

COPIA PRIVADA CON FINES DOCENTES

VISIÓN POR COMPUTADOR

ARTURO DE LA ESCALERA HUESO



4. Transformaciones matemáticas.

Las transformaciones matemáticas que se describen en el presente capítulo, son aquellos algoritmos de los que se obtiene una imagen $g(x,y)$ a partir de la original $f(x,y)$. El fin perseguido es favorecer alguna característica presente en la imagen original o eliminar alguna otra que esté ocultando lo que se busca en la imagen. Dos serán los tipos de transformaciones que pueden llevarse a cabo:

- Transformaciones en el dominio del espacio: convolución y correlación.
- Transformaciones en el dominio de la frecuencia: transformada de Fourier.

A continuación se describirán con más detalle ambos tipos de transformaciones. Aunque se estudian por separado, no son enfoques excluyentes. Además, como se verá al final del capítulo, los resultados pueden pasarse de un dominio a otro con las transformaciones correspondientes.

4.1. Procesamiento espacial

El procesamiento espacial lo constituyen aquellas técnicas que operan directamente sobre los valores de los píxeles de la imagen. Serán transformaciones del tipo:

$$S(x,y)=F(I(x,y))$$

siendo $I(x,y)$ la imagen original, $S(x,y)$ la resultante y F la transformación. Estas transformaciones suelen aplicarse a un entorno cuadrado del píxel.

4.1.1. Convoluciones.

Para el caso unidimensional, se define como convolución de la función $f(x)$ respecto a la función $h(x)$ una nueva función $g(x)$ tal que:

$$g(x) = h(x) * f(x) = \int_{i=-\infty}^{i=\infty} f(i) h(x-i) di$$

Si se tratara del caso discreto con dos secuencias $f(x)$ y $h(x)$ se obtendría una secuencia de salida $g(x)$:

$$g(x) = h(x) * f(x) = \sum_{i=-\infty}^{i=\infty} f(i) h(x-i)$$

En la figura 4.1 puede observarse la convolución entre dos funciones:

$$f(x) = \begin{cases} 1 & 0 < x < 4 \\ 0 & \text{en cualquier otro caso} \end{cases}$$
$$h(x) = \begin{cases} x & -1 < x < 1 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

Lo primero es obtener la función $h(-x)$ simétrica respecto al eje x . Esta señal se irá desplazando por el eje x , y el valor de la convolución para cada valor de x será el producto del área común a las dos señales. Así:

$$-1 < x < 1$$

$$g(x) = \int_0^{x+1} h(x-i) di = \int_0^{x+1} (x-i) di = \frac{1}{2} (x^2 - 1)$$

1 < x < 3

$$g(x) = \int_{x-1}^{x+1} (x-i) di = 0$$

3 < x < 5

$$g(x) = \int_{x-1}^4 (x-i) di = -\frac{x^2}{2} + 4x - \frac{15}{2}$$

La señal $g(x)$ irá decreciendo conforme $f(x)$ y $h(x)$ se vayan solapando, hasta alcanzar un mínimo cuando toda la parte negativa haya entrado. El efecto de la parte positiva de $h(x)$ implica que el valor de $g(x)$ aumenta hasta hacerse cero. Conforme su parte negativa abandona a $f(x)$, los valores serán positivos hasta alcanzar un máximo, tendiendo de nuevo a cero.

Para el caso de imágenes digitales se usan convoluciones bidimensionales discretas, cuya fórmula es:

$$g(x,y) = h(x,y) * f(x,y) = \sum_{i=-\infty}^{i=\infty} \sum_{j=-\infty}^{j=\infty} f(i,j) h(x-i, y-j)$$

