

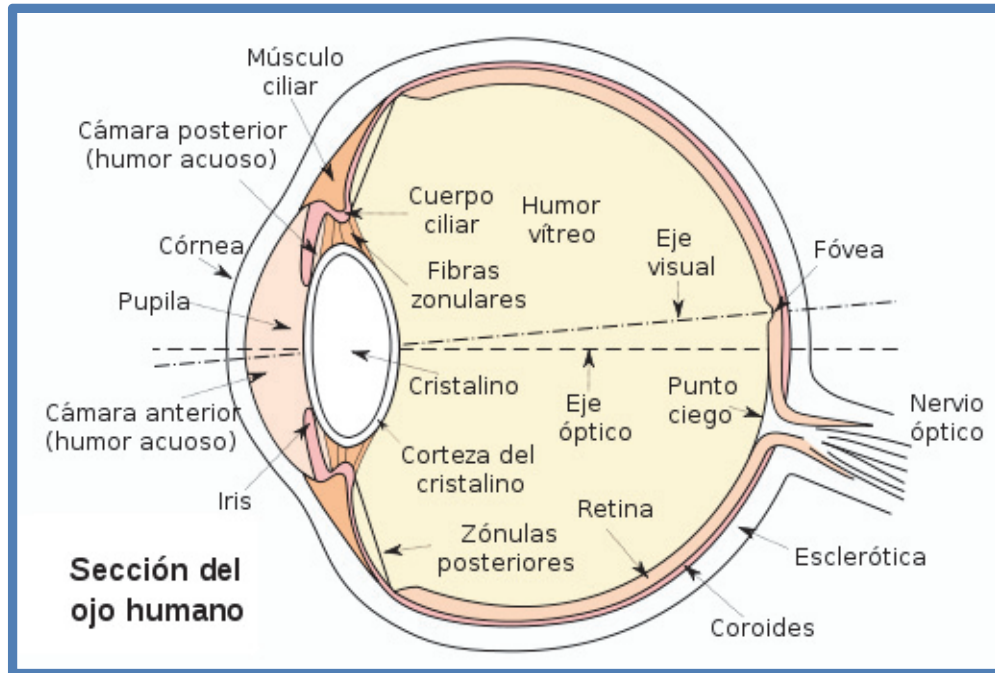
2. Fundamentos de imágenes digitales (primera parte)

Agenda

2. Fundamentos de imágenes digitales

- 2.1. Cómo se crean las imágenes?.
- 2.2. Luz y espectro electromagnético.
- 2.3. Modelo simplificado de formación de imágenes en una cámara fotográfica.
- 2.4. Muestreo y cuantización de imágenes análogas.
- 2.5. Términos de interés: tamaño del sensor óptico, número de píxeles en un fotograma, número de bits de cuantización por pixel, número de canales de información, número de cuadros por segundo en un video, tamaño de imágenes y videos sin comprimir.
- 2.6. Convenciones para las coordenadas espaciales y de intensidad en imágenes, y para las coordenadas temporales en video.
- 2.7. Relaciones básicas entre píxeles.

Elementos de percepción visual



- Diámetro aprox. 20 mm.
- 3 membranas:
 - ✓ Cornea – esclerótica.
 - ✓ Coroides.
 - ✓ Retina.
- Coroides:
 - ✓ Cuerpo ciliar.
 - ✓ Iris (su apertura central es la pupila).
- Cristalino (lens): Cambia su forma.

- Receptores de luz ubicados en la retina:
 - ✓ Conos : 6-7 millones por ojo. Responsables de la visión a color. Conexión a nervios individuales. Se ubican principalmente en la fóvea. Conos Rojos, Verdes y Azules (fotopigmento). Curva característica de absorción respecto a λ .
 - ✓ Bastones: 75-150 millones por ojo. Responsables de la visión en condiciones de baja luminosidad. Se conectan en grupo. Imagen general sin detalles.

FIGURE 2.1
Simplified
diagram of a cross
section of the
human eye.

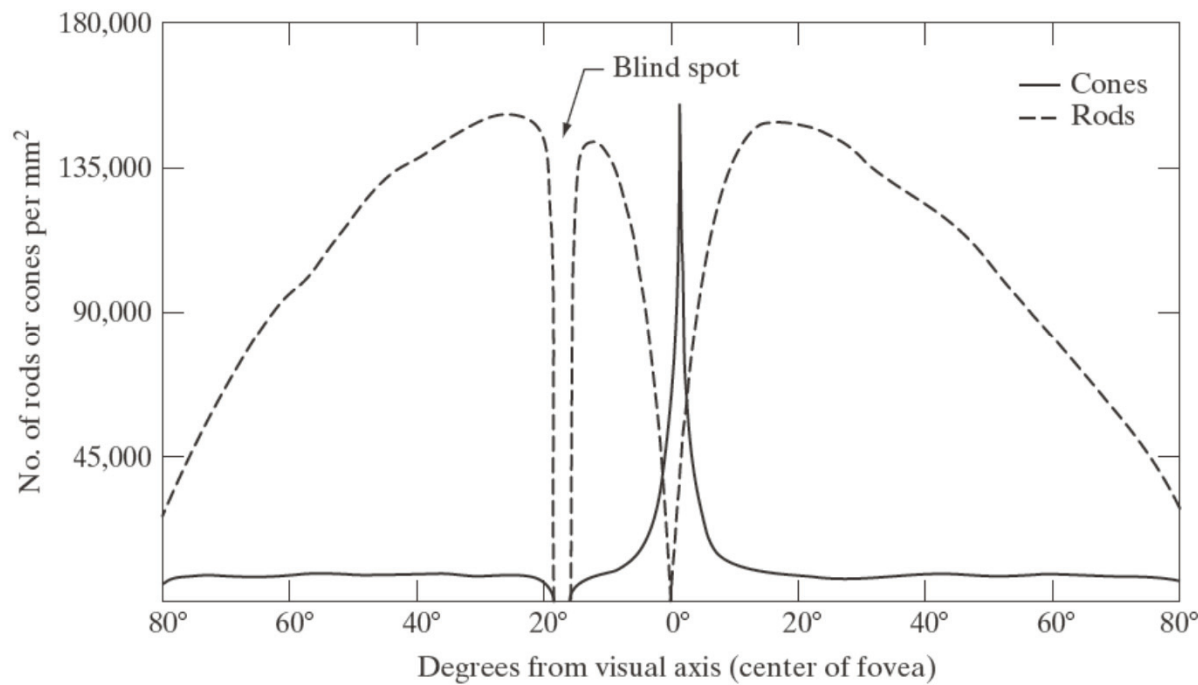
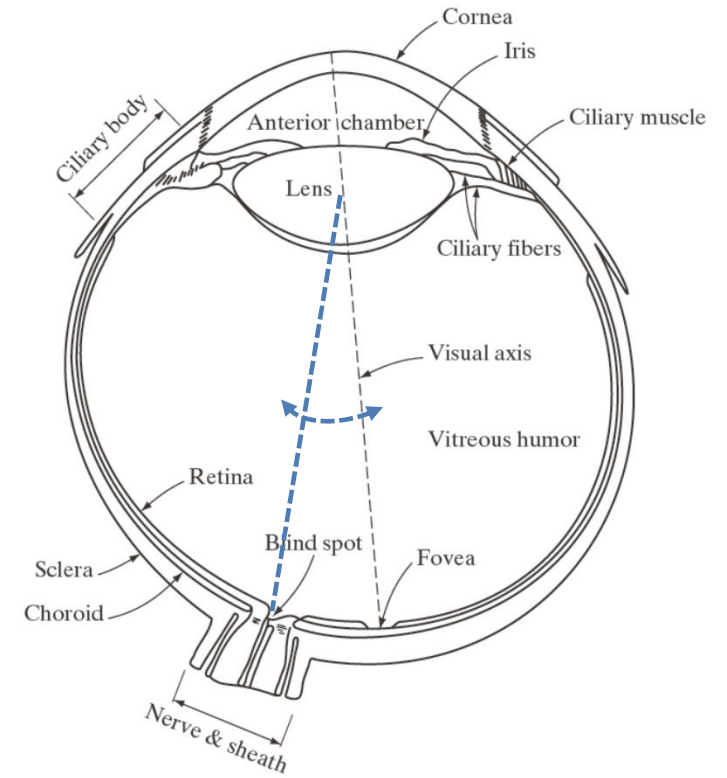


FIGURE 2.2
Distribution of
rods and cones in
the retina.

