Capítulo 2

Representación de las imágenes y operaciones

Bibliografía:

- González, R.C., Wintz, P. (1996), *Procesamiento digital de imágenes.* Addison-Wesley, Tema 2, pág 23-56.
- Jain, R., Kasturi, R. y Schunck, B. (1995), *Machine Vision* , McGraw-Hill International Editions, Tema 1, pág 1-18.

Contenido

1. Modelo de imagen.
2. Adquisición de imágenes: muestreo y cuantificación.
3. Relaciones de vecindad, distancias y conectividad.
4. Operaciones sobre y entre imágenes.

2.1. Definición y representación de imágenes

Una **imagen** a niveles de gris es una función bidimensional de la intensidad de luz y

$$f : \mathcal{Z} \times \mathcal{Z} \to \mathcal{Z}$$

cuyos valores se han obtenido muestreando la intensidad sobre una retícula rectangular. Una imagen la denotaremos como $f(x, y)$, donde x e y son las cordenadas espaciales y el valor de f en cada punto (x, y) es proporcional a la intensidad de luz (nivel de gris) de ese punto.
Podemos decir que una imagen $f(x, y)$ está formada por dos componentes: una es la cantidad de luz incidente en la escena y la otra es la cantidad de luz reflejada por los objetos. Estas dos componentes se llaman: **iluminación**, que denotaremos por $i(x, y)$ y **reflectancia**, que denotaremos por $r(x, y)$. Entonces

$$f(x,y) = i(x,y)r(x,y)$$

con $0 < i(x,y) < \infty$ y $0 < r(x,y) < 1$. La iluminación $i(x,y)$ está determinada por las características de la fuente que emite la luz y la reflectancia $r(x,y)$ por las caracterísitcas del objeto. Por ejemplo, los valores normales que tenemos son: $i(x,y) = 100$ para una oficina comercial, $i(x,y) = 0{,}01$ para una noche clara. $r(x,y) = 0{,}01$ para el terciopelo negro y $r(x,y) = 0{,}93$ para la nieve.
Llamaremos **nivel de gris**, l, a la intensidad de luz de una imagen f. El rango de variación de l será:

$$L_{min} \leq l \leq L_{max}$$

En la práctica $L_{min} = i_{min}r_{min} \sim 0{,}005$ y $L_{max} = i_{max}r_{max} \sim 100$ en la mayoría de las aplicaciones de procesamiento de imágenes. Llamaremos al intervalo $[L_{min}, L_{max}]$ **escala de grises** y normalmente se desplaza al intervalo $[0, L]$, donde $l = 0$ se considera **negro** y $l = L$ se considera **blanco**. El resto de valores son variaciones de grises que varian de forma continua desde el negro hasta el blanco.

2.2. Muestreo y cuantificación

Una **imagen digital** es una imagen que ha sido discretizada en el espacio y en los valores de intensidad que puede tomar. Podemos considerar una imagen digital como una matriz en la que la fila y la columna representan un punto en la imagen y el valor del elemento de la matriz corresponde al nivel de gris de ese punto. Los elementos d e la matriz se llaman **puntos o pixels**.
A la discretización de las coordenadas espaciales le llamaremos **muestreo** de la imagen, y a la discretización de la amplitud le llamaremos **cuantización** del nivel de gris.
Supongamos que la imágen es muestreada espacialmente en las dos direcciones x e y de manera que tenemos una matriz NxN, donde los elementos toman valores discretos.

$$f(x,y) = \begin{pmatrix} f(0,0) & f(0,1) & \dots & f(0,N-1) \\ \vdots & \vdots & \vdots & \vdots \\ f(N-1,0) & f(N-1,1) & \dots & f(N-1,N-1) \end{pmatrix} \tag{2.1}$$

Normalmente se toma valores de N y para el número de niveles de gris G potencias de 2.

$$N = 2^n = 64, 128, 256, 512, 1024$$

$$G = 2^m = 2, ..., 256$$

Por tanto el número de bits para almacenar una imagen es:

$$b = NxNxm$$

¿Qué resolución espacial y niveles de gris necesitamos para tener una buena imagen?
La resolución (grado para discernir detalles) está estrechamente relacionado con n y m. Cuanto mayor sea el número de niveles de gris que tengamos para representar la imagen y mayor sea el número de puntos, mayor nivel de detalle se apreciará. Los niveles adecuados dependeran de cada aplicación. Como orientación, para tener una calidad semejante a la TV hemos de tener una imagen de 512x512 puntos con 128 niveles de gris.