Lo más normal es usar convoluciones de 3x3 elementos. Entonces la expresión anterior puede concretarse en:

$$g(x,y) = h(x,y) * f(x,y) = \sum_{i=0}^{i=2} \sum_{j=0}^{j=2} f(i,j) h(x-i, y-j)$$

que por ejemplo para $g(2,2)$ será:

$$g(2,2) = \sum_{i=0}^{i=2} \sum_{j=0}^{j=2} f(i,j) h(2-i, 2-j) =$$

$$f(0,0)h(2,2) + f(0,1)h(2,1) + f(0,2)h(2,0) +$$

$$f(1,0)h(1,2) + f(1,1)h(1,1) + f(1,2)h(1,0) +$$

$$f(2,0)h(0,2) + f(2,1)h(0,1) + f(2,2)h(0,0)$$

En el caso de la visión por computador, en lugar de hablar de convolucionar la imagen $f(x,y)$ con la transformación $h(x,y)$, se habla de convolucionar con una máscara que tiene la forma:

$h(0,0)$	$h(0,1)$	$h(0,2)$
$h(1,0)$	$h(1,1)$	$h(1,2)$
$h(2,0)$	$h(2,1)$	$h(2,2)$

Para obtener $h(-i,-j)$ hay que realizar dos giros. Así, se obtiene $h(i,-j)$:

$h(2,0)$	$h(2,1)$	$h(2,2)$
$h(1,0)$	$h(1,1)$	$h(1,2)$
$h(0,0)$	$h(0,1)$	$h(0,2)$

Y $h(-i,-j)$

$h(2,2)$	$h(2,1)$	$h(2,0)$
$h(1,2)$	$h(1,1)$	$h(1,0)$

$h(0,2)$	$h(0,1)$	$h(0,0)$
----------	----------	----------

La porción inicial de la imagen correspondiente es:

$f(0,0)$	$f(0,1)$	$f(0,2)$
$f(1,0)$	$f(1,1)$	$f(1,2)$
$f(2,0)$	$f(2,1)$	$f(2,2)$

y el valor del píxel de la imagen transformada será el mismo que se obtuvo antes.

Gráficamente se trata de poner las dos matrices, una encima de la otra, multiplicar cada celda por la inferior y sumar los productos. Para obtener el siguiente valor de $g(x,y)$ la máscara $h(x,y)$ se desplaza un elemento a la derecha y se repite la operación. Al acabar las columnas se vuelve al principio a una fila inferior, y así sucesivamente hasta llegar al último píxel.

De la figura 4.1 puede deducirse que el rango de la señal $g(x)$ será siempre más grande que el de $f(x)$. En el ejemplo $f(x)$ toma valores distintos de cero entre 1 y 4 mientras que $g(x)$ está entre -1 y 5. En el caso de imágenes no es así porque, por convenio, la operación se realiza para todos los píxeles de la imagen menos para los píxeles del borde, en los que no es posible aplicar la convolución al no solaparse totalmente la máscara y la imagen. Se tiene así un marco de anchura dos píxeles cuyo valor se suele tomar igual a cero. En el caso general, siempre que se realice una convolución se obtendrá una imagen más pequeña de dimensiones $(N-(n-1)) \times (M-(m-1))$, donde N y M son las dimensiones de la imagen original y n y m las de la máscara.

En el caso de la señal de salida $g(x)$ en 4.1 se observa que presenta dos picos, uno positivo y otro negativo, en los dos extremos de la señal $f(x)$. Ha servido por lo tanto para detectar los bordes de esa señal, o cuándo la señal $f(x)$ pasa de tomar un valor constante a otro. Es de esperar que si se aplica una convolución parecida en el caso bidimensional dé resultados análogos. En la figura 4.2 se ve el caso de aplicar la siguiente máscara:

-1	0	1
-1	0	1
-1	0	1

Queda ahora ver como se representan los píxeles con valores negativos. Como en la pantalla todos los píxeles deben ser positivos, primero se normalizan los resultados entre -128 y 127, para después desplazar el origen 128 posiciones a la derecha. Así en el caso de la figura 4.2 derecha, píxeles oscuros corresponden a valores negativos, los claros a positivos, y los intermedios a valores en torno al cero.

