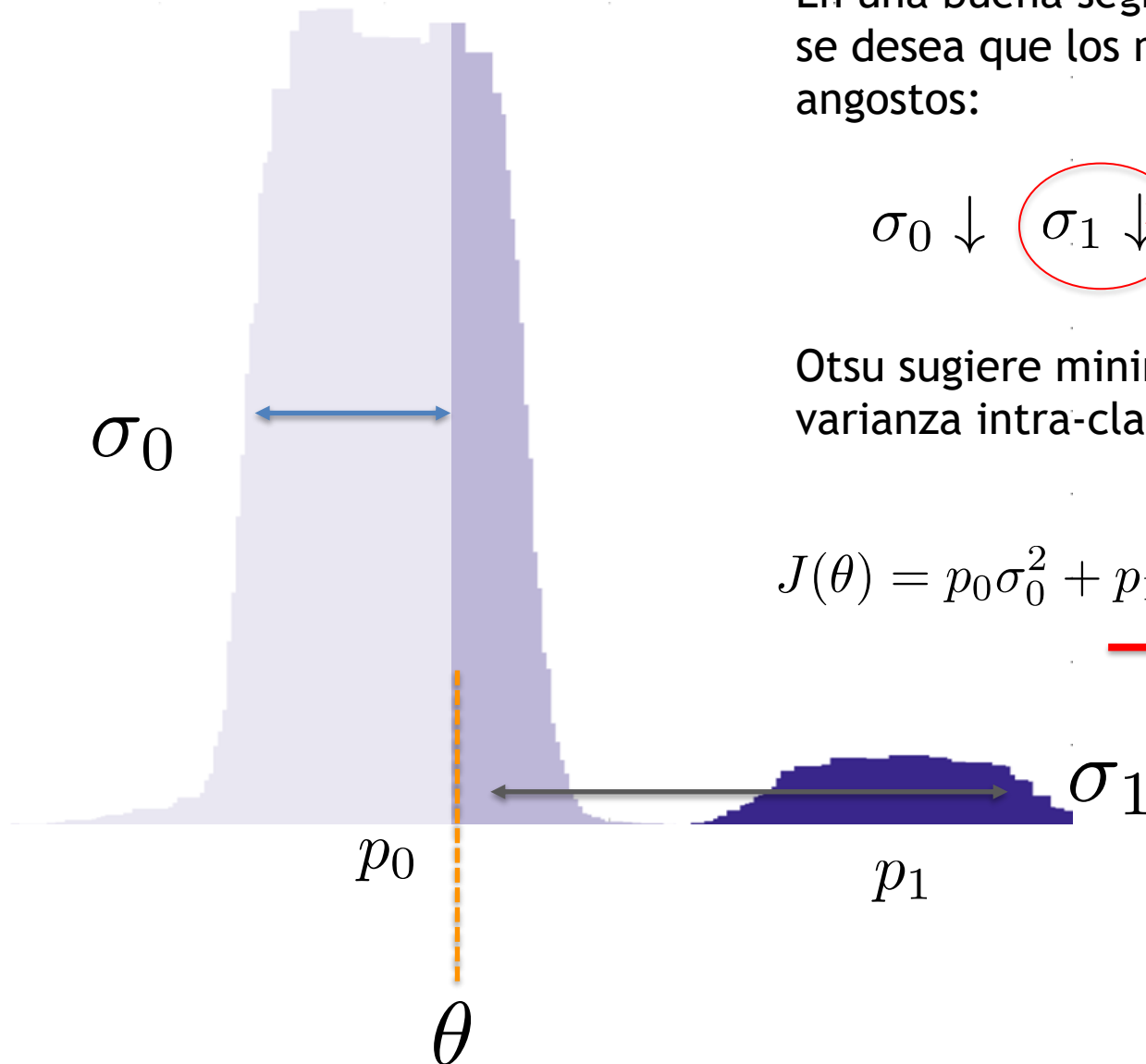


En una buena segmentación,
se desea que los modos sean
angostos:

$$\sigma_0 \downarrow \quad \sigma_1 \downarrow^{\times}$$

Otsu sugiere minimizar la
varianza intra-clase:

$$J(\theta) = p_0 \sigma_0^2 + p_1 \sigma_1^2 \rightarrow \min_{\theta}$$

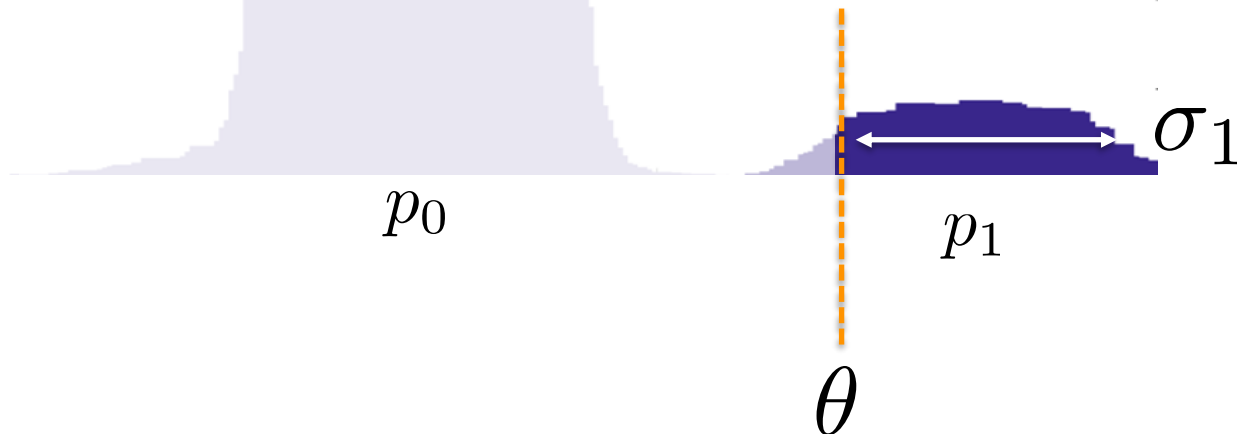


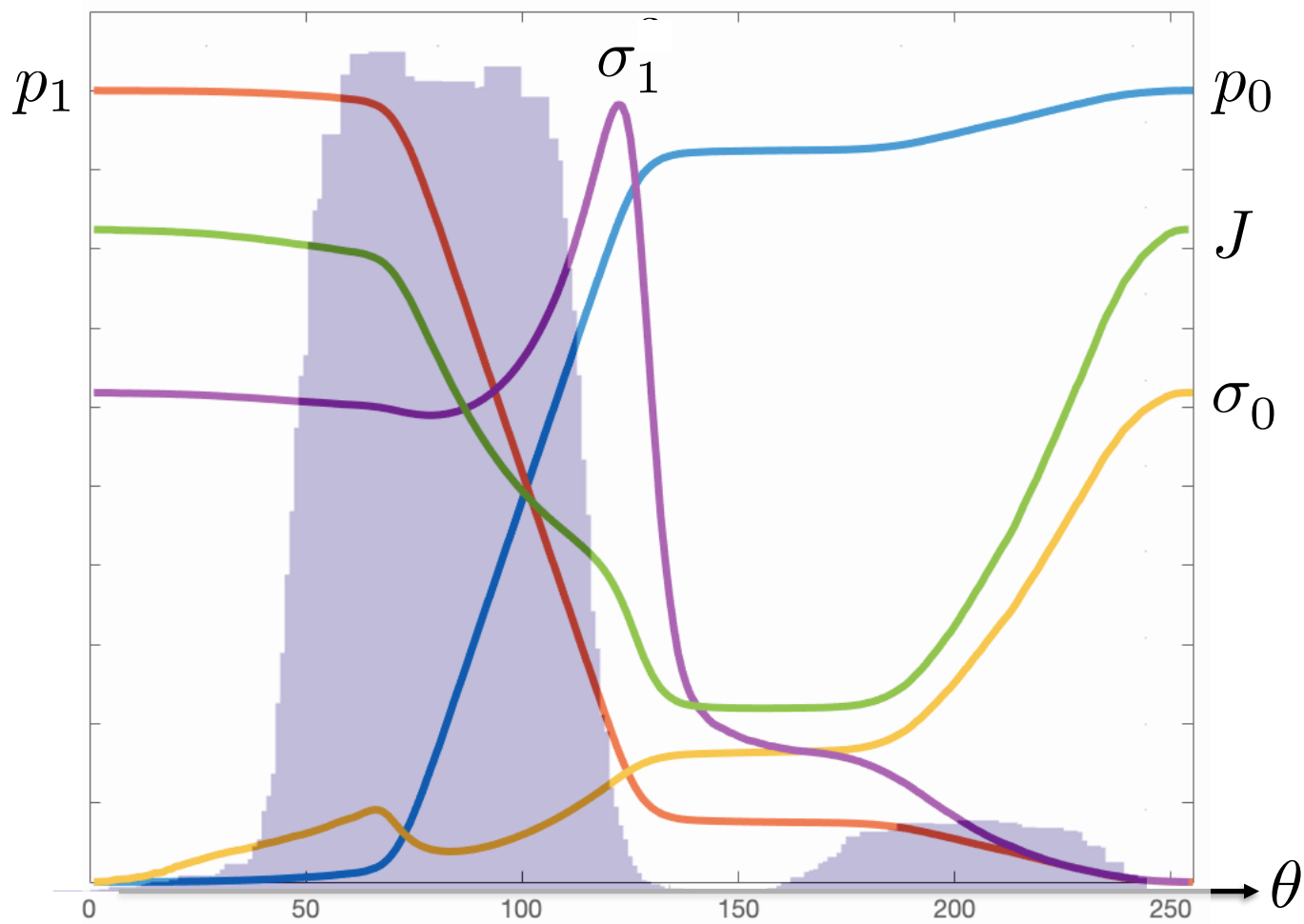
En una buena segmentación,
se desea que los modos sean
angostos:

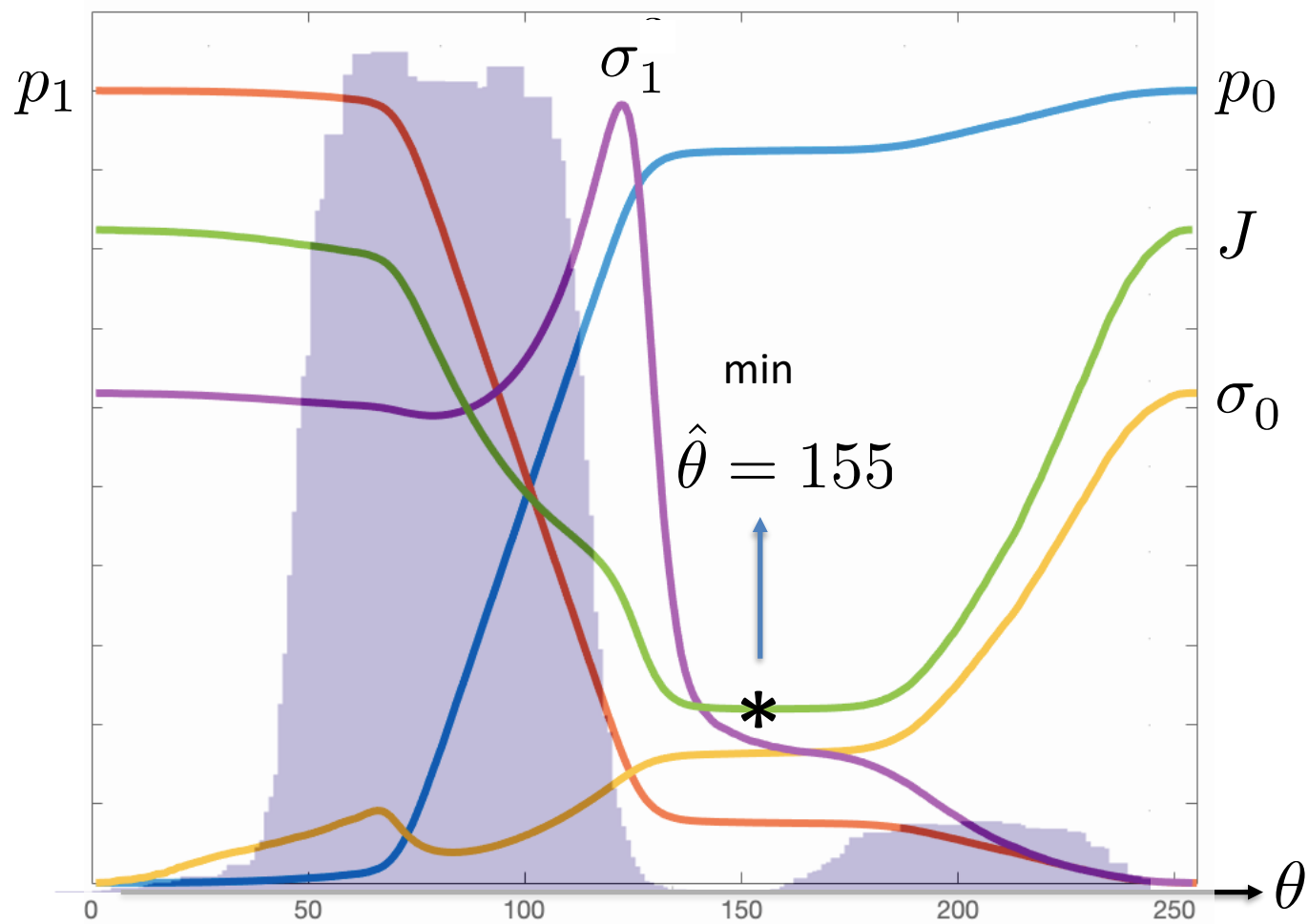
$$\sigma_0 \downarrow^{\times} \sigma_1 \downarrow$$

σ_0  Otsu sugiere minimizar la
varianza intra-clase:

$$J(\theta) = p_0 \sigma_0^2 + p_1 \sigma_1^2 \rightarrow \min_{\times}$$





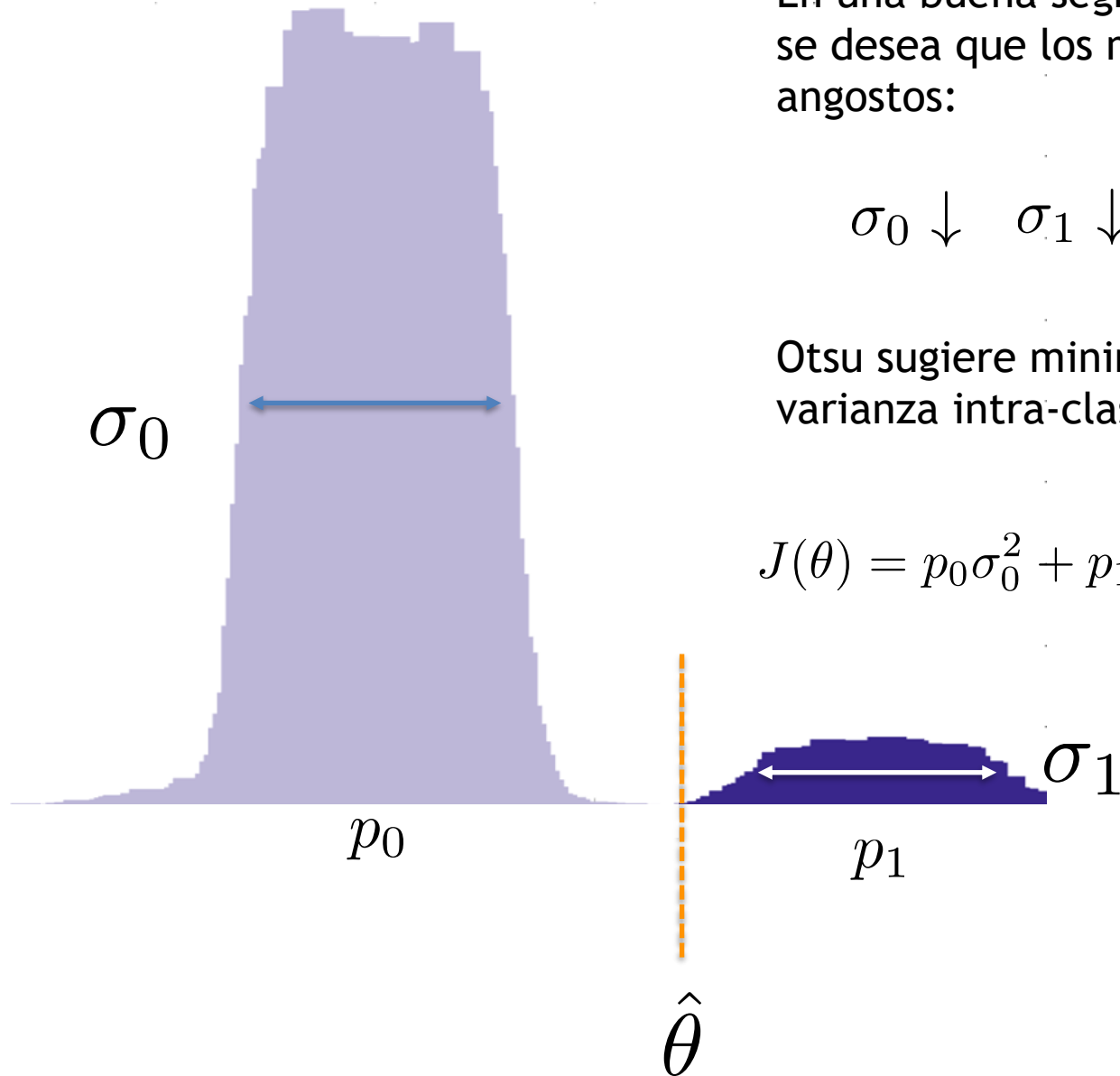


En una buena segmentación,
se desea que los modos sean
angostos:

$$\sigma_0 \downarrow \quad \sigma_1 \downarrow$$

Otsu sugiere minimizar la
varianza intra-clase:

$$J(\theta) = p_0\sigma_0^2 + p_1\sigma_1^2 \rightarrow \min$$



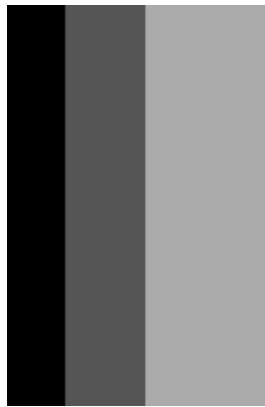
Ejemplos

Método de Otsu

Umbralización de 2 niveles

- **Ejemplo:**

- Consideremos la siguiente imagen con los siguientes parámetros que la definen.



$$L = 4 \rightarrow [0, 85, 171, 255]$$

$$f_1 = 10, f_2 = 20$$

$$f_3 = 30, f_4 = 40$$

$$N = 100 (10 \times 10)$$

- A continuación se calcula la varianza entre clases de la imagen para todo valor de umbral posible (4 en nuestro caso).