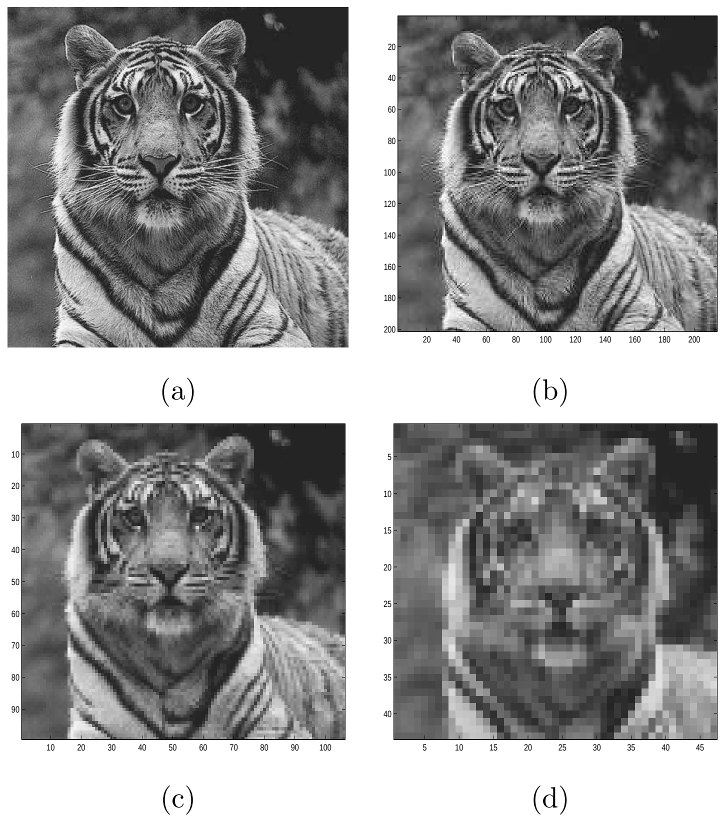
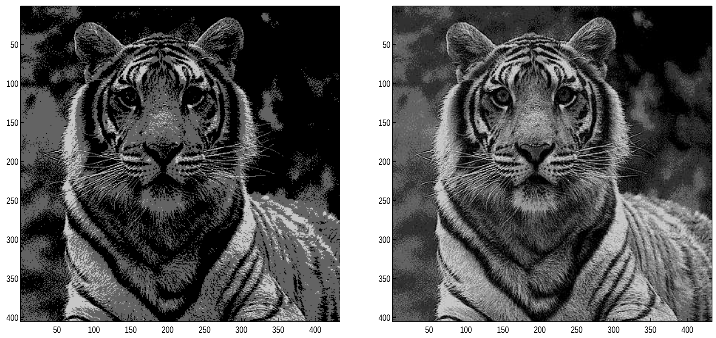


Figura 2.1: (a) Imagen original; (b) Imagen muestreada cada 2 pixels; (c)Imagen muestreada cada 4; (d) Imagen muestreada cada 8



(a) (b)

Figura 2.2: (a) Imagen original sólo con 3 niveles de gris; (b) con 6 niveles de gris;

Aparecen falsos contornos, debido al número reducido de níveles de gris.

2.3. Relaciones de vecindad, distancias y conectividad

Una imagen la denotaremos como $f(x,y)$, los puntos los denotaremos por letras minúsculas p o q y denotaremos por S un conjunto de puntos de la imagen $f(x,y)$.
Vecinos de un pixel
Un punto p de cordenadas (x,y) tiene 4 vecinos verticales y horizontales cuyas coordenadas son:

$$(x+1,y),(x-1,y),(x,y+1),(x,y-1)$$

A este conjunto de puntos se le llama los 4-vecinos de p y lo denotaremos como $\mathbf{N_4(p)}$.

Los 4 vecinos diagonales de p tienen como cordenadas:

$$(x+1,y+1),(x+1,y-1),(x-1,y+1),(x-1,y-1)$$

y los denotaremos como $\mathbf{N_D(p)}$. A estos 4 vecinos diagonales junto con $N_4(p)$ le llamaremos los 8-vecinos de p, denotado como $\mathbf{N_8(p)}$. Sus cordenadas son:

$(x-1,y-1)$	$(x,y-1)$	$(x+1,y-1)$
$(x-1,y)$	(x,y)	$(x+1,y)$
$(x-1,y+1)$	$(x,y+1)$	$(x+1,y+1)$

(2.2)

Conectividad

Para establecer si dos puntos están conectados debemos determinar si son adyacentes en algún sentido (por ejemplo 4-vecinos) y si su nivel de gris satisface algún criterio de similaridad. Sea V el conjunto de niveles de gris usados para definir la conectividad. Por ejemplo: sólo interesa la conectividad entre puntos con niveles de gris en el rango $1,\ldots,100$.

