

Imágenes Digitales

Extracto de “Visión por Computador. Fundamentos y métodos”. Arturo de la Escalera. Prentice Hall, 2001

Copia para el alumno con fines didácticos

3. Imágenes digitales

El concepto de imagen está asociado a una función bidimensional $f(x,y)$, cuya amplitud o valor será el grado de iluminación (intensidad de la luz) en el espacio de coordenadas (x,y) de la imagen para cada punto. El valor de esta función depende de la cantidad de luz que incide sobre la escena vista, así como de la parte que sea reflejada por los objetos que componen dicha escena. Estos componentes son llamados *iluminación* y *reflexión*, siendo descritos por $i(x,y)$ y $r(x,y)$ respectivamente. El producto de ambas funciones proporciona la función $f(x,y)$:

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

Siendo el intervalo de definición:

- $0 < i(x,y) < \infty$
- $0 < r(x,y) < 1$

por lo que $f(x,y)$ estará acotada:

$$0 < f(x,y) < \infty$$

La naturaleza de la iluminación viene determinada por la fuente de luz, mientras que la reflexión depende de las características del objeto en la escena.

3.1 Muestreo espacial y niveles de gris.

Una señal es una función que depende de alguna variable que tiene un significado físico. Dependiendo del número de variables puede ser unidimensional, o bidimensional, como el caso de las imágenes. Dependiendo de las dimensiones de la función pueden ser escalares (como las imágenes en blanco y negro), o vectoriales (como las imágenes en color, que tienen tres componentes). Las funciones tienen un dominio y un rango. Si el dominio y rango son continuos, la señal es continua o analógica; si el dominio es discreto pero el rango no, la señal también será discreta; y si el dominio y el rango son discretos, como en el caso de las imágenes, la señal es digital.

La función de imagen $f(x,y)$ es digitalizada en la memoria del computador, tanto espacialmente como en amplitud. La digitalización de las coordenadas espaciales (x,y) está asociada al concepto de *muestreo*, mientras que la digitalización de la amplitud al de cuantificación de los *niveles de gris*. Una imagen $f(x,y)$ está por tanto almacenada en una matriz de $N \times M$ elementos¹. Entenderemos que el origen de coordenadas de la imagen es la esquina superior izquierda, el eje x el horizontal y el eje y el vertical.

$$f(x,y) = \begin{bmatrix} f(0,0) & f(1,0) & \cdots & f(N-2,0) & f(N-1,0) \\ f(0,1) & f(1,1) & \cdots & f(N-2,1) & f(N-1,1) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ f(0,M-2) & f(1,M-2) & \cdots & f(N-2,M-2) & f(N-1,M-2) \\ f(0,M-1) & f(1,M-1) & \cdots & f(N-2,M-1) & f(N-1,M-1) \end{bmatrix}$$

El muestreo es la conversión que sufren las dos dimensiones espaciales de la señal analógica, y que genera la noción de píxel. La imagen en la memoria del computador se comporta como una matriz donde el valor de los elementos es su nivel de

¹ En general N será igual a M y ambas potencias de 2.

gris, y la fila y la columna su posición en el espacio. La cuantificación es la conversión que sufre la amplitud de la señal analógica; así se genera el concepto de nivel de gris o intensidad. Para el caso de tener 256 niveles de gris (0-255), el 0 corresponde a un objeto no iluminado o que absorbe todos los rayos luminosos que inciden sobre él (negro), y el nivel 255 a un objeto muy iluminado o que refleja todos los rayos que inciden sobre él (blanco). En la figura 3.1 viene reflejado el proceso de digitalización de un objeto a varias resoluciones. Al efecto de disminuir la resolución y mantener las dimensiones físicas se le conoce como efecto *tablero de ajedrez*.

En el proceso de digitalización es preciso establecer el valor de N y M así como el número de niveles de gris asignados para cada píxel. Es una práctica común en el proceso de digitalización de imágenes que estas cantidades sean números enteros potencias de dos. El número de bits b requeridos para almacenar la imagen digitalizada viene dado por el producto:

$$b = N M g$$

siendo 2^g el número de niveles de gris.

El concepto de resolución surge asociado al número de muestras y niveles de gris necesarios para tener una buena aproximación de la imagen. Es interesante apreciar el efecto de las variaciones de los parámetros citados sobre la calidad de una misma imagen, y cómo es difícil establecer el valor óptimo de cada uno de ellos para los requerimientos de diferentes aplicaciones. Si la imagen posee detalles de poca resolución, las nociones de cuantificación y muestreo estarán íntimamente ligadas.

