

# Extracción de Características

**Domingo Mery**

Departamento de Ciencia de la Computación  
Universidad Católica de Chile  
Santiago de Chile  
6 de abril de 2006

**Departamento de Ciencia de la Computación  
Universidad Católica de Chile**

**Domingo Mery:**

Av. Vicuña Mackenna 4860(143)

eMail: [dmery@ieee.org](mailto:dmery@ieee.org)

<http://www.ing.puc.cl/~dmery>

# Índice general

<b>Indice General</b>	<b>4</b>
<b>1. Introducción</b>	<b>5</b>
<b>2. Características geométricas</b>	<b>7</b>
2.1. Altura y ancho ( $h$ y $w$ ) . . . . .	7
2.2. Área ( $A$ ) . . . . .	7
2.3. Perímetro ( $L$ ) . . . . .	8
2.4. Redondez ( $R$ ) . . . . .	8
2.5. Momentos . . . . .	8
2.6. Descriptores de Fourier . . . . .	10
2.7. Elipse . . . . .	11
2.8. Diámetro $d$ . . . . .	13
2.9. Distancias al borde . . . . .	13
<b>3. Características Cromáticas</b>	<b>15</b>
3.1. Color promedio . . . . .	15
3.2. Gradiente promedio en el borde . . . . .	16
3.3. Promedio de la segunda derivada . . . . .	16
3.4. Contraste . . . . .	17
3.5. Momentos . . . . .	18
3.6. Características de textura . . . . .	19



# Capítulo 1

## Introducción

En este capítulo se describen algunas características que pueden ser empleadas para describir cuantitativamente regiones que hayan sido segmentadas en una imagen. Se entenderá por *región* aquel conjunto de píxels que pertenezcan a una misma zona de la imagen y que esté limitado por bordes. Se asumirá que los bordes no pertenecen a la región.

Para explicar las características que se detallarán a continuación se usará el ejemplo de la Figura 1.1. En este ejemplo se presenta una región circular que ha sido segmentada. La región entonces se conforma por los píxels que pertenecen al círculo (pero no a su perímetro), es decir los píxels que han sido marcados con color gris en la Figura 1.1b. Los bordes de la región definen el límite de la región.

Las características que se pueden extraer de una región se dividen en dos categorías: *características geométricas* y *características cromáticas*.

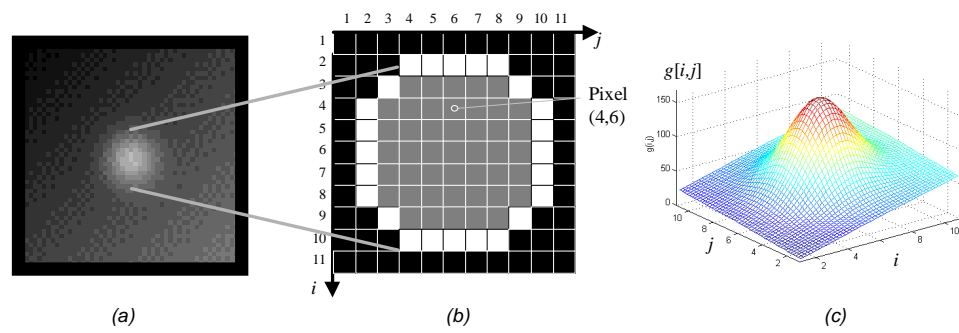


Figura 1.1: Ejemplo de una región: a) Imagen. b) Región segmentada. c) Representación 3D de los valores de gris de la región y su entorno.

## Capítulo 2

# Características geométricas

Las características geométricas contienen información de forma, posición, tamaño y orientación de la región.

A continuación se enumeran algunas características geométricas que se usan comúnmente en el reconocimiento de patrones.

### 2.1. Altura y ancho ( $h$ y $w$ )

La altura y el ancho de una región se definen como:

$$h = i_{max} - i_{min} + 1 \quad \text{y} \quad w = j_{max} - j_{min} + 1 \quad (2.1)$$

donde  $i_{max}$  e  $i_{min}$  representan el valor máximo y mínimo que toma la coordenada  $i$  en la región (ver Figura 1.1), y lo mismo es válido para  $j_{max}$  y  $j_{min}$ . En el ejemplo mostrado  $h = w = 7$  píxels.