Formación de imágenes en el ojo

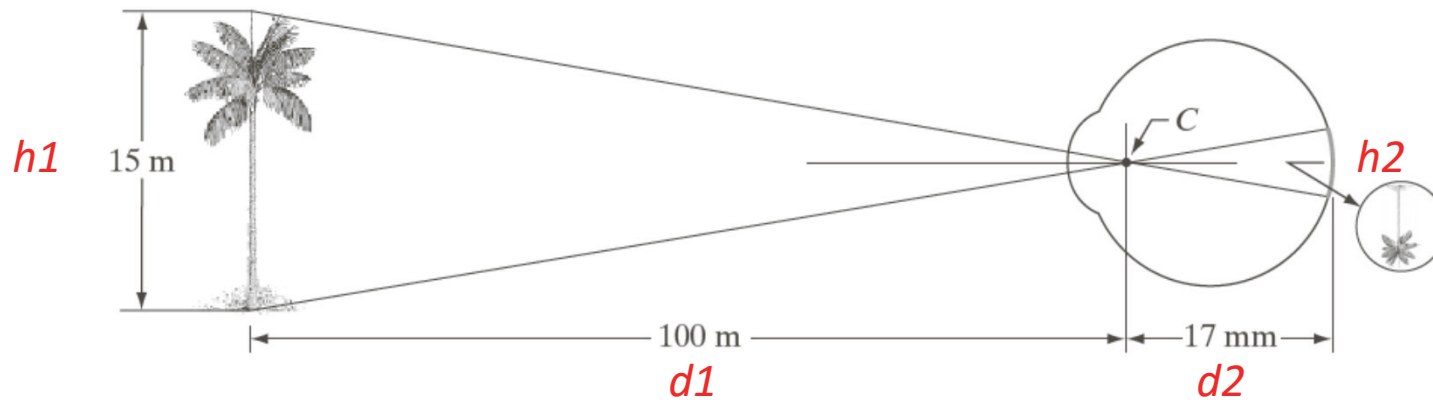


FIGURE 2.3

Graphical representation of the eye looking at a palm tree. Point C is the optical center of the lens.

- Distancia focal en el ojo humano
 - ✓ 14-17 mm (medido desde el centro del cristalino hasta la retina a lo largo del eje visual).
 - ✓ 17 mm: ojo relajado y enfocando en distancias mayores a 3 metros
- En el ejemplo la altura de la palma que se forma en la retina (h_2) se obtiene como
 - ✓ $h_1/d_1 = h_2/d_2$
 - ✓ Luego $h_2 = 2.55\text{ mm}$.

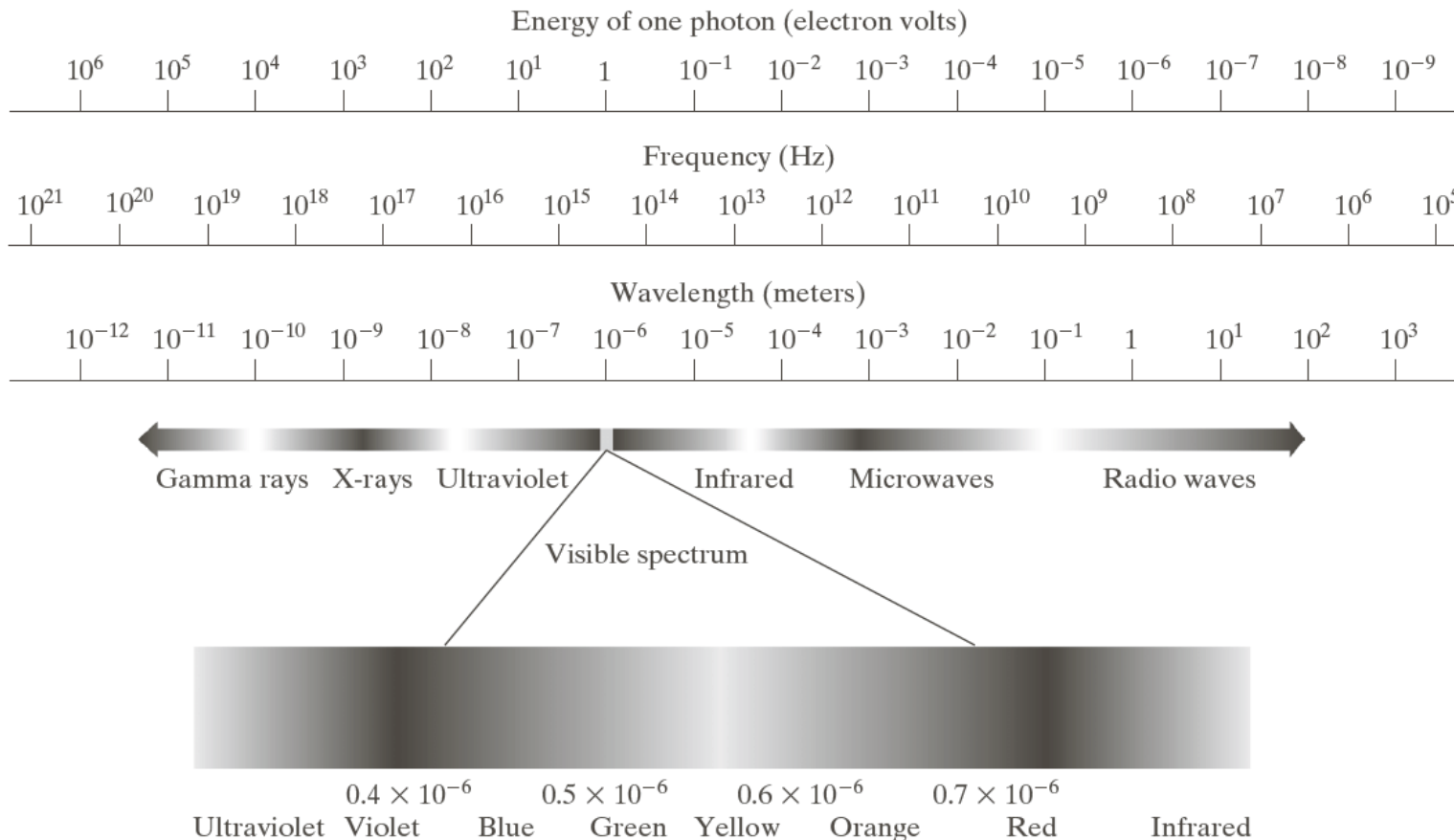
Calibración de una cámara

- **Calibración intrínseca:** Permite determinar los parámetros propios de la cámara (f_x, f_y, c_x, c_y) y los coeficientes que caracterizan la distorsión radial y tangencial del lente de acuerdo con unos modelos matemáticos predefinidos (k_1, k_2, k_3, p_1, p_2). **Lectura para la casa:**

http://docs.opencv.org/doc/tutorials/calib3d/camera_calibration/camera_calibration.html

- **Calibración extrínseca.** Encontrar la matriz de rotación **R** y el vector translación **T**.

Luz y espectro electromagnético



$$\lambda = \frac{c}{f}$$

$$E = hf$$

FIGURE 2.10 The electromagnetic spectrum. The visible spectrum is shown zoomed to facilitate explanation, but note that the visible spectrum is a rather narrow portion of the EM spectrum.

Para recordar....

- El color de un objeto depende de la naturaleza de la luz reflejada por el objeto.
- Objeto Blanco: *refleja* bien en todas las longitudes de onda.
- Objeto Negro: *absorbe* todas las longitudes de onda.
- Objeto verde: Refleja luz con longitudes de onda entre 500 – 700 nm y absorbe la mayor parte de la energía con otras longitudes de onda.
- Luz sin color: monocromática (o acromática). Sólo se caracteriza por su intensidad.

- Nivel de gris (Negro-gris-blanco): Intensidad monocromática.
- Imagen monocromática: imagen en escala de grises.
- Luz cromática (a color): Longitudes de onda entre $0.43 - 0.79 \mu\text{m}$.