Método de Otsu

Umbralización de 2 niveles

- Ejemplo:**

— Por ejemplo, comenzamos para $t = 85$:



$$C_1 \leftrightarrow [0,85]$$

$$C_2 \leftrightarrow [171,255]$$

$$\omega_1(t = 2) = \sum_{i=1}^t p_i = \frac{1}{10} + \frac{1}{5} = \frac{3}{10}$$

$$\omega_2(t = 2) = \sum_{i=t+1}^L p_i = \frac{3}{10} + \frac{4}{10} = \frac{7}{10}$$

$$p_1 = \frac{f_1}{N} = \frac{10}{100} = \frac{1}{10}$$

$$p_2 = \frac{f_2}{N} = \frac{20}{100} = \frac{1}{5}$$

$$p_3 = \frac{f_3}{N} = \frac{30}{100} = \frac{3}{10}$$

$$p_4 = \frac{f_4}{N} = \frac{40}{100} = \frac{4}{10}$$

Método de Otsu

Umbralización de 2 niveles

- **Ejemplo:**

— Por ejemplo, comenzamos para $t = 85$:



$$C_1: \frac{p_1}{\omega_1(t=2)}, \frac{p_2}{\omega_1(t=2)} = \left\{ \frac{1}{3}, \frac{2}{3} \right\}$$

$$C_2: \frac{p_3}{\omega_2(t=2)}, \frac{p_4}{\omega_2(t=2)} = \left\{ \frac{3}{7}, \frac{4}{7} \right\}$$

$$\mu_1 = \sum_{i=1}^t \frac{ip_i}{\omega_1(t=2)} = \frac{1}{3} + 2 \times \frac{2}{3} = \frac{5}{3}$$

$$\mu_2 = \sum_{i=t+1}^L \frac{ip_i}{\omega_2(t=2)} = 3 \times \frac{3}{7} + 4 \times \frac{4}{7} = \frac{25}{7}$$

Método de Otsu

Umbralización de 2 niveles

- **Ejemplo:**

- Por ejemplo, comenzamos para $t = 85$:



$$\mu_T = \omega_1(t = 2)\mu_1 + \omega_2(t = 2)\mu_2 = 3$$

$$\sigma_B^2 = \omega_1(t = 2)(\mu_1 - \mu_T)^2 + \omega_2(t = 2)(\mu_2 - \mu_T)^2 \approx 0.7619$$

- Resultando para el resto de umbrales:

$$\sigma_B^2(t = 1) = 0.4444$$

$$\sigma_B^2(t = 3) = 0.6667$$

Método de Otsu

Umbralización de 2 niveles

- **Ejemplo:**
 - En consecuencia, el umbral óptimo según Otsu para este caso sería $t = 85$, resultando la imagen umbralizada con este valor (la mayor varianza entre clases se obtiene con dicho umbral):



Método de Otsu

Umbralización de M niveles

- Generalización:**

- En este caso, al existir M clases, existirán $M-1$ umbrales distintos, generalizando el caso particular anteriormente descrito. Por tanto, en este caso habremos de obtener el conjunto multinivel que maximice la varianza entre clases de la forma:

$$\{t_1^*, t_2^*, \dots, t_{M-1}^*\} = \underset{t_1, t_2, \dots, t_{M-1}}{\text{Max}} \{ \sigma_B^2(t_1, t_2, \dots, t_{M-1}) \}$$

$$1 \leq t_1 < \dots < t_{M-1} < L$$

Donde:

$$\omega_k = \sum_{i \in C_k} p_i \quad \mu_k = \sum_{i \in C_k} \frac{i \cdot p_i}{\omega_k}$$

$$\sigma_B^2 = \sum_{k=1}^M \omega_k \cdot (\mu_k - \mu_T)^2$$

Método de Otsu

Ejemplo

- **Ejemplo (2 niveles):**
 - Binarización mediante umbral subjetivo $t = 45$.