### 2.2. Área ( $A$ )

El área de una región se define como el número de los píxels de la región. En el ejemplo  $A = 45$  píxels.

### 2.3. Perímetro ( $L$ )

El perímetro de una región puede ser definido de varias maneras. Una definición práctica, mas no exacta, es tomar el perímetro como el número de píxels que pertenecen al borde de la región<sup>1</sup>. En el ejemplo de la Figura 1.1b,  $L$  es el número de píxels marcados en color blanco, es decir  $L = 24$ .

### 2.4. Redondez ( $R$ )

Esta característica que indica la calidad de redondo de una región es una medida de su forma. La redondez se define como [5]:

$$R = \frac{4 \cdot A \cdot \pi}{L^2} \quad (2.2)$$

La redondez  $R$  de una región estará entre los valores 0 y 1. Teóricamente  $R = 1$  para un círculo (perfecto); y  $R = 0$  para una región que tenga altura y/o ancho igual a cero. En la práctica sin embargo, debido al muestreo en el espacio de la región estos valores presentan desviaciones como se puede ver en la región circular de nuestro ejemplo. En este caso  $R = 4 \cdot 45 \cdot \pi / 24^2 = 0,98$ .

### 2.5. Momentos

Los momentos estadísticos se definen como

$$m_{rs} = \sum_{i,j \in \mathfrak{R}} i^r j^s \quad \text{para } r, s \in \mathcal{N} \quad (2.3)$$

donde  $\mathfrak{R}$  es el conjunto de píxels de la región. En el ejemplo de la Figura 1.1b el pixel cuyas coordenadas son  $(i = 4, j = 6)$  pertenece a este conjunto. El parámetro  $r + s$  denota el orden del momento. El momento de orden cero

---

<sup>1</sup>Otras definiciones más exactas consideran el factor  $\sqrt{2}$  para píxels del borde de la región que estén en diagonal, como por ejemplo en el caso de un borde que contenga los píxels  $(i, j)$  y  $(i + 1, j + 1)$  [1]. En este caso existe un compromiso entre la precisión y el costo computacional que requiere su cálculo.



$m_{00}$  corresponde al área de la región  $A$ . El centro de gravedad de una región queda definido por:

$$\bar{i} = \frac{m_{10}}{m_{00}} \quad \bar{j} = \frac{m_{01}}{m_{00}} \quad (2.4)$$

Con ayuda de las coordenadas del centro de gravedad se definen los momentos centrales que son invariantes al desplazamiento de la región en la imagen.

$$\mu_{rs} = \sum_{i,j \in \mathfrak{R}} (i - \bar{i})^r (j - \bar{j})^s \quad \text{para } r, s \in \mathcal{N} \quad (2.5)$$

Muy conocidos en la teoría de reconocimiento de patrones son las características derivadas de los momentos centrales, denominados momentos de Hu [6, 14]:

$$\begin{aligned} \phi_1 &= \eta_{20} + \eta_{02} \\ \phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\ \phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\ \phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\ \phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\ &\quad (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ \phi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + \\ &\quad 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ \phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - \\ &\quad (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (2.6)$$

con

$$\eta_{rs} = \frac{\mu_{rs}}{\mu_{00}^t} \quad t = \frac{r+s}{2} + 1.$$

Los momentos de Hu son invariantes a la traslación, rotación y escalamiento. Esto quiere decir que dos regiones que tengan la misma forma pero que sean de distinto tamaño y que estén ubicados en posiciones y orientaciones distintas en la imagen tendrán momentos de Hu iguales.