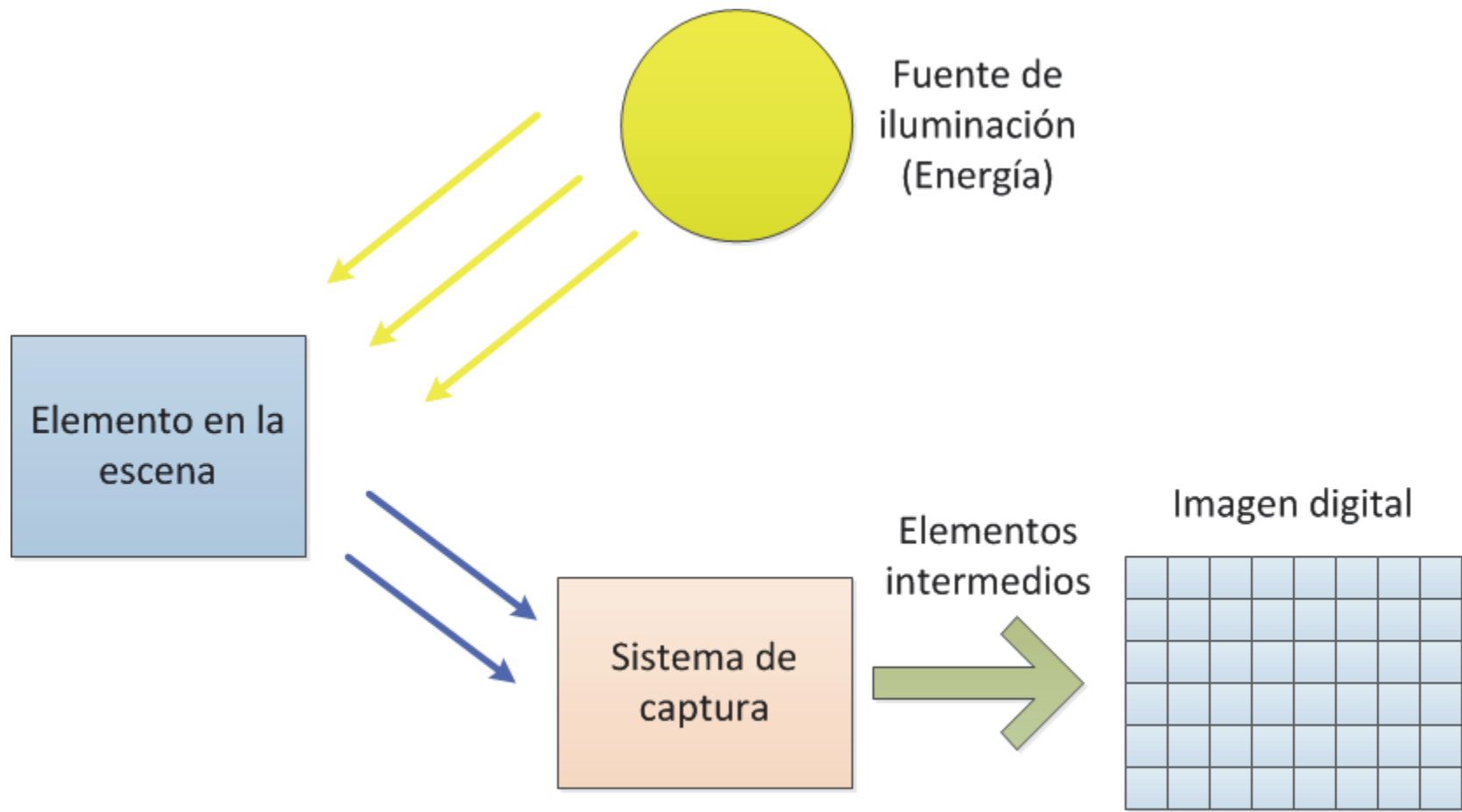
Calidad de una fuente de luz cromática

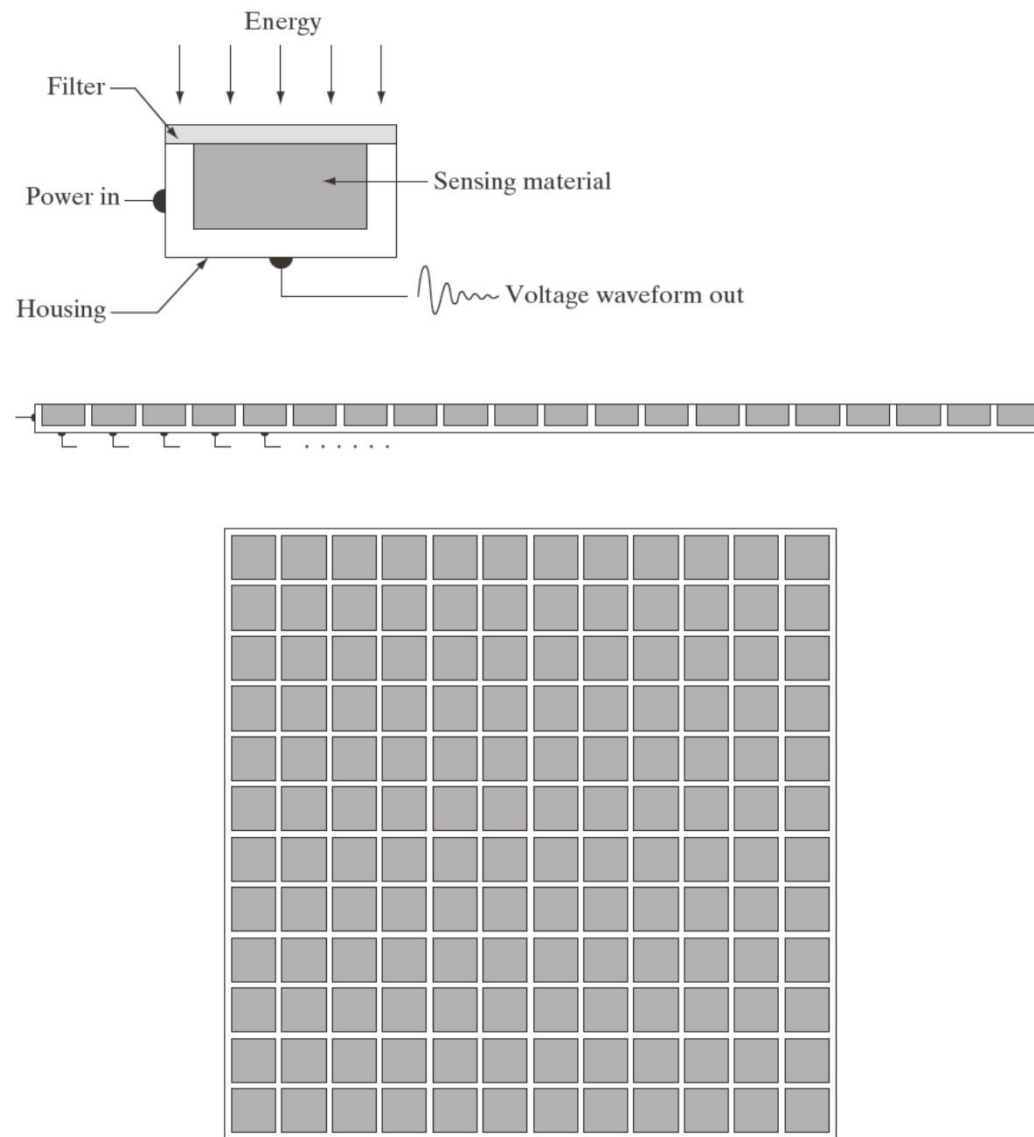
- Frecuencia, radiancia, luminancia, brillo.
- **Radiancia** (**radiance**): Cantidad total de energía que fluye desde la fuente de luz. Se mide en Watts (W).
- **Luminancia** (**Luminance**): Da una medida de la cantidad de energía que percibe un observador de una fuente de luz (contempla la sensibilidad variable del ojo humano a las diferentes longitudes de onda de la luz). Se mide en lumen (lm). *Ejemplo*: una fuente IR puede tener una alta radiancia pero un luminancia casi nula.
- **Brillo** (**brightness**): Descriptor subjetivo de la percepción de luz. Difícil de medir. Incorpora la noción acromática de intensidad. Aspecto clave para describir la sensación de color.

Elección de la banda EM

- La longitud de onda de una onda electromagnética requerida para “ver” un objeto debe ser del mismo tamaño o más pequeña que el objeto mismo.

Sensado de imágenes y adquisición





a
b
c

FIGURE 2.12
 (a) Single imaging sensor.
 (b) Line sensor.
 (c) Array sensor.

Adquisición usando un solo sensor

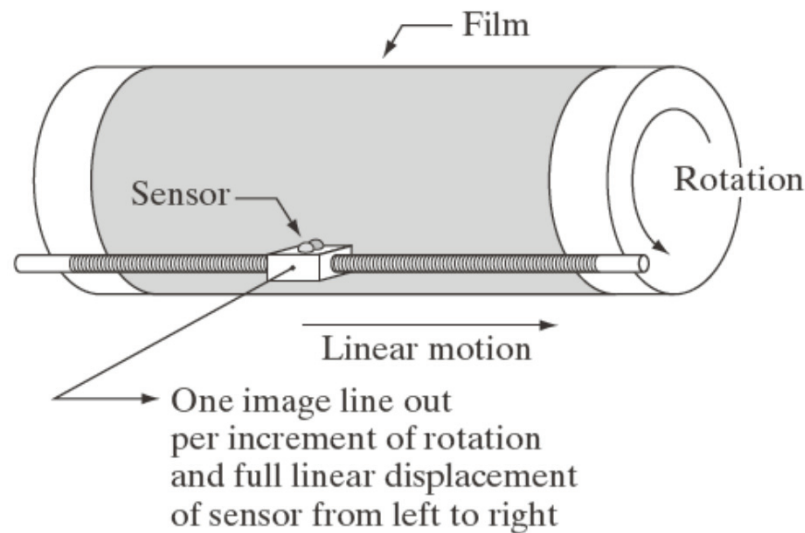


FIGURE 2.13
Combining a single sensor with motion to generate a 2-D image.

- Fotodiodo: Voltaje de salida proporcional a la luz.
- Desplazamientos en x e y .
- Alta resolución, económico, lento.