En la imagen 3.2 aparece el mismo objeto digitalizado con diversos niveles de gris. Es importante notar que aparentemente no existen diferencias entre las primeras imágenes. Esto es debido a que el ojo humano es incapaz de distinguir más de dos docenas de niveles de gris distintos. Conforme disminuye el número de niveles de gris aparece el efecto conocido como *falsos contornos*.

Una vez definidas las resoluciones puede hablarse del *histograma* de una imagen. El histograma contiene el número de píxeles que tienen el mismo nivel de gris (figura 3.3). Puede entenderse como la probabilidad de que un valor de gris determinado aparezca en la imagen. Tiene la gran ventaja que condensa la información de la imagen en muchos menos valores (256 si cada píxel son 8 bits) pero, aunque en los capítulos siguientes se verá que es una de las herramientas más importantes a la hora de procesar y analizar una imagen, el histograma tiene el gran inconveniente de que por sí solo no aporta ninguna información sobre la localización espacial de los píxeles. Así se tendrán infinidad de imágenes distintas que tienen el mismo histograma. A lo largo del libro se verán otras utilidades del histograma como obtener el número de objetos en una imagen, su área y las diversas texturas.

Un último elemento de la imagen son los *planos de bits* (figura 3.4). Si una imagen tiene 256 niveles de gris, cada uno de los píxeles ocupa un *byte* (8 *bits*):

Pixel x	Bit 7	Bit 6	Bit 5	Bit 4	Bit 3	Bit 2	Bit 1	Bit 0
Pixel x+1	Bit 7	Bit 6	Bit 5	Bit 4	Bit 3	Bit 2	Bit 1	Bit 0

Cada plano de bit es una imagen formada por un determinado bit de cada píxel. La importancia de los planos de bits es que reflejan muy bien la influencia del ruido en la imagen. Así puede observarse como para planos de bits altos se pueden distinguir los distintos objetos que forman la imagen, pero para valores bajos la impresión que dan es una distribución aleatoria de puntos.

3.2 Relaciones entre píxeles: vecindad, conectividad.

Un píxel p de coordenadas (x,y) presenta un total de cuatro vecinos en el plano vertical y horizontal, siendo sus coordenadas:

	$x, y-1$	
$x-1, y$	x, y	$x+1, y$
	$x, y+1$	

Este conjunto de píxeles se denomina vecindad de tipo 4 del píxel p , y se representa por $N_4(p)$. Además se puede considerar la existencia de otros cuatro vecinos asociados a las diagonales, cuyas coordenadas son:

x-1,y-1		x+1,y-1
	x,y	
x-1,y+1		x+1,y+1

los cuales se representan por $ND(p)$. La suma de los anteriores define los ocho vecinos del píxel p , $N8(p)$.

Mediante el concepto de *conectividad* se quiere expresar que dos pixeles pertenecen al mismo objeto, por lo que está relacionado con el de vecindad. Dos pixeles están conectados si son adyacentes (vecinos) y si sus niveles de gris satisfacen algún criterio de especificación (por ejemplo ser iguales).

Existen tres tipos:

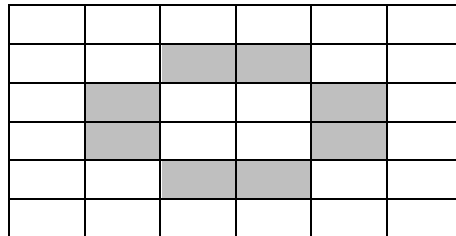
- Conectividad-4. Dos pixeles p y q presentan una conectividad-4 si q pertenece al $N4(p)$.
- Conectividad-8. Dos pixeles p y q presentan una conectividad-8 si q pertenece al $N8(p)$.
- Conectividad-m (conectividad mixta). Dos pixeles p y q presentan una conectividad-m si:
 1. q pertenece a $N4(p)$, o
 2. q pertenece a $ND(p)$ y el conjunto $N4(p) \cap N4(q)$ es el conjunto vacío.

La conectividad mixta es una modificación de la conectividad-8 cuya utilidad es eliminar las conexiones múltiples que en ocasiones aparecen cuando la conectividad-8 es utilizada. Por ejemplo:

1	1	1		1	1	1		1	1	1	
					\		/				
1	-	1	0	1	-	1	0	1	-	1	0
						\				\	
0	0	1		0	0	1		0	0	1	
Conectividad-4				Conectividad-8				Conectividad-m			

En este caso, para ir desde el pixel central al inferior derecha se utilizaría conectividad-8. Sin embargo para ir al pixel superior derecha se podría ir de dos formas, directamente o a través del pixel superior. Esta indeterminación se evita con la conectividad mixta. Ahora el pixel superior derecha no está conectado al central ya que sí lo está al superior.