A veces sin embargo, es necesario contar con características que además sean invariantes a las transformadas afines. Un conjunto alternativo de características que son invariantes a la traslación, rotación, escalamiento y también a transformaciones afines se puede derivar de los momentos de segundo y tercer orden [13]:

$$I_1 = \frac{\mu_{20}\mu_{02} - \mu_{11}^2}{\mu_{00}^4}$$

$$\begin{aligned}
I_2 &= \frac{\mu_{30}^2 \mu_{03}^2 - 6\mu_{30}\mu_{21}\mu_{12}\mu_{03} + 4\mu_{30}\mu_{12}^3 + 4\mu_{21}^3\mu_{03} - 3\mu_{21}^2\mu_{12}^2}{\mu_{00}^{10}} \\
I_3 &= \frac{\mu_{20}(\mu_{21}\mu_{03} - \mu_{12}^2) - \mu_{11}(\mu_{30}\mu_{03} - \mu_{21}\mu_{12}) + \mu_{02}(\mu_{30}\mu_{12} - \mu_{21}^2)}{\mu_{00}^7} \quad (2.7) \\
I_4 &= \frac{(\mu_{20}^3\mu_{03}^2 - 6\mu_{20}^2\mu_{11}\mu_{12}\mu_{03} - 6\mu_{20}^2\mu_{02}\mu_{21}\mu_{03} + 9\mu_{20}^2\mu_{02}\mu_{12}^2 \\
&\quad 12\mu_{20}\mu_{11}^2\mu_{21}\mu_{03} + 6\mu_{20}\mu_{11}\mu_{02}\mu_{30}\mu_{03} - 18\mu_{20}\mu_{11}\mu_{02}\mu_{21}\mu_{12} \\
&\quad - 8\mu_{11}^3\mu_{30}\mu_{03} - 6\mu_{20}\mu_{02}^2\mu_{30}\mu_{12} + 9\mu_{20}\mu_{02}^2\mu_{21} \\
&\quad + 12\mu_{11}^2\mu_{02}\mu_{30}\mu_{12} - 6\mu_{11}\mu_{02}^2\mu_{30}\mu_{21} + \mu_{02}^3\mu_{30}^2)/\mu_{00}^{11}}{\mu_{00}^{11}}
\end{aligned}$$

## 2.6. Descriptores de Fourier

Una buena caracterización de la forma de una región se logra utilizando los *descriptores de Fourier* [2, 12, 15]. Las coordenadas  $(i_k, j_k)$  de los píxeles del borde, para  $k = 0, \dots, L - 1$  de una región se agrupan en un sentido de giro conformando un número complejo  $(i_k + j \cdot j_k)$  con  $j = \sqrt{-1}$ , donde  $L$  es el perímetro de la región definido como el número de píxeles del borde de la región. La línea continua formada por estas coordenadas corresponden a una señal periódica que puede ser transformada al dominio de Fourier por medio de la Transformada Discreta de Fourier (DFT) [1]:

$$F_n = \sum_{k=0}^{L-1} (i_k + j \cdot j_k) e^{-j \frac{2\pi kn}{L}} \quad \text{para } n = 0, \dots, L - 1. \quad (2.8)$$

Los descriptores de Fourier corresponden al módulo de los coeficientes complejos de Fourier. Como se puede apreciar los descriptores de Fourier son invariantes a la rotación de la región. El primer descriptor de Fourier  $|F_0|$  da información de la ubicación de la región en la imagen. Los descriptores que son invariantes a la posición de la región son los siguientes descriptores. La fase de los coeficientes de Fourier proporciona información acerca de la orientación y de la simetría de las regiones.

En la Figura 2.1 se muestran los descriptores de Fourier para el ejemplo de la Figura 1.1. En este ejemplo el píxel de partida es  $(i_0, j_0) = (6, 10)$ . En el caso de un círculo ideal los descriptores serían  $|F_n| = 0$  para  $1 < n < L$ , ya que la representación de las coordenadas  $(i_k, j_k)$  corresponderían a una senoide perfecta.

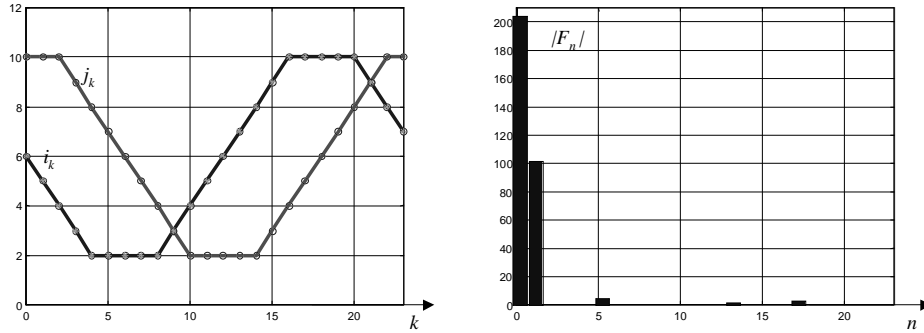


Figura 2.1: Coordenadas del borde de la región de la Figura 1.1 y sus descriptores de Fourier.