Adquisición usando arreglos lineales y circulares de sensores

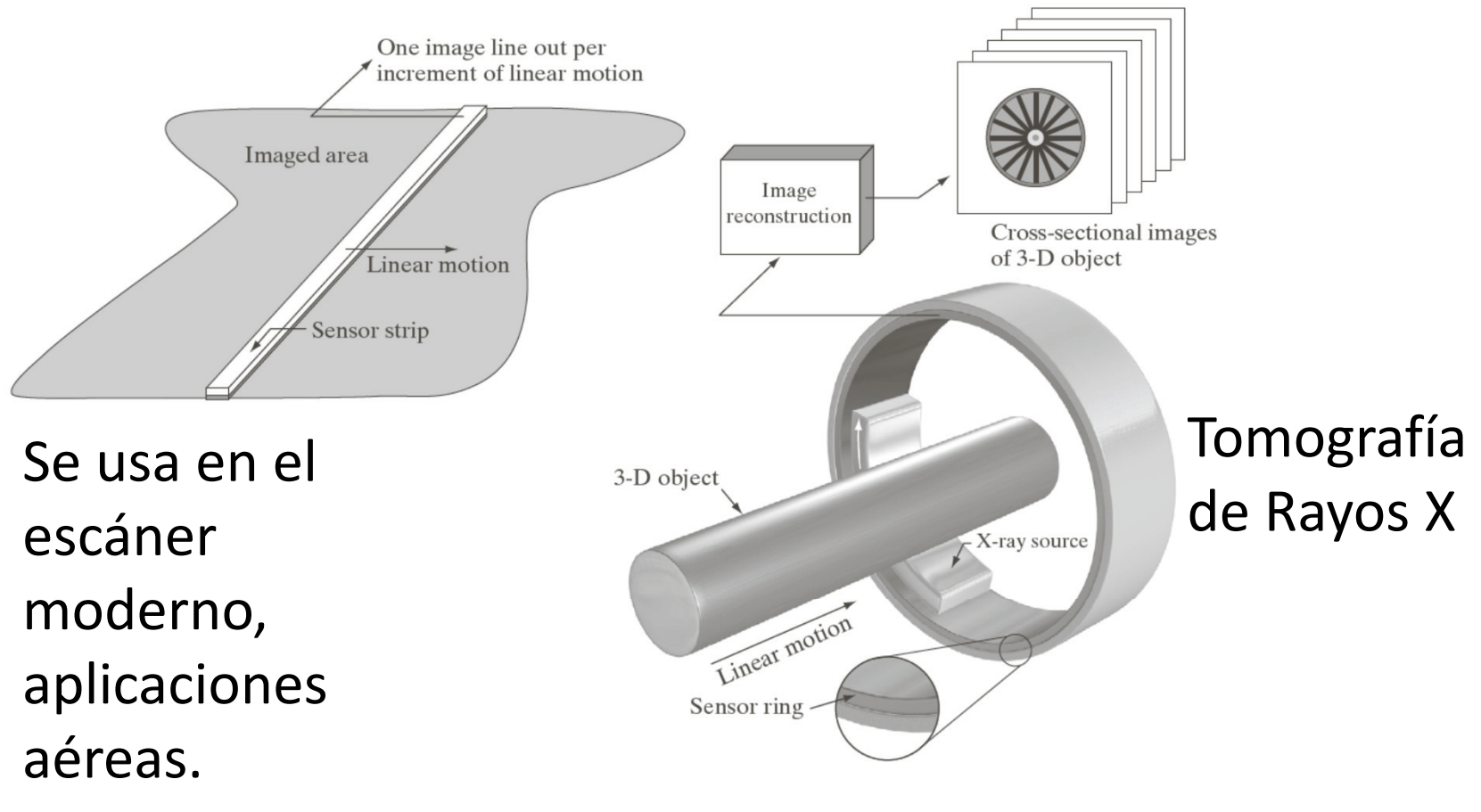


FIGURE 2.14 (a) Image acquisition using a linear sensor strip. (b) Image acquisition using a circular sensor strip.

Adquisición usando arreglos de sensores en dos dimensiones

- CMOS, CCD.
- Más de 4000 x 4000 elementos.
- ✓ Respuesta proporcional a la integral de la energía lumínica proyectada sobre la superficie del sensor.
- ✓ No requiere movimiento.

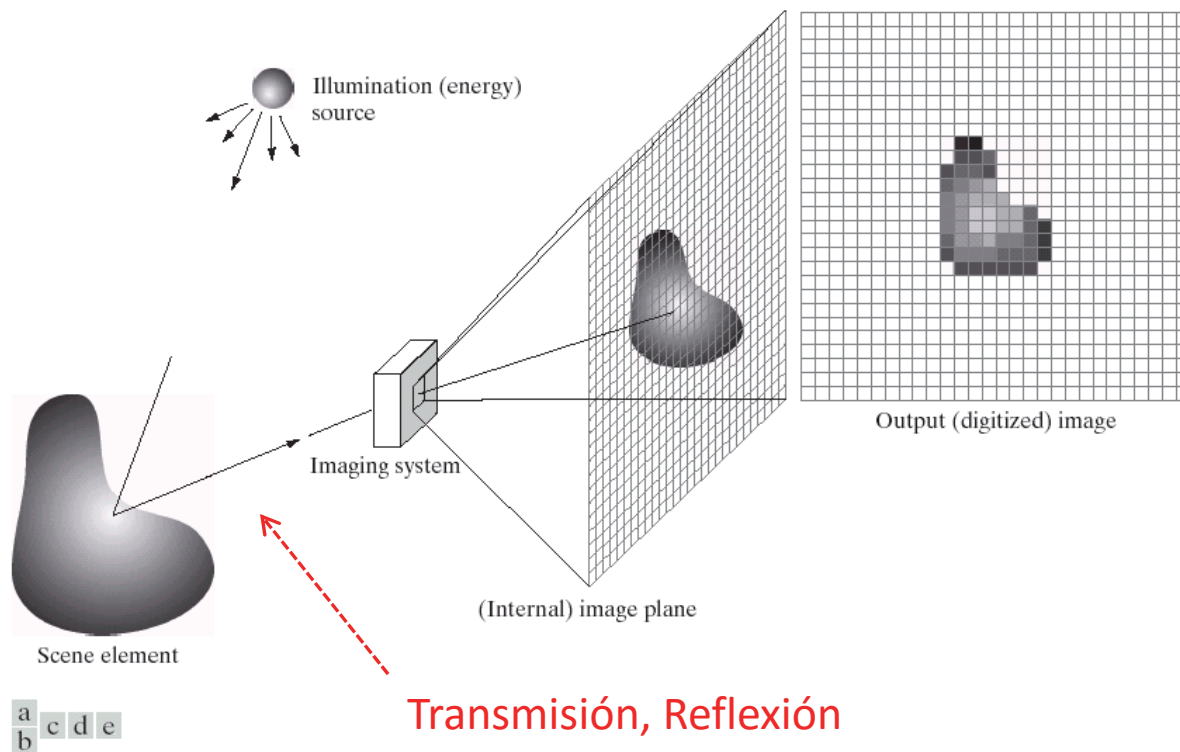


FIGURE 2.15 An example of the digital image acquisition process. (a) Energy (“illumination”) source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

Modelo simplificado de la formación de una imagen

$$f(x, y)$$

Imagen

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

Cantidad escalar positiva

$$i(x, y)$$

Iluminación

$$r(x, y)$$

Reflectancia

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

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

$$0 < r(x, y) < 1$$

{ 0: Absorción total.
1: Reflexión total.

Valores típicos de iluminación

- Lumen (lm): Unidad de flujo luminoso.
 $1 \text{ (lm)} = 1 \text{ (cd} \cdot \text{sr)} = 1 \text{ (lx} \cdot \text{m}^2 \text{)}.$
- En la superficie terrestre debido al sol/luna:
 - ✓ Día soleado: $i=90\,000 \text{ (lm/ m}^2\text{)}.$
 - ✓ Día nublado: $i=10\,000 \text{ (lm/ m}^2\text{)}.$
 - ✓ Noche despejada con luna llena: $i=0.1 \text{ (lm/ m}^2\text{)}.$
- Iluminación en una oficina comercial: $i=1000 \text{ (lm/ m}^2\text{)}.$

Valores típicos de reflectancia

- Terciopelo negro: $r=0.01$
- Acero inoxidable: $r=0.65$
- Pared plana pintada de blanco: $r=0.80$
- Placa de plata: $r=0.90$
- Nieve: $r=0.93$

Escala de grises en imágenes monocromáticas

$$l = f(x_0, y_0)$$

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

$$L_{\min} = i_{\min} r_{\min}$$

$$L_{\max} = i_{\max} r_{\max}$$

$$[L_{\min}, L_{\max}]$$

$$[0, L - 1]$$

Para una oficina aprox. 10.

Para una oficina aprox. 1000.