Método de Otsu

Ejemplo

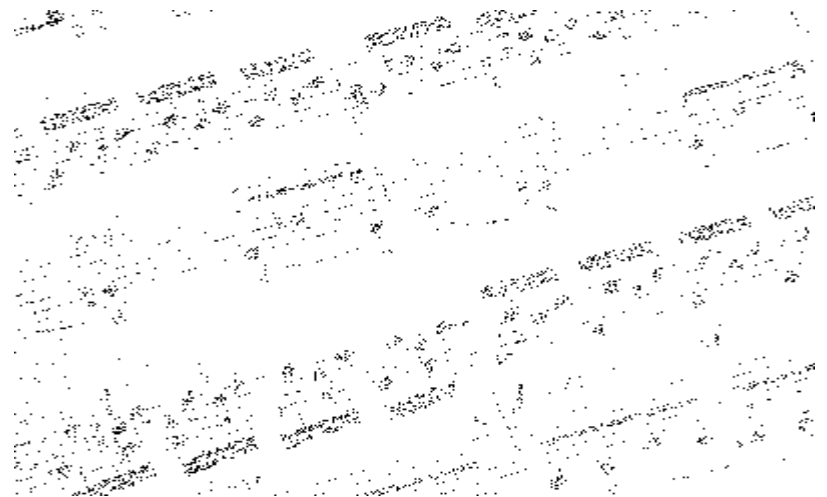
- **Ejemplo (2 niveles):**
 - Binarización mediante umbral óptimo según el método de Otsu de $t = 79$.



Método de Otsu

Ejemplo

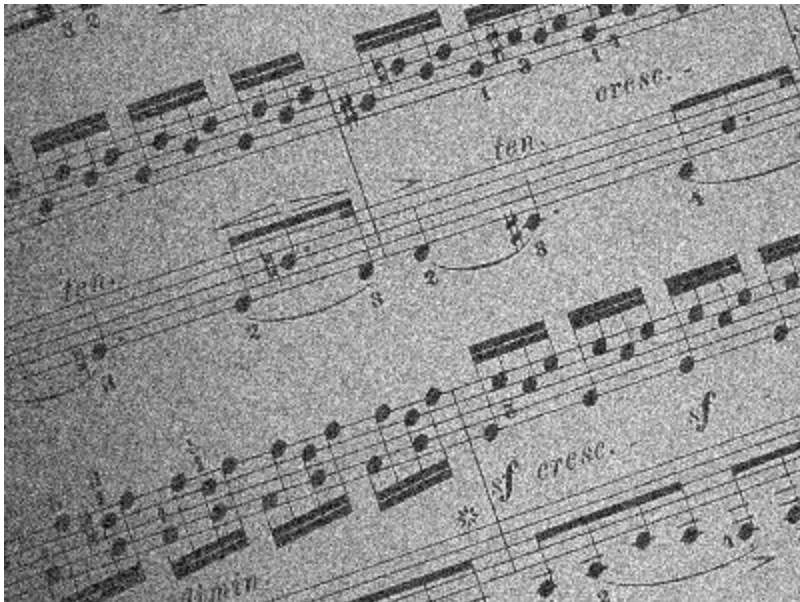
- **Ejemplo (2 niveles):**
 - Repetimos la binarización con el mismo umbral subjetivo habiendo añadido ruido blanco gaussiano a la imagen original con una densidad de 0.2:



Método de Otsu

Ejemplo

- **Ejemplo (2 niveles):**
 - Ahora el umbral óptimo de Otsu para la imagen con ruido blanco gaussiano es de $t = 133$:



Bibliografía relacionada

- **Nobuyuki Otsu, “A threshold selection method from gray-level histogram”, *IEEE Transactions on System Man Cybernetics*, Vol. SMC-9, No. 1, 1979.**
- ***Digital Image Processing Second Edition*, Rafael C. González – Richard E. Woods, capítulo 10.**