## 2.7. Elipse

Características muy usadas son aquellas que se derivan de una elipse que se aproxima al contorno de la región. Es decir, se busca una elipse que sea lo más parecida posible al borde de la región y luego se extraen características de la elipse tales como eje mayor, eje menor, orientación y posición de la elipse. A continuación se explica una metodología para extraer estas características.

Los píxeles del borde de una región son definidos como  $(x_i, y_i)$  para  $i = 1, \dots, L$ . Es sabido que la ecuación de una elipse está dada por:

$$ax^2 + bxy + cy^2 + dx + ey + f = 0 \quad (2.9)$$

que podría ser escrita como  $\mathbf{a}^T \mathbf{x} = 0$ , donde  $\mathbf{a} = [a \ b \ c \ d \ e \ f]^T$  es un vector con los parámetros de la elipse y  $\mathbf{x} = [x^2 \ xy \ y^2 \ x \ y \ 1]^T$  es un vector que incluye las coordenadas de un punto  $(x, y)$  de la elipse.

Si la región fuese elíptica, para cada punto  $(x_i, y_i)$  se cumpliría  $\mathbf{a}^T \mathbf{x}_i = 0$  con  $\mathbf{x}_i = [x_i^2 \ x_i y_i \ y_i^2 \ x_i \ y_i \ 1]^T$ . En la realidad las regiones no son elípticas, no sólo porque las regiones reales tienen formas distintas a las ideales, sino que además hay un error en la cuantificación espacial de la posición de los píxeles, por lo tanto se busca un vector  $\mathbf{a}$  tal que haga que  $\mathbf{a}^T \mathbf{x}_i$  sea cercano a cero. Esto se logra minimizando:

$$\| \mathbf{Xa} \| \rightarrow \text{mín} \quad (2.10)$$

donde  $\mathbf{X}$  es una matriz de  $L$  filas cuya fila  $i$  es  $\mathbf{x}_i$ . Una solución típica a (2.10) se consigue restringiendo  $\|\mathbf{a}\| = 1$ . En este caso  $\mathbf{a}$  es la última columna de  $\mathbf{V}$ , donde  $\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^T$  es la descomposición de valor singular<sup>2</sup> de  $\mathbf{X}$ .

Las características mencionadas de la elipse se pueden obtener re-escribiendo (2.9) de la siguiente manera:

$$\left(\frac{x - x_0}{a_e}\right)^2 + \left(\frac{y - y_0}{b_e}\right)^2 = 1 \quad (2.11)$$

donde

$$a_e = \frac{1}{\sqrt{s} a_p}, \quad b_e = \frac{1}{\sqrt{s} b_p} \quad (2.12)$$

con

$$s = \frac{1}{v - f}$$

$$v = \mathbf{t}^T \mathbf{T} \mathbf{t}$$

$$\mathbf{T} = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}$$

$$\mathbf{t} = \frac{1}{2} \mathbf{T}^{-1} \begin{bmatrix} d \\ e \end{bmatrix}$$

$$\begin{aligned} a_p &= a \cos^2(\alpha) + b \cos(\alpha) \sin(\alpha) + c \sin^2(\alpha) \\ b_p &= a \sin^2(\alpha) - b \cos(\alpha) \sin(\alpha) + c \cos^2(\alpha) \end{aligned}$$

y

$$\alpha = \frac{1}{2} \arctan\left(\frac{b}{a - c}\right) \quad (2.13)$$

$$x_0 = t_1, y_0 = t_2 \quad (2.14)$$

---

<sup>2</sup>SVD: *singular decomposition value*. En este caso  $\mathbf{S}$  es una matriz diagonal de la misma dimensión que  $\mathbf{X}$  con elementos no negativos ordenados en forma descendente, y las matrices  $\mathbf{U}$  y  $\mathbf{V}$  son matrices unitarias tale que  $\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^T$  [10].