Consideramos dos tipos de conectividad:

4-Conectividad: dos puntos p y q con valores de niveles de gris en V están 4-conectados si q está en el conjunto $N_4(p)$.

8-Conectividad: dos puntos p y q con valores de niveles de gris en V están 8-conectados si q está en el conjunto $N_8(p)$.

0	0	0	0	0	0
0	13	43	0	34	22
0	12	0	54	98	67
0	21	45	0	86	0
0	0	0	0	76	0

(2.3)

Usando 4-conectividad tenemos dos objetos; usando 8 conectividad tenemos sólo un objeto.

2.3.1. Medidas de distancia

Dados los puntos p, q, con coordenadas $(x, y), (s, t)$, respectivamente:

Distancia Euclidea

Se define la distancia Euclidea entre dos puntos p y q como:

$$D(p,q) = \sqrt{(x-s)^2 + (y-t)^2}$$

Los puntos que están a una distancia D igual o menor a un valor r de (x, y) son puntos contenidos en un disco de radio r y centrado en (x, y).

D_4 **Distancia**

Se define la D_4 distancia entre dos puntos p y q como:

$$D_4(p,q) = |x-s| + |y-t|$$

Los puntos que están a una distancia D_4 de (x, y) igual o menor a un valor r forma un diamante centrado en (x, y). Por ejemplo, los puntos con una distancia $D_4 \leq 2$ de (x, y) son los siguientes. El número que aparece en la celda es el valor de la distancia.

		2		
	2	1	2	
2	1	0	1	2
	2	1	2	
		2		

(2.4)

D_8 **Distancia**

Se define la D_8 distancia entre dos puntos p y q como:

$$D_8(p,q) = max(|x-s|, |y-t|)$$

Los puntos que están a una distancia D_8 de (x, y) igual o menor a un valor r forman un cuadrado centrado en (x, y). Por ejemplo, los puntos con una distancia $D_8 \leq 2$ de (x, y) son:

2	2	2	2	2
2	1	1	1	2
2	1	0	1	2
2	1	1	1	2
2	2	2	2	2

(2.5)

2.4. Operaciones sobre imágenes: aritméticas y lógicas

Operaciones aritméticas:

- Suma: $p + q$ (por eje. en el promediado para la eliminación de ruido).
- Resta: $p - q$ (para la eliminación de información estática en la detección de movimiento).
- Multiplicación: αp (para aumentar los niveles de gris).
- Division: p/α.

Operaciones lógicas:

- AND: $pANDq$
- OR: $pORq$
- COMPLEMENT: $NOTq$

Las operaciones lógicas sólo se pueden aplicar sobre imágenes binarias.
Operaciones sobre una imagen

- a nivel de punto (por ejemplo, umbralización),

$$g(i, j) = O_{point}\{f(i, j)\}$$

- a nivel local sobre una vecindad mediante las máscaras de convolución (por ejemplo, suavizado o detección de bordes),

$$g(i, j) = O_{local}\{f(i_k, j_l); (i_k, j_l) \in N(i, j)\}$$

- y las operaciones a nivel global que dependen de toda la imagen (por ejemplo, la transformada de Fourier o el histograma).

2.5. Ejercicios

1. Leer la imagen *monedas.tif* y visualizala. *Funciones: imread, image.*
2. Hacer una función que, dada una imagen que se le pasa como parámetro, devuelva una nueva imagen cuya resolución espacial se ha dividido por dos, tanto en las filas como en las columnas.
3. Hacer una función, que dada una imagen que se le pasa como parámetro, devuelva una nueva imagen, donde se ha reducido a la mitad el número de niveles posibles de gris que podemos utilizar.