En la figura siguiente se observa que si se utiliza la conectividad de tipo 4, en la figura existen cuatro objetos, cuando de manera intuitiva se observa que es uno solo. Si entonces se toma la de tipo 8 el anterior problema desaparece; pero ahora tanto el fondo exterior como el interior están conectados. Por ello cuando se quiere determinar si dos pixeles pertenecen al mismo objeto se toma conectividad de tipo 8 y para el fondo conectividad de tipo 4 o viceversa.



3.3 Distancia.

Con la distancia se quiere obtener el mínimo número de pasos elementales que se necesitan para ir de un punto a otro. Dados tres pixeles p , q y z , con coordenadas (x,y) , (s,t) y (u,v) respectivamente, se puede definir una función de distancia D si cumple:

- $D(p, q) \geq 0$, ($D(p,q) = 0$, si $p=q$).
- $D(p, q) = D(q, p)$
- $D(p, z) \leq D(p, q) + D(q, z)$

Las funciones de distancia usadas comúnmente son:

La *distancia Euclídea* entre p y q se define:

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

con esta definición, las distancias serán:

$$\begin{array}{ccccc} \sqrt{8} & \sqrt{5} & 2 & \sqrt{5} & \sqrt{8} \\ \sqrt{5} & \sqrt{2} & 1 & \sqrt{2} & \sqrt{5} \\ 2 & 1 & 0 & 1 & 2 \\ \sqrt{5} & \sqrt{2} & 1 & \sqrt{2} & \sqrt{5} \\ \sqrt{8} & \sqrt{5} & 2 & \sqrt{5} & \sqrt{8} \end{array}$$

La distancia euclídea es la más exacta pero presenta el inconveniente de su gran número de cálculos. Por otro lado no tiene en cuenta el concepto de vecindad antes descrito. Es por ello que se han definido otras.

Distancia Manhattan. Se toman solamente en cuenta los vecinos de orden 4.

$$D = |x - s| + |y - t|$$

Las distancias serán por tanto:

4	3	2	3	4
3	2	1	2	3
2	1	0	1	2
3	2	1	2	3
4	3	2	3	4

Con esta definición los vecinos de tipo 4 están a la distancia unidad. Si lo que se quiere es que los vecinos de tipo 8 estén a la misma distancia se toma:

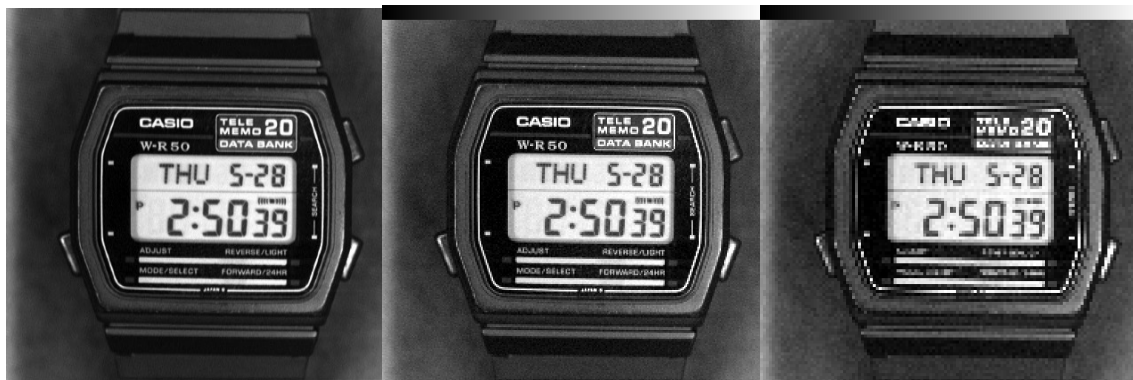
$$D(p,q)=\max(x-s,y-t)$$

Obteniéndose la distancia *tablero de ajedrez*. Y por tanto se tiene:

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

3.5 Bibliografía.

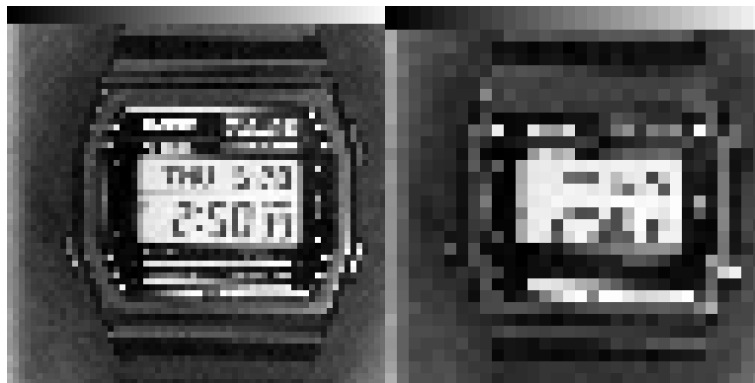
- González, R. C.; Woods, R E.. *Digital image processing*, Addison-Wesley, 1993
- Pratt, W. K. *Digital image processing*,. 2ª edición. John Wiley & Sons, cop. 1991.
- Jain, A. K. *Fundamentals of digital image processing*, Prentice-Hall International 1989
- Ohta, Y., Kanade, T., Sakai, T. Color information for region segmentation. *Computer Graphics and Image Processing* 13, 222-241,1980.
- Russ, J. C. *The image processing handbook*,. 2ª edición. CRC Press, 1995



512x512

256x256

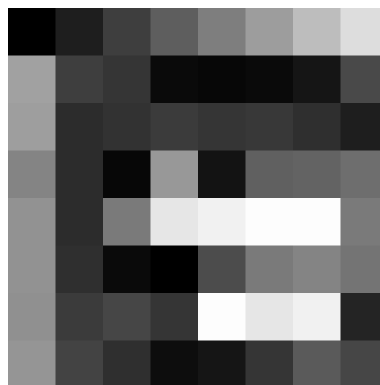
128x128



64x64

32x32

16x16

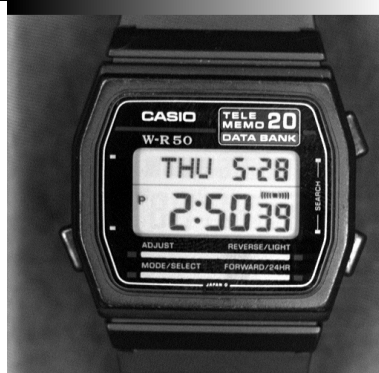


8x8

Figura 3.1 Diversos muestreos espaciales

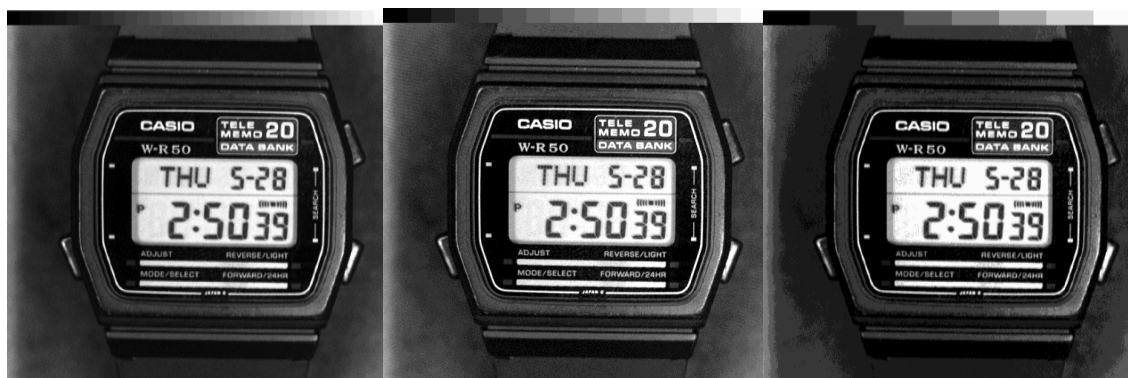


256



128

64



32

16

8



4

2

Figura 3.2 Diversos niveles de gris.