Los ejes de la elipse quedan definidos como  $a_e$  y  $b_e$ , el centro de la origen está en  $(x_0, y_0)$  y la orientación de la elipse es  $\alpha$ .

A partir de los ejes se puede definir la excentricidad como:

$$e_x = \frac{\min(a_e, b_e)}{\max(a_e, b_e)} \quad (2.15)$$

Para regiones circulares, al igual que la redondez (ver Sección 2.4), la excentricidad es uno, y para regiones más ovaladas la excentricidad disminuye su valor:  $0 \leq e_x \leq 1$ .

## 2.8. Diámetro $d$

Una forma de hacer referencia al tamaño de una región es midiendo su *diámetro*, que se define como la distancia máxima entre dos píxeles de la región, esto se traduce a la distancia máxima entre dos píxeles del borde.

## 2.9. Distancias al borde

Otras medidas que sirven para caracterizar la forma de una región son:

- $D_{max}$ : distancia máxima del centro de masa al borde.
- $D_{min}$ : distancia mínima del centro de masa al borde.
- $D_{mean}$ : distancia promedio del centro de masa al borde.

Como las anteriores características no son invariantes al cambio de escala se definen las métricas:

$$\frac{D_{max}}{D_{min}}, \frac{D_{max}}{D_{mean}} \text{ y } \frac{D_{min}}{D_{mean}}. \quad (2.16)$$

Una métrica adicional es  $\beta$  definida como la distancia promedio al borde:

$$\beta = \frac{1}{A} \sum_{i,j \in \mathcal{R}} d_{i,j} \quad (2.17)$$

donde  $A$  es el área de la región (número de píxeles de la región), y  $d_{i,j}$  es la distancia mínima del píxel  $(i, j)$  al borde de la región. Una forma parecida de expresar  $\beta$  es mediante una característica invariante a la escala denominada complejidad de forma:

$$f = \frac{A}{\beta^2} \quad (2.18)$$

## Capítulo 3

# Características Cromáticas

Antes de entrar a definir las características del color es necesario saber si la imagen que se pretende analizar es a color o en blanco y negro. En el primer caso el color se descompone en tres componentes (rojo, verde y azul) para cada píxel de la imagen, en el segundo caso se cuenta sólo con el tono de gris en cada píxel. Las características que se mencionan a continuación son para una sola variable de color. Esta variable puede ser cada una de las componentes del color, una combinación lineal de las tres componentes o bien simplemente el tono de gris. La información entonces necesaria para calcular estas características es el valor de esta variable de color en cada píxel que es representada como  $x[i, j]$  para el píxel  $(i, j)$  de la imagen. Es posible que en algunas aplicaciones sea necesario analizar de manera independiente dos variables de color, por ejemplo la componente en rojo y la componente en azul. En este tipo de aplicaciones será necesario extraer las características de color para cada una de las variables de color requeridas.

### 3.1. Color promedio

( $G$ ): Esta característica es el promedio de la variable de color que se define como:

$$G = \frac{1}{A} \sum_{i,j \in \mathfrak{R}} x[i, j] \quad (3.1)$$

donde  $\mathfrak{R}$  denota el conjunto de píxels de la región y  $x[i, j]$  el valor de la variable de color en el píxel  $(i, j)$ .

El número de píxels de la región  $\mathfrak{R}$  es  $A$ , el área de la región. Una representación 3D de la variable de color de una región y su entorno se muestra en la Figura 1.1c. En este caso se trabaja con el valor de gris, ya que la imagen es blanco y negro. Para este ejemplo el promedio es  $G = 121, 90$  ( $G=0$  significa 100 % negro y  $G=255$  corresponde a 100 % blanco).