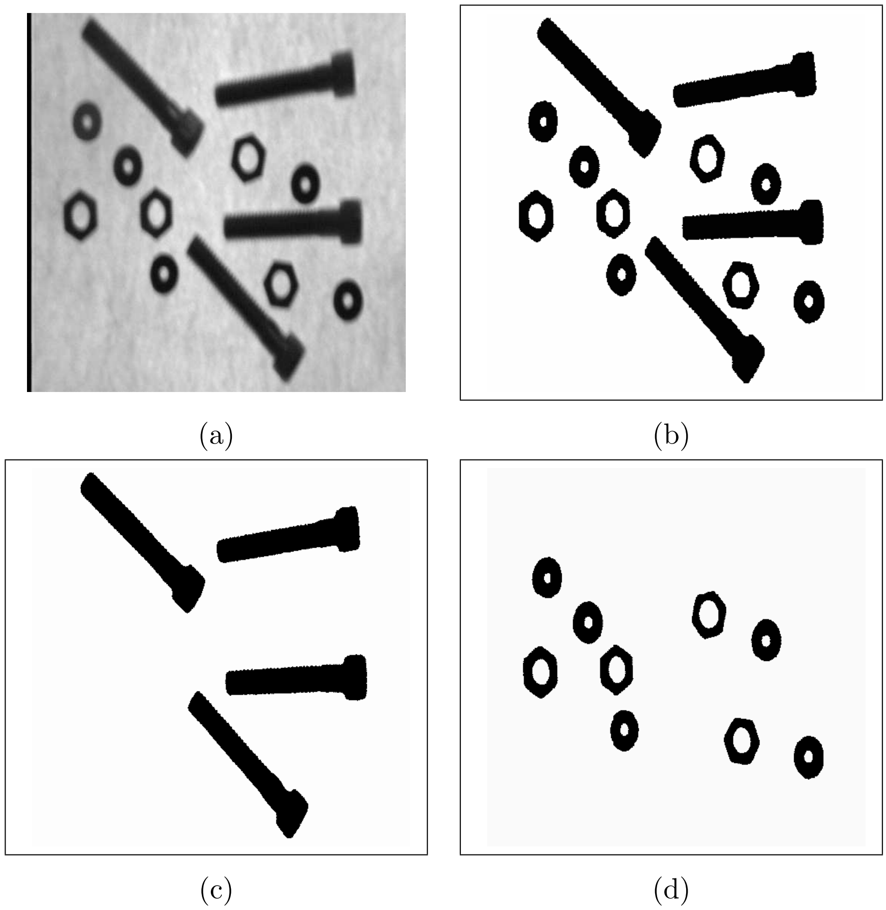
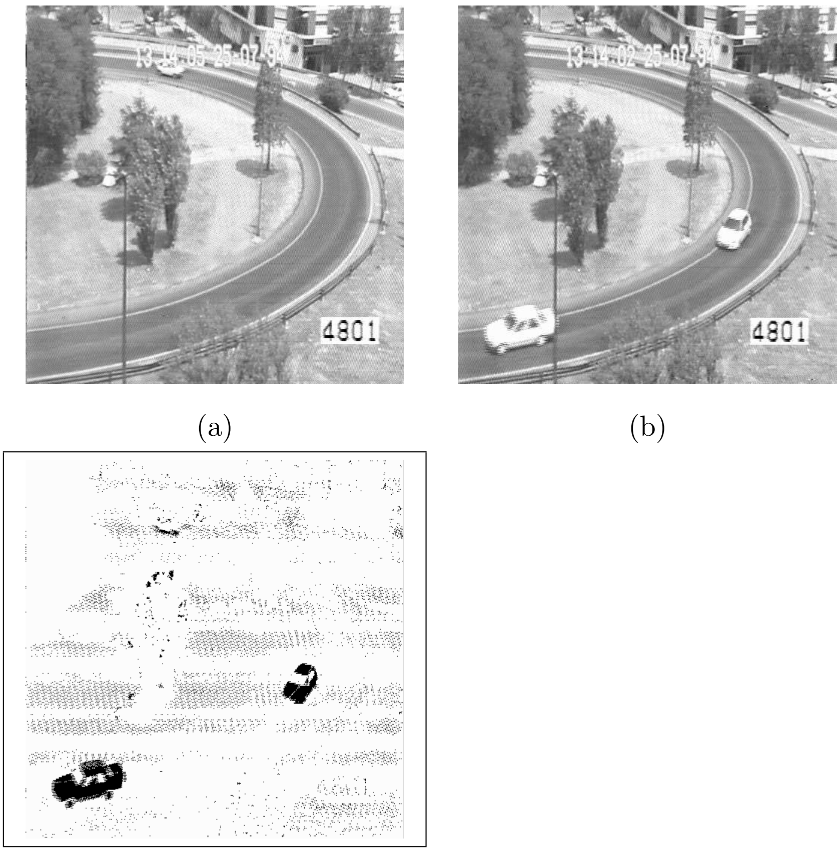


Figura 2.3: (a) Imagen original; (b) Imagen binarizada; (c) es xor(b,d)



(a) (b)

(c)

Figura 2.4: (a) Imagen de referencia $t = 0$; (b) Imagen en tiempo t; (c) Imagen Diferencia umbralizada