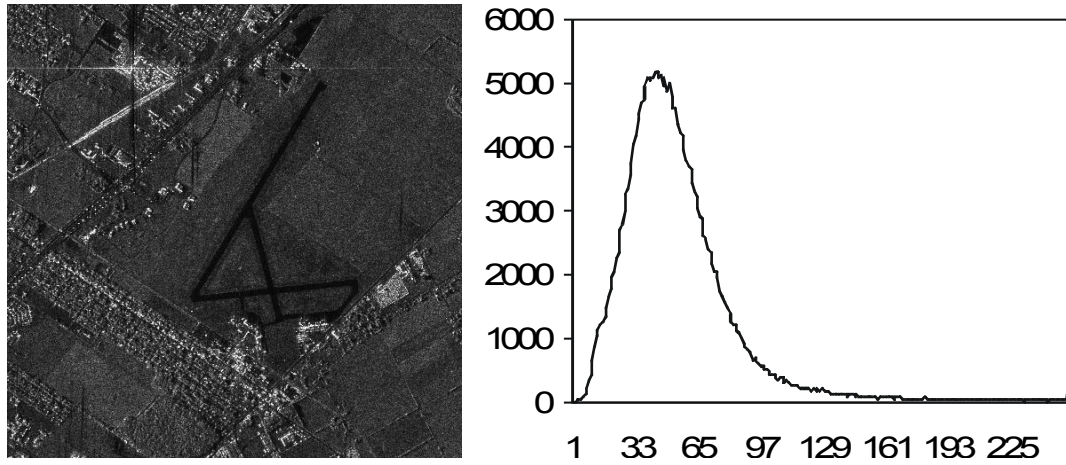


Figura 3.3. Imagen aérea de un aeropuerto y su histograma.

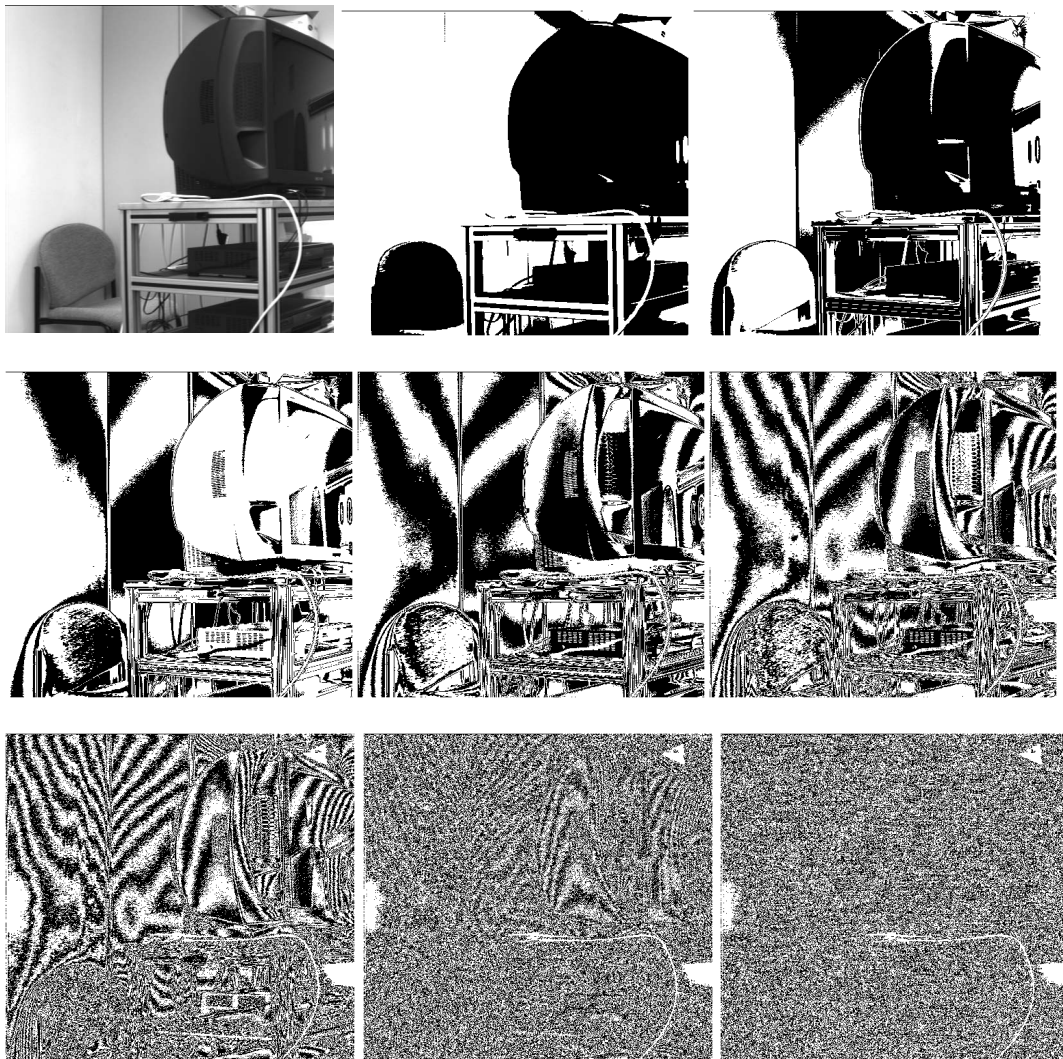


Figura 3.4. planos de bits