### 3.2. Gradiente promedio en el borde

( $C$ ): Esta característica toma el valor promedio del gradiente de la variable de color en el borde de la región. Con esta característica se puede medir qué tan abrupto es el cambio en la coloración en la región con respecto a su entorno. El gradiente promedio en el borde se calcula como:

$$C = \frac{1}{L} \sum_{i,j \in \ell} x'[i, j] \quad (3.2)$$

donde  $x'[i, j]$  es el módulo del gradiente de la variable de color del píxel  $(i, j)$ . Los píxels a evaluar pertenecen exclusivamente al borde. Estos píxels conforman el conjunto  $\ell$ . El número de píxels del conjunto  $\ell$  es  $L$ , el perímetro de la región. El gradiente puede ser calculado utilizando el operador de gradiente de Gauss [1], en este caso para el ejemplo de la Figura 1.1  $C = 35, 47$ .

### 3.3. Promedio de la segunda derivada

( $D$ ): Esta característica se calcula como el promedio de la segunda derivada de la variable de color en la región:

$$D = \frac{1}{A} \sum_{i,j \in \mathfrak{R}} x''[i, j] \quad (3.3)$$

donde  $x''[i, j]$  denota el módulo de la segunda derivada de la variable de color en el píxel  $(i, j)$ , y  $\mathfrak{R}$  el conjunto de píxels que pertenecen a la región. Para calcular la segunda derivada se puede utilizar el operado LoG (*Laplacian-of-Gauss*) [9, 1]. Es necesario observar que  $D < 0$  significa que la región es más clara que su entorno (es decir que su variable de color es mayor en la región que fuera de ella). Así mismo  $D > 0$  indica una región más oscura que su entorno.



### 3.4. Contraste

El contraste de una región es concebido como una medida para la diferencia de color entre la región y su entorno. ‘Región’ y ‘entorno’ no tienen píxeles en común, y conforman una ‘zona’, que puede ser definida como un rectángulo: La zona entonces queda definida como la ventana

$$g[i, j] = x[i + i_r, j + j_r] \quad (3.4)$$

para  $i = 1, \dots, 2h+1$  y  $j = 1, \dots, 2w+1$ , donde  $h$  y  $w$  representan la altura y el ancho de la región respectivamente (ver ecuación (2.1)). Los puntos centrales de estas zonas se definen como  $i_r = \bar{i} - h - 1$  y  $j_r = \bar{j} - b - 1$ , donde  $(\bar{i}, \bar{j})$  corresponde al centro de gravedad de la región (ver ecuación (2.4)).

Entre más pequeña sea la diferencia de la variable de color en la región con respecto a su entorno, más pequeño será el contraste. Para visualizar el contraste se pueden representar la variable de color de una zona como una función 3D, donde el eje  $x$  y el eje  $y$  representan el eje  $i$  y el eje  $j$  de la imagen, y el eje  $z$  el valor de la variable de color que toma el pixel correspondiente, es decir  $g[i, j]$ . La Figura 1.1c muestra esta representación para el ejemplo de la Figura 1.1a. Se reconoce en este ejemplo una región de alto contraste.

El contraste se define matemáticamente de diversas formas. Una definición comúnmente usada es utilizando características de textura [1], que serán explicadas posteriormente. Otras definiciones de contraste [8, 13] se dan a continuación:

$$K_1 = \frac{G - G_e}{G_e}, \quad K_2 = \frac{G - G_e}{G + G_e} \quad \text{y} \quad K_3 = \ln(G/G_e), \quad (3.5)$$

donde  $G$  y  $G_e$  representan el promedio de la variable de color en la región y en el entorno respectivamente.

Una nueva forma de calcular el contraste se muestra en [11]. El cálculo de esta característica se obtiene en tres pasos:

*i)* Extracción del color en los ejes principales de la zona: se calculan dos funciones de color  $P_1$  y  $P_2$ . La primera función  $P_1$  toma los valores de la variable de color en la dirección  $i$  y la segunda función  $P_2$  en la dirección  $j$ . Ambas funciones se centran en los centros de gravedad. En el ejemplo de la Figura 1.1b el centro de gravedad está en el píxel (6,6), esto quiere decir que