Escala de gris o escala de intensidad.

$$\left\{ \begin{array}{ll} l = 0 & \text{:Negro.} \\ l = L - 1 & \text{:Blanco.} \end{array} \right.$$

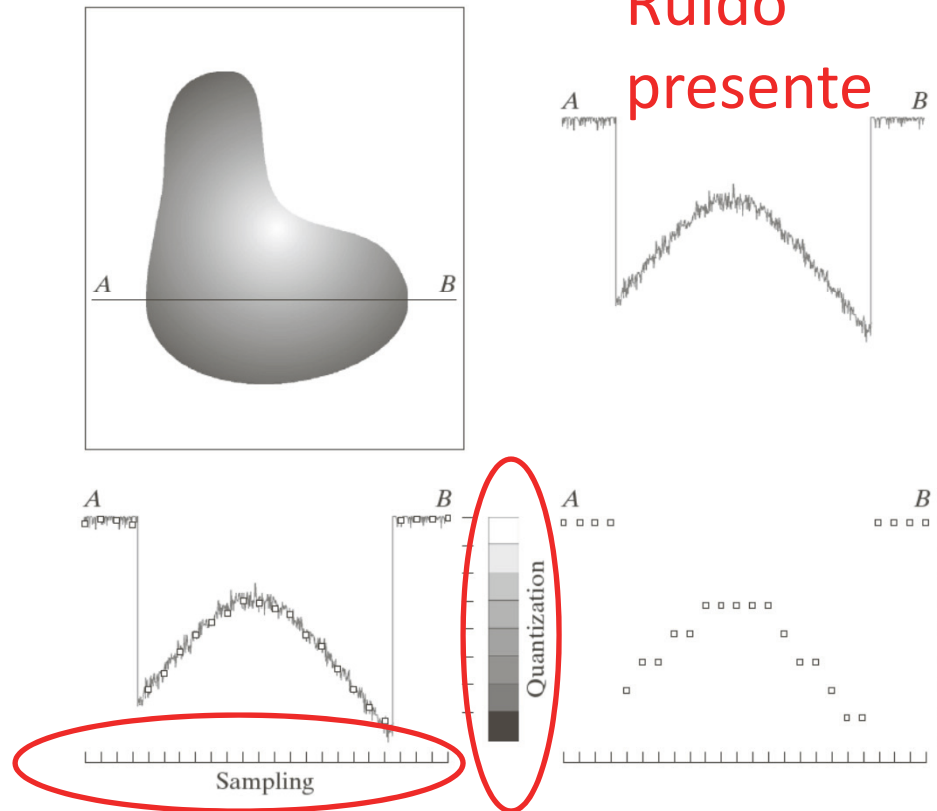
Muestreo y cuantización

- Convertir información continua de los sensores a un formato digital.

- Dos procesos de digitalización:

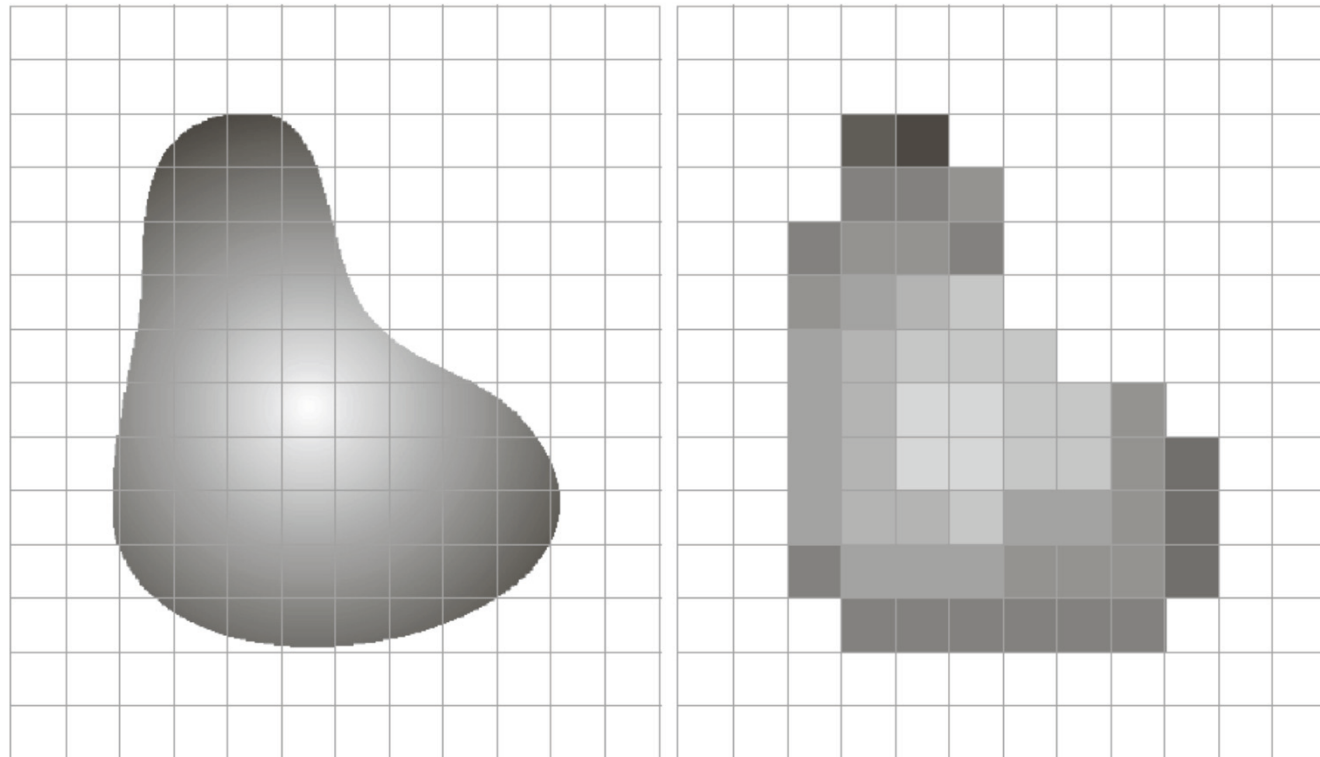
- ✓ Coordinadas espaciales:
Muestreo.

- ✓ Amplitud:
Cuantización.



a	b
c	d

FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from *A* to *B* in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

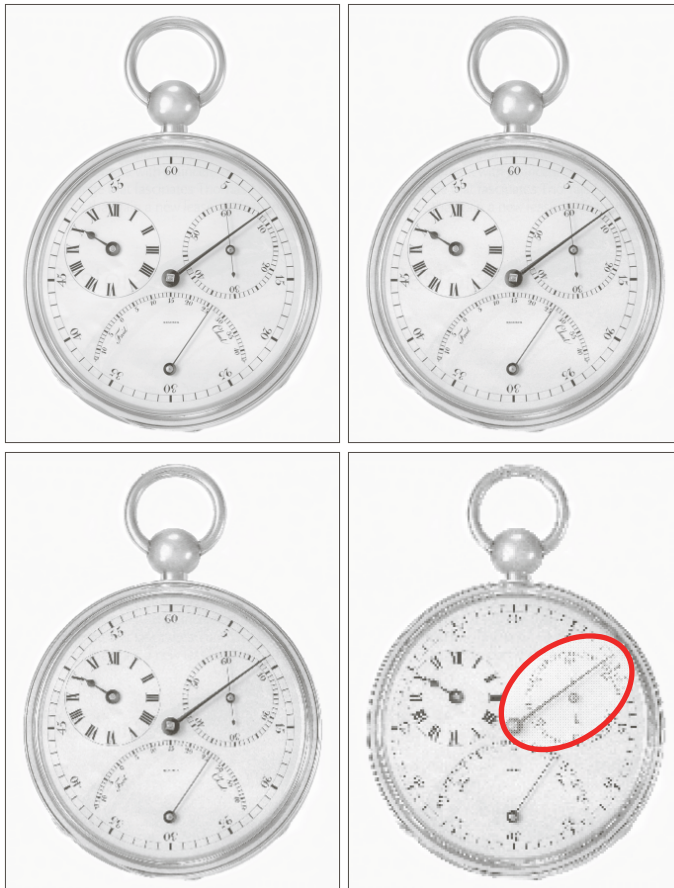


a b

FIGURE 2.17 (a) Continuous image projected onto a sensor array. (b) Result of image sampling and quantization.

Resolución espacial y resolución de intensidad

Resolución espacial



a b
c d

FIGURE 2.20 Typical effects of reducing spatial resolution. Images shown at: (a) 1250 dpi, (b) 300 dpi, (c) 150 dpi, and (d) 72 dpi. The thin black borders were added for clarity. They are not part of the data.

➤ Medida del detalle más pequeño discernible en una imagen.

✓ Pares de líneas por unidad de distancia. Ej: | |

w w

En este caso el ancho de un par de líneas es $2w$. Luego: $1/(2w)$ pares de líneas por unidad de distancia.

✓ Puntos (píxeles) por unidad de distancia. Ejemplo:
Dot per inch (dpi).