4.1.2. Correlación.

La correlación $g(x)$ entre dos funciones $f(x)$ y $h(x)$ es:

$$g(x, y) = h(x, y) \circ f(x, y) = \sum_{i=-\infty}^{i=\infty} \sum_{j=-\infty}^{j=\infty} f^*(i, j)h(x + i, y + j)$$

donde f^* es el complejo conjugado¹. En la imagen 4.3 puede verse la correlación entre dos señales. El funcionamiento es parecido al de la convolución. Así, si se tienen las funciones:

$$f(x) = \begin{cases} 1 & 0 < x < 1 \\ 2 & 1 < x < 3 \\ 1 & 3 < x < 4 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

$$h(x) = \begin{cases} 2 & -1 < x < 1 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

Entonces, el resultado es:

$$-1 < x < 0$$

$$g(x) = \int_0^{x+1} 2 \cdot 1 \, di = 2(x+1)$$

$$0 < x < 1$$

$$g(x) = \int_0^1 2 \cdot 1 \, di + \int_1^{x+1} 2 \cdot 2 \, di = 2 + 4x$$

$$1 < x < 2$$

$$g(x) = \int_{x-1}^1 2 \cdot 1 \, di + \int_1^{x+1} 2 \cdot 2 \, di = 4 + 2x$$

$$2 < x < 3$$

$$g(x) = \int_{x-1}^3 2 \cdot 2 \, di + \int_3^{x+1} 2 \cdot 1 \, di = 12 - 2x$$

$$3 < x < 4$$

$$g(x) = \int_{x-1}^3 2 \cdot 2 \, di + \int_3^4 2 \cdot 1 \, di = 18 - 4x$$

$$4 < x < 5$$

$$g(x) = \int_{x-1}^4 2 \cdot 1 \, di = 2(5-x)$$

Observando el resultado anterior puede deducirse que una posible aplicación de la correlación es la localización de un patrón dentro de una imagen, ya que el valor donde la correlación es máxima corresponde a las coordenadas donde ese patrón se encuentra (figura 4.4). Si embargo si se toma el segundo caso de la misma figura se ve que si un objeto engloba a otro dará la misma respuesta. Otro caso en el que la correlación fallaría sería si buscamos un objeto grande y oscuro si también está presente otro objeto más pequeño, pero mucho más claro. Para solucionar esto se tiene la *correlación normalizada*, que evita además posibles cambios en los niveles de gris entre el modelo que se busca y el objeto presente en la imagen.

¹ En el caso de imágenes, al ser números reales f^* es igual a f .

$$g(x,y) = \frac{\sum_i \sum_j (f(x+i, y+i) - \bar{f})(h(i, j) - \bar{h})}{\sqrt{\sum_j \sum_j (f(x+i, y+i) - \bar{f})^2 \sum_j \sum_j (h(i, j) - \bar{h})^2}}$$

En la figura 4.5 puede verse el uso de la correlación normalizada. En este caso se trata de buscar un componente electrónico en la imagen. Puede observarse que la correlación no es inmune a cambios en la escala, ni a la rotación del objeto.

4.2. Dominio de la frecuencia.

4.2.1 Series de Fourier; Error! Marcador no definido..

Dada una función $f(t)$ periódica en el tiempo, la lógica lleva a concluir que puede expresarse como la suma de otras funciones periódicas. Las series de Fourier dan la fórmula matemática que expresa esa relación con las funciones periódicas seno y coseno. Así, cualquier función $f(t)$ de periodo T_0 puede expresarse como:

$$f(t) = a_0 + \sum_{k=1}^{\infty} a_k \cos k\omega_0 t + b_k \sen k\omega_0 t$$

Donde:

$$a_0 = \frac{1}{T_0} \int_0^{T_0} f(t) dt$$

$$a_k = \frac{2}{T_0} \int_0^{T_0} f(t) \cos k\omega_0 t dt$$

$$b_k = \frac{2}{T_0} \int_0^{T_0} f(t) \sen k\omega_0 t dt$$

Siendo T_0 el periodo y $\omega_0 = \frac{2\pi}{T_0}$

Dada la función de la figura 4.6.a:

$$f(t) = \begin{cases} 1 & 0 < t < \pi \\ -1 & \pi < t < 2\pi \end{cases}$$

puede descomponerse en el sumatorio:

$$f(t) = \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{\sen(2k+1)t}{2k+1}$$

Como puede verse en la figura 4.7, conforme se añaden componentes al sumatorio la señal se aproxima a la original. También puede observarse cómo el nivel de detalle mejora conforme la frecuencia es más alta. En este ejemplo puede verse que se han necesitado dos informaciones: las correspondientes a los coeficientes de los senos y cosenos.