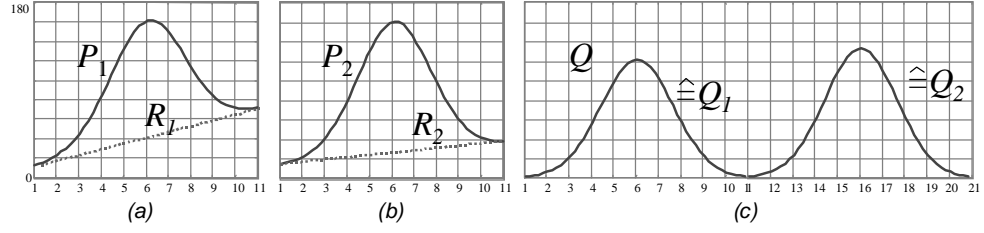


Figura 3.1: Cálculo del contraste para la Figura 1.1: a) valor de gris de la zona y rampa en dirección  $i$ ; b) valor de gris de la zona y rampa en dirección  $j$ ; c) fusión de las funciones de a) y b) sin rampas.

$P_1$  y  $P_2$  son los valores del tono de gris de la zona de la columna 6 y de la fila 6 respectivamente, tal como se muestra en la Figura 3.1a-b.

ii) Aislamiento de la región: Para aislar la región de su entorno se trata de eliminar el fondo de la región, que se modela como una función lineal de primer orden, es decir una rampa. Se asume entonces que los valores extremos de  $P_1$  y  $P_2$  pertenecen las rampas  $R_1$  y  $R_2$ , tal como se ilustra en la Figura 3.1a-b. Las rampas serán substraídas de las funciones originales para conformar  $Q_1 = P_1 - R_1$  y  $Q_2 = P_2 - R_2$  que se fusionan en la nueva función  $Q$  como se muestra en la Figura 3.1c.

iii) Cálculo del contraste: a partir de la nueva función  $Q$  se definen dos nuevos contrastes:

$$K_\sigma = \sigma_Q \quad \text{y} \quad K = \ln(Q_{\max} - Q_{\min}) \quad (3.6)$$

donde  $\sigma_Q$ ,  $Q_{\max}$  y  $Q_{\min}$  representan la desviación estándar, el máximo y el mínimo de  $Q$  respectivamente.

### 3.5. Momentos

Los momentos definidos en la Sección 2 pueden ser utilizados en el análisis de color de la región si en la ecuación del cálculo de los momentos (2.3) se incorpora la información de la variable de color:

$$m'_{rs} = \sum_{i,j \in \mathbb{R}} i^r j^s x[i, j] \quad \text{para } r, s \in N. \quad (3.7)$$

La sumatoria se calcula sobre los pixels  $(i, j)$  de la región. De esta manera se pueden calcular las características indicadas en (2.6) y (2.7) que incorporen la información del color según (3.7), el resultado será el conjunto de características  $\phi'_1 \dots \phi'_7$  e  $I'_1 \dots I'_4$ .

### 3.6. Características de textura

Las características de textura proporcionan información sobre la distribución espacial del color en la imagen. Para el análisis de regiones se pueden aplicar las características de textura no a la imagen entera sino sólo a las zonas (región y entorno) como se definió en (3.4).

Una característica simple de textura es la varianza local [7] definida como:

$$\sigma_g^2 = \frac{1}{4hb + 2h + 2b} \sum_{i=1}^{2h+1} \sum_{j=1}^{2b+1} (g[i, j] - \bar{g})^2 \quad (3.8)$$

donde  $\bar{g}$  denota el valor promedio de la variable de color en la zona.

Otras características de textura se obtienen por medio de la *matriz de coocurrencia*<sup>1</sup>. La matriz de coocurrencia se denotará como  $\mathbf{P}_{kl}$ , donde el elemento  $P_{kl}[i, j]$  otorga el valor de frecuencia (divido por  $N_T$ ) de ocurrencia de los valores de color  $i$  y  $j$  en dos píxels ubicados en una posición relativa dada por el vector  $(k, l)$ . La variable  $N_T$  significa el número de píxels que fueron necesarios para calcular  $\mathbf{P}_{kl}$ , con esto se normaliza la matriz de coocurrencia ya que la suma de todos sus elementos es uno.