Valores típicos

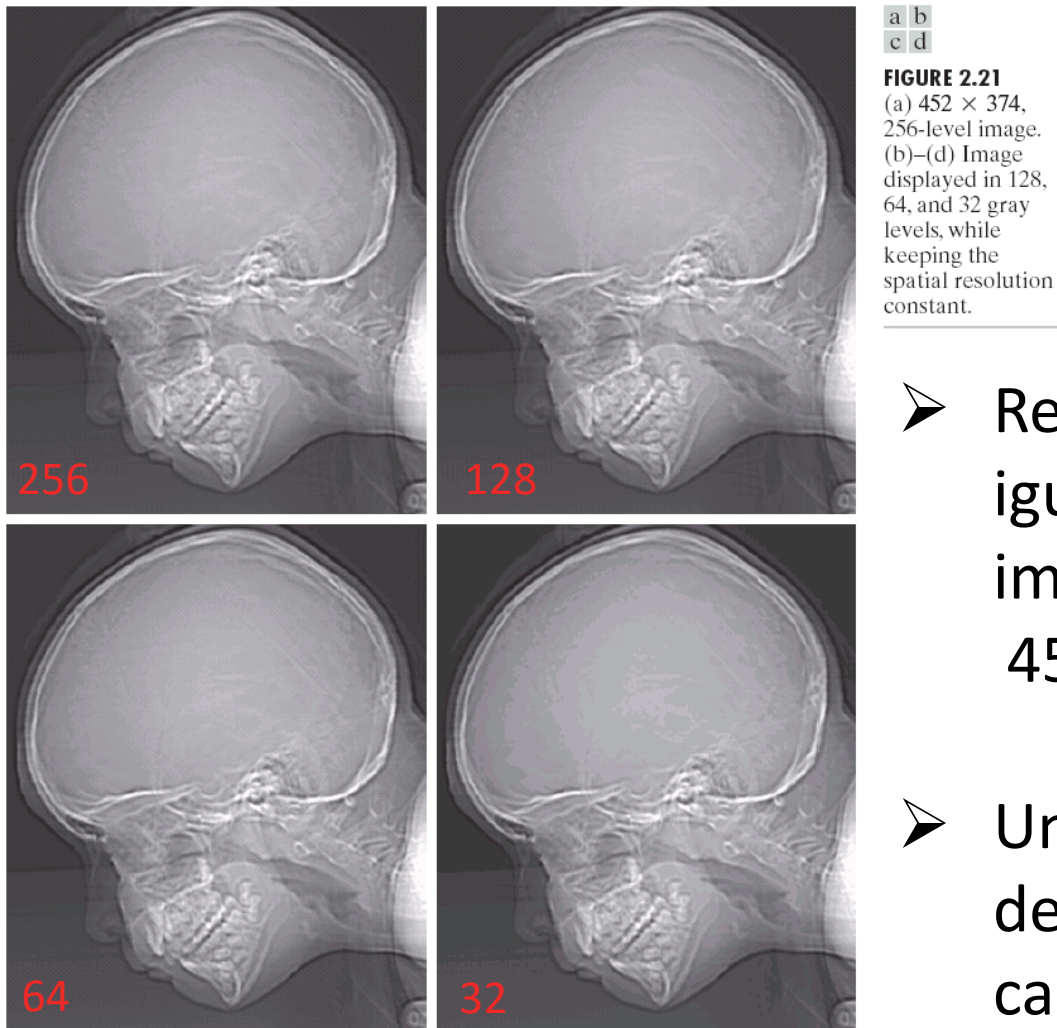
- Periódicos: 75 dpi.
- Revistas: 133 dpi.
- Folletos: 175 dpi.
- Libro con imágenes de alta calidad: 2400 dpi.

Resolución de la intensidad

- Medida del cambio discernible más pequeño en el nivel de intensidad de una imagen.
- Típicamente es una potencia de 2 por razones de construcción del sistema de captura. Ejemplo 2^8 , 2^{16} .
- Número de bits utilizados para la cuantización. Ej: 256 niveles \longrightarrow “8 bits de resolución”.

- La resolución espacial y de intensidad interactúan cambiando nuestra percepción de la calidad de una imagen digital.

Efectos de la resolución de intensidad



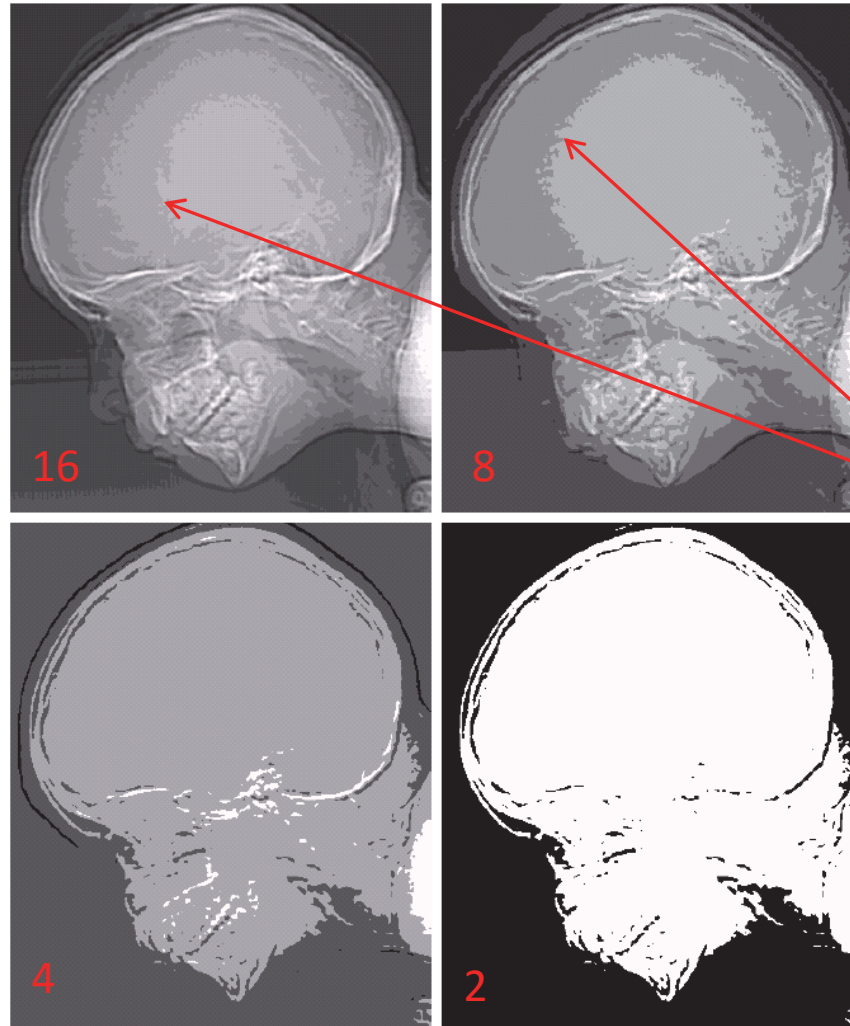
- Resolución espacial igual en todas las imágenes: 452×374 .
- Un número diferente de niveles de gris en cada imagen.

e	f
g	h

FIGURE 2.21

(Continued)

(e)–(h) Image displayed in 16, 8, 4, and 2 gray levels. (Original courtesy of Dr. David R. Pickens, Department of Radiology & Radiological Sciences, Vanderbilt University Medical Center.)



Falsos
contornos

Cambios simultáneos en la resolución espacial y de intensidad

Evaluación subjetiva de la calidad de las imágenes



a b c

FIGURE 2.22 (a) Image with a low level of detail. (b) Image with a medium level of detail. (c) Image with a relatively large amount of detail. (Image (b) courtesy of the Massachusetts Institute of Technology.)

Nivel de detalle

Curvas de igual calidad subjetiva

k : Número de bits usados para codificar la intensidad.

$N \times N$: Número de píxeles en la imagen.

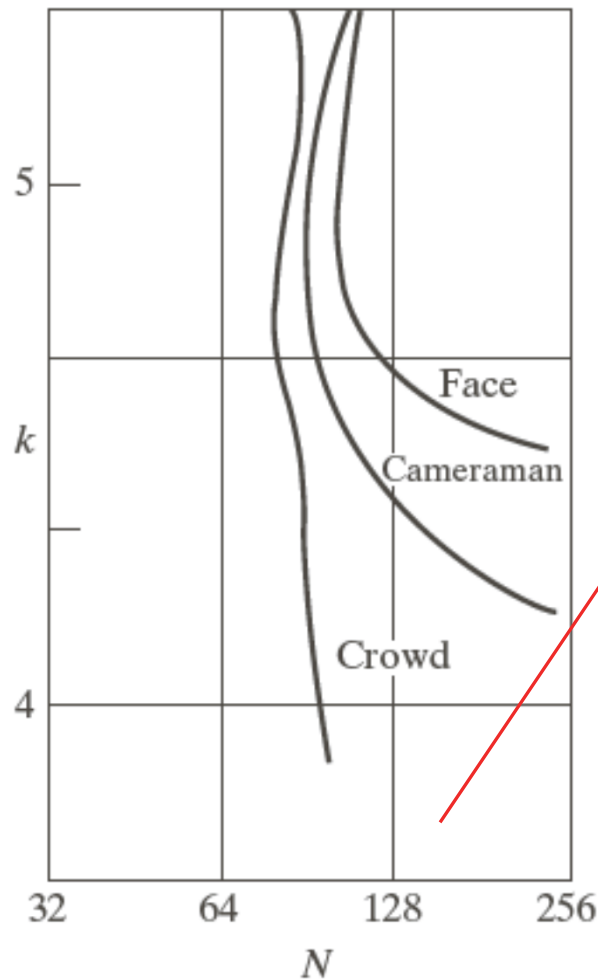
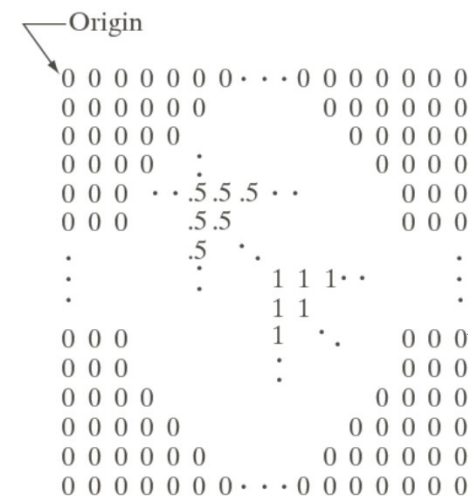
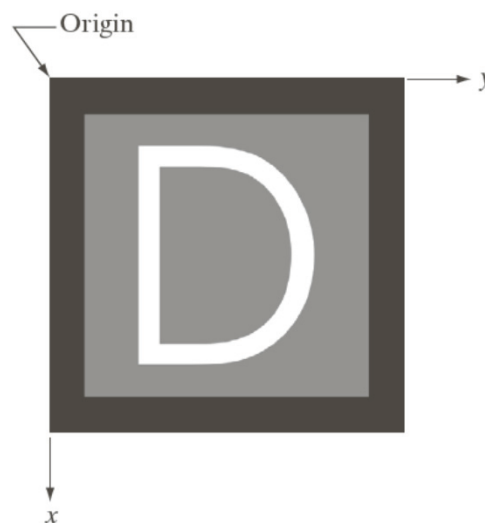
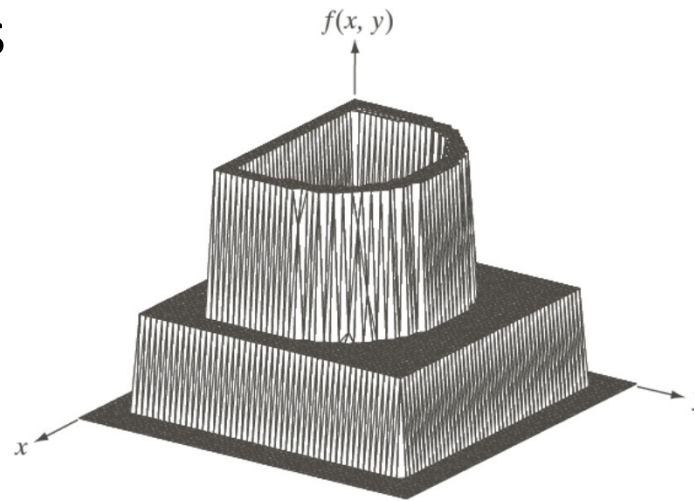


FIGURE 2.23
Typical isopreference curves for the three types of images in Fig. 2.22.

Mejor calidad

Representando imágenes digitales

- x, y : Variables espaciales o coordenadas espaciales.
- Dominio espacial.
- Índices enteros: (filas, columnas).
- Origen: arriba a la izquierda.
 - ✓ (0,0)
 - ✓ (1,1)
 - ✓ Depende del software.



a
b c

FIGURE 2.18
(a) Image plotted as a surface.
(b) Image displayed as a visual intensity array.
(c) Image shown as a 2-D numerical array (0, .5, and 1 represent black, gray, and white, respectively).

pixel

Rango dinámico y contraste

➤ **Rango dinámico:** Cociente de la máxima intensidad que se puede medir a la mínima intensidad que se puede detectar en el sistema.

- ✓ Max: Saturación.
- ✓ Min: Ruido.

➤ **Contraste:** Diferencia en intensidad entre el nivel de intensidad más alto y más bajo en una imagen.

$$L = 2^k$$

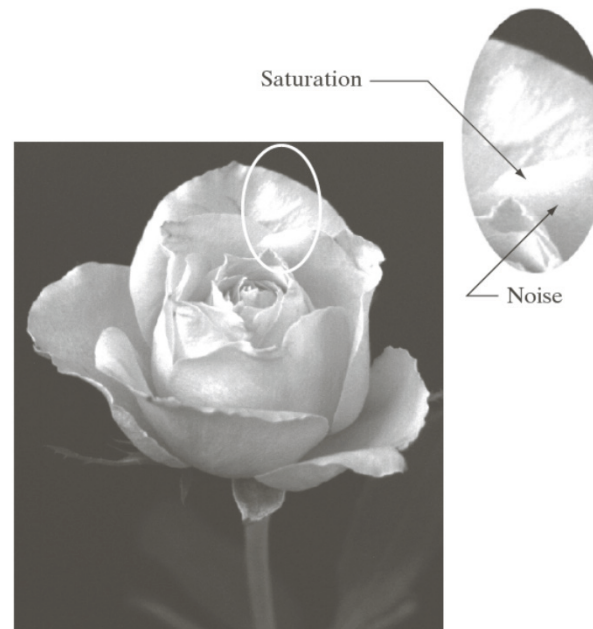


FIGURE 2.19 An image exhibiting saturation and noise. Saturation is the highest value beyond which all intensity levels are clipped (note how the entire saturated area has a high, *constant* intensity level). Noise in this case appears as a grainy texture pattern. Noise, especially in the darker regions of an image (e.g., the stem of the rose) masks the lowest detectable true intensity level.

Tamaño de imágenes digitales en escala de grises en **bits**

$$b = MNk$$

Donde:

b es el tamaño en bits.

M es el número de filas de la imagen.

N es el número de columnas de la imagen.

Tamaño de imágenes digitales a color en **bits**

$$b = cMNk$$

Donde:

b es el tamaño en bits.

c es el número de planos de color (típicamente es igual a 3).

M es el número de filas de la imagen.

N es el número de columnas de la imagen.

Tamaño de un video en escala de grises en **bits**

$$b = MNkft$$

Donde:

- b es el tamaño en bits.
- M es el número de filas de un fotograma.
- N es el número de columnas de un fotograma.
- f es el número de fotogramas/segundo o cuadros/segundo.
- t es el tiempo en segundos del video.

Tamaño de un video a color en **bits**

$$b = cMNkft$$

Donde:

- b es el tamaño en bits.
- c es el número de planos de color (típicamente es igual a 3).
- M es el número de filas de un fotograma.
- N es el número de columnas de un fotograma.
- f es el número de fotogramas/segundo o cuadros/segundo.
- t es el tiempo en segundos del video.

Prefijos en el sistema binario y en el sistema decimal

Nombre	Símbolo	Potencias binarias y valores decimales	Valores en el SI
Kilo	K	$2^{10} = 1\,024$	$10^3 = 1\,000$
Mega	M	$2^{20} = 1\,048\,576$	$10^6 = 1\,000\,000$
Giga	G	$2^{30} = 1\,073\,741\,824$	$10^9 = 1\,000\,000\,000$
Tera	T	$2^{40} = 1\,099\,511\,627\,776$	$10^{12} = 1\,000\,000\,000\,000$
Peta	P	$2^{50} = 1\,125\,899\,906\,842\,624$	$10^{15} = 1\,000\,000\,000\,000\,000$
Exa	E	$2^{60} = 1\,152\,921\,504\,606\,846\,976$	$10^{18} = 1\,000\,000\,000\,000\,000\,000$
Zetta	Z	$2^{70} = 1\,180\,591\,620\,717\,411\,303\,424$	$10^{21} = 1\,000\,000\,000\,000\,000\,000\,000$
Yotta	Y	$2^{80} = 1\,208\,925\,819\,614\,629\,174\,706\,176$	$10^{24} = 1\,000\,000\,000\,000\,000\,000\,000\,000$

Nota: Recuerde que 1 byte = 8 bits

Ejercicios: tamaño de imágenes y videos sin comprimir

1. Desarrolle una ecuación general para calcular el tamaño de una imagen en escala de grises en kilobytes (kB).
2. Desarrolle una ecuación que le permita calcular el tamaño de una imagen a color en Megabytes (MB).
3. Desarrolle una ecuación general para calcular el tamaño de una imagen en escala de grises en Megabytes (MB) a partir del número de píxeles de la imagen y el número de bits usados en el proceso de cuantización.

4. Desarrolle una función en Matlab que permita calcular el tamaño de una imagen sea esta en escala de grises o a color según lo indique el usuario. La función debe entregar el tamaño en: bits, kb, kB, y MB.

5. Desarrolle una función en Matlab que permita calcular el tamaño de un video en escala de grises o a color según lo indique el usuario. La función debe entregar el tamaño en: kB, MB, y GB.

Relaciones básicas entre pixeles

4-vecinos de un pixel

- p, q : pixeles.
- Un pixel p con coordenadas (x, y) tiene 4 vecinos horizontales y verticales con coordenadas:

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

- Este conjunto de pixeles
 - ✓ Se llama 4-vecinos de p .
 - ✓ Se denota $N_4(p)$.

4-vecinos diagonales de un pixel

- Los 4 vecinos diagonales de p tienen coordenadas:

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

- Se denotan $N_D(p)$.

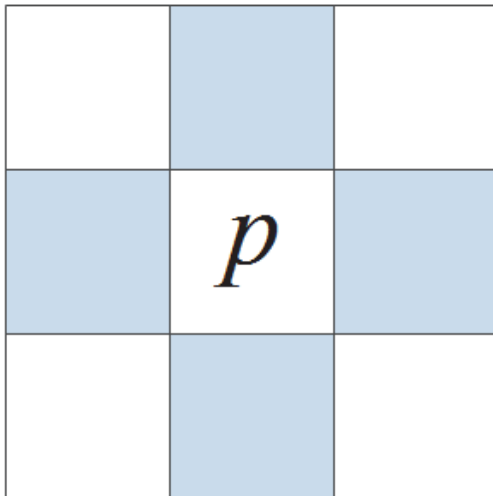
8-vecinos de un pixel

- Combinando los pixeles horizontales, verticales y diagonales se obtienen los 8-vecinos del pixel p : $N_8(p)$.

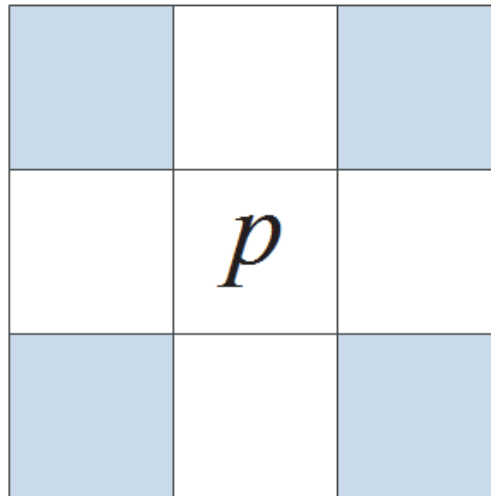
$$N_8(p) = N_4(p) \cup N_D(p)$$

Vecinos de un pixel

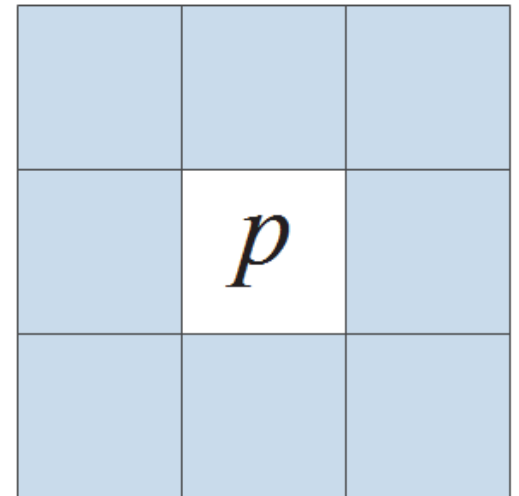
$$N_4(p)$$



$$N_D(p)$$



$$N_8(p)$$

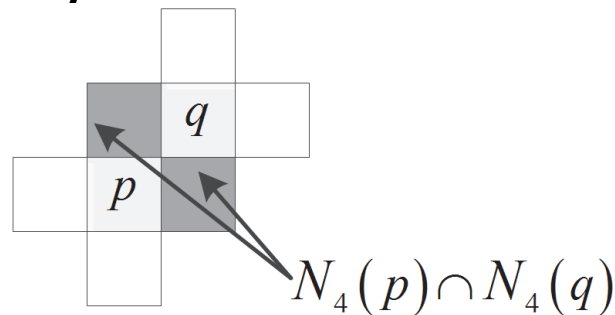


Adyacencia

- V : Conjunto de valores de intensidad usado para definir la adyacencia.
- En una imagen binaria $V=\{1\}$ si nos referimos a la adyacencia de pixeles con valor 1.
- En una imagen en escala de grises V tiene más elementos.
- Si la intensidad de una imagen tiene un rango de $[0,255]$, el conjunto V puede ser cualquier subconjunto de esos 256 valores.

Tipos de adyacencia

- Dos pixeles p y q con valores de V tienen
- **Adyacencia-4** si q está en $N_4(p)$.
- **Adyacencia-8** si q está en $N_8(p)$.
- **Adyacencia-m**, o adyacencia mixta si:
 - ✓ q está en $N_4(p)$ **O**
 - ✓ q está en $N_D(p)$ **Y** el conjunto $N_4(p) \cap N_4(q)$ **NO** tiene pixeles cuyos valores estén en V .



Trayectoria o curva digital

- Una trayectoria o curva digital desde un pixel p con coordenadas (x,y) hasta un pixel q con coordenadas (s,t) es una secuencia de pixeles distintos con coordenadas $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$

Donde: $(x_0, y_0) = (x, y)$

$$(x_n, y_n) = (s, t)$$

y los pixeles (x_i, y_i) y (x_{i-1}, y_{i-1}) son adyacentes para $1 \leq i \leq n$. Donde n es la longitud de la trayectoria

- Si $(x_0, y_0) = (x_n, y_n)$ la trayectoria es cerrada.

- Dependiendo del tipo de adyacencia:
 - ✓ Trayectoria-4.
 - ✓ Trayectoria-8.
 - ✓ Trayectoria-m.

Conexión

- Sea S un subconjunto de píxeles en una imagen. Decimos que dos píxeles p y q están conectados si existe una trayectoria entre ellos formada únicamente con píxeles de S .
- Para cualquier píxel p en S , el conjunto de píxeles que están conectados a él en S , se conoce como componente conexo de S .
- Si S solo tiene un componente conexo, entonces S se conoce como conjunto conexo.

Región

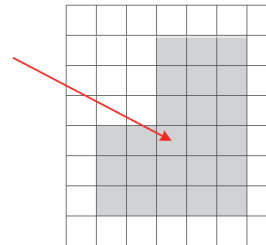
- Sea R un subconjunto de píxeles en una imagen.
- Decimos que R es una **región** de la imagen si R es un **conjunto conexo**.
- Dos **regiones** R_i y R_j son **adyacentes** si su unión forma un conjunto conexo.
- Las regiones que no son adyacentes, se conocen como **regiones disjuntas**.
- **Se debe especificar el tipo de adyacencia.**

Primer plano (*foreground*) y fondo (*background*)

- Suponga que una imagen contiene K regiones disjuntas R_k con $k=1,2,3,\dots,K$ y que ninguna de ellas toca el borde de la imagen.
- Sea R_u la unión de todas las K regiones y denotemos por $(R_u)^c$ su complemento.
 - ✓ **Primer plano:** Todos los puntos en R_u .
 - ✓ **Fondo de la imagen:** todos los puntos en $(R_u)^c$.

Contorno o frontera interna de una región R

- Conjunto de puntos que son adyacentes a los puntos en el complemento de R .
- El contorno de una región es el conjunto de píxeles en la región que tienen al menos un píxel vecino en el fondo de la imagen.
- Se debe especificar la conectividad que se usa para definir la adyacencia. Por lo general se usa **adyacencia-8** para evitar situaciones como



Contorno o frontera externa de una región

- Contorno correspondiente pero en el fondo de la imagen.
- Permite obtener contornos que formen trayectorias cerradas.

Para tener en cuenta...

- Si R corresponde a una imagen completa, entonces su contorno se **define** como el conjunto de píxeles en la primera y última fila y columna de la imagen.

Bordes

- Un borde se forma a partir de pixeles cuyos valores de derivadas exceden un umbral predeterminado. Pensar en discontinuidades de intensidad.
- No confundir el contorno (*boundary*), que es un concepto global, con el borde (*edge*) que es un concepto local .
- Los bordes y contornos en general sólo corresponden en imágenes binarias dependiendo de la conectividad y el operador de borde que se use.

Trayectoria con ambigüedad para adyacencia 8 usando

$V=\{1\}$

Trayectoria sin
ambigüedad con
adyacencia m
usando $V=\{1\}$

Dos regiones sin
adyacencia 4
pero con
adyacencia 8
dado que $V=\{1\}$

```
0 1 1
0 1 0
0 0 1
```

```
0 1 1
0 1 0
0 0 1
```

```
0 1 1
0 1 0
0 0 1
```

```
1 1 1
1 0 1
0 1 0
0 0 1
1 1 1
1 1 1
```

R_i

R_j

```
0 0 0 0 0
0 1 1 0 0
0 1 1 0 0
0 1 1 1 0
0 1 1 1 0
0 0 0 0 0
```

```
0 0 0
0 1 0
0 1 0
0 1 0
0 1 0
0 0 0
```

No sería parte del
contorno si se usa
la adyacencia-4
para $V=\{1\}$

a	b	c
d	e	f

FIGURE 2.25 (a) An arrangement of pixels. (b) Pixels that are 8-adjacent (adjacency is shown by dashed lines; note the ambiguity). (c) m -adjacency. (d) Two regions that are adjacent if 8-adjacency is used. (e) The circled point is part of the boundary of the 1-valued pixels only if 8-adjacency between the region and background is used. (f) The inner boundary of the 1-valued region does not form a closed path, but its outer boundary does.

Medidas de distancia

- Sean $p(x,y)$, $q(s,t)$ y $z(v,w)$: pixeles.
- D es una función de distancia o métrica si se cumplen las siguientes tres condiciones:

$$D(p,q) \geq 0 \quad (D(p,q) = 0 \quad \text{si} \quad p = q)$$

$$D(p,q) = D(q,p)$$

$$D(p,z) \leq D(p,q) + D(q,z)$$

Distancias entre pixeles p y q

- Distancia Euclídea: $D_e(p, q) = \sqrt{(x-s)^2 + (y-t)^2}$
- Distancia D_4 : $D_4(p, q) = |x-s| + |y-t|$
- Distancia D_8 : $D_8(p, q) = \max(|x-s|, |y-t|)$

Las distancias anteriores únicamente consideran las coordenadas de los pixeles p y q . El valor de intensidad de los pixeles no las afecta.

- Distancia D_m : trayectoria-m más corta entre los puntos (considera la adyacencia-m).
 - ✓ Depende de los valores de los pixeles a lo largo de la trayectoria, así como de los valores de sus vecinos.

Preguntas??