Si ahora se toma una nueva señal (figura 4.6.b):

$$f(t) = \begin{cases} -1 & 0 < t < \frac{\pi}{2} \\ 1 & \frac{\pi}{2} < t < \frac{3\pi}{2} \\ -1 & \frac{3\pi}{2} < t < 2\pi \end{cases}$$

puede descomponerse en el sumatorio:

$$f(t) = \frac{-4}{\pi} \sum_{k=0}^{\infty} \frac{\cos(2k+1)t}{2k+1}$$

donde ahora el coeficiente para los senos sería cero. Si se quisiera encontrar una expresión similar a la primera, teniendo en cuenta que:

$$-\cos \vartheta = \operatorname{sen}\left(\vartheta - \frac{\pi}{2}\right)$$

se tendría:

$$f(t) = \frac{-4}{\pi} \sum_{k=0}^{\infty} \frac{\cos(2k+1)t}{2k+1} = \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{\operatorname{sen}\left((2k+1)t - \frac{\pi}{2}\right)}{2k+1}$$

Que indica que las senoides necesarias son las mismas en ambos casos, sólo que en el segundo están desfasadas noventa grados. De nuevo se han necesitado dos coeficientes: el de los senos y el de los desfases.

Si la función fuese (figura 4.6.c):

$$f(t) = \begin{cases} 1 & 0 < t < \pi \\ -0.5 & \pi < t < 2\pi \end{cases}$$

puede descomponerse en el sumatorio:

$$f(t) = \frac{1}{4} + \frac{3}{\pi} \sum_{k=0}^{\infty} \frac{\operatorname{sen}(2k+1)t}{2k+1}$$

Ahora ha aparecido un nuevo valor, 1/4, que es el nivel medio de la señal durante el periodo.

Si ahora se desplaza de nuevo la señal en el tiempo (figura 4.6.d):

$$f(t) = \begin{cases} -0.5 & 0 < t < \frac{\pi}{4} \\ 1 & \frac{\pi}{4} < t < \frac{5\pi}{4} \\ -0.5 & \frac{5\pi}{4} < t < 2\pi \end{cases}$$

puede descomponerse en el sumatorio:

$$f(t) = \frac{1}{4} + \frac{3}{\pi} \sum_{k=0}^{\infty} \frac{1}{2k+1} \left[-\operatorname{sen}\left(\frac{(2k+1)\pi}{4}\right) \cos((2k+1)t) + \cos\left(\frac{(2k+1)\pi}{4}\right) \operatorname{sen}((2k+1)t) \right]$$

como:

$$\operatorname{sen}(A-B) = \operatorname{sen}A \cos B - \operatorname{sen}B \cos A$$

se puede reescribir:

$$f(t) = \frac{1}{4} + \frac{3}{\pi} \sum_{k=0}^{\infty} \frac{1}{2k+1} \operatorname{sen}\left((2k+1)\left(t - \frac{\pi}{4}\right)\right)$$

que es el mismo caso anterior, sólo que ahora las senoides están retrasadas 45°.

Otra posibilidad es poner la serie de Fourier en notación exponencial. Para ello se toman en cuenta las ecuaciones de Euler:

$$\cos k\omega_0 t = \frac{e^{jk\omega_0 t} + e^{-jk\omega_0 t}}{2} \quad \text{sen } k\omega_0 t = \frac{e^{jk\omega_0 t} - e^{-jk\omega_0 t}}{2}$$

Así los dos últimos casos quedarían:

$$f(t) = \frac{