Si la variable de color tiene una resolución de 256, por ejemplo de 0 a 255, el tamaño de la matriz de coocurrencia  $\mathbf{P}_{kl}$  será  $256 \times 256$ . Ya que esto implica un costo computacional muy alto, es común que se utilicen matrices más pequeñas empleando sólo los bits más significativos de la variable de color [1]. A manera de ejemplo, se puede tener una matriz de coocurrencia de  $8 \times 8$  agrupando el valor de la variable de color  $x$  en  $[0, \dots, 31]$ ,  $[32, \dots, 63]$ , ...  $[224, \dots, 255]$ .

Algunas características de textura para imágenes (o zonas) cuyas matrices de coocurrencia sean de  $N_x \times N_x$  elementos se presentan a continuación [1, 13, 4, 3]:

---

<sup>1</sup>del inglés co-occurrence matrix.

Entropía:

$$H_{kl} = \sum_{i=1}^{N_x} \sum_{j=1}^{N_x} P_{kl}[i, j] \log(P_{kl}[i, j]) \quad (3.9)$$

Inercia o contraste:

$$I_{kl} = \sum_{i=1}^{N_x} \sum_{j=1}^{N_x} (i - j)^2 P_{kl}[i, j] \quad (3.10)$$

Homogenidad o energía:

$$E_{kl} = \sum_{i=1}^{N_x} \sum_{j=1}^{N_x} [P_{kl}[i, j]]^2 \quad (3.11)$$

Momento de diferencia inverso:

$$Z_{kl} = \sum_{i=1}^{N_x} \sum_{j=1}^{N_x} \frac{P_{kl}[i, j]}{1 + (i - j)^2} \quad (3.12)$$

# Bibliografía

- [1] K.R. Castleman. *Digital image processing*. Prentice-Hall, Englewood Cliffs, New Jersey, 1996.
- [2] R. Chellappa and R. Bagdazian. Fourier coding of image boundaries. *IEEE Trans. Pattern Analysis and Machine Intelligence*, PAMI-6(1):102–105, 1984.
- [3] O. Faugeras and W. Pratt. Decorrelation methods of texture feature extraction. *IEEE Trans. Pattern Analysis and Machine Intelligence*, PAMI-2(4):323–332, 1980.
- [4] R.M. Haralick. Statistical and structural approaches to texture. *Proc. IEEE*, 67(5):786–804, 1979.
- [5] I. Hartmann. *Mustererkennung*. Skriptreihe Regelungstechnik und Bildverarbeitung, Technische Universität Berlin, 1996.
- [6] M.-K. Hu. Visual pattern recognition by moment invariants. *IRE Trans. Info. Theory*, IT(8):179–187, 1962.
- [7] B. Jähne. *Digitale Bildverarbeitung*. Springer, Berlin, Heidelberg, 2 edition, 1995.
- [8] K.-F Kamm. Grundlagen der Röntgenabbildung. In K. Ewen, editor, *Moderne Bildgebung: Physik, Gerätetechnik, Bildbearbeitung und -kommunikation, Strahlenschutz, Qualitätskontrolle*, pages 45–62, Stuttgart, New York, 1998. Georg Thieme Verlag.
- [9] D. Marr and E. Hildreth. Theory of edge detection. *Proc. Roy. Soc. London*, B(207):187–217, 1980.
- [10] MathWorks. *Image Processing Toolbox for Use with MATLAB: User’s Guide*. The MathWorks Inc., January 2003.
- [11] D. Mery. *Automated Flaw Detection in Castings from Digital Radioscopic Image Sequences*. Verlag Dr. Köster, Berlin, 2001. (Ph.D. Thesis in German).
- [12] E. Persoon and K.S. Fu. Shape discrimination using Fourier descriptors. *IEEE Trans. Systems, Man, and Cybernetics*, SMC-7(3):170–179, 1977.
- [13] M. Sonka, V. Hlavac, and R. Boyle. *Image Processing, Analysis, and Machine Vision*. PWS Publishing, Pacific Grove, CA, 2 edition, 1998.
- [14] C.H. Teh and R.T. Chin. On digital approximation of moment invariants. *Computer Vision, Graphics and Image Processing*, 33(3):318–326, 1986.

- [15] C.T. Zahn and R.Z. Roskies. Fourier descriptors for plane closed curves. *IEEE Trans. Computers*, C-21(3):269–281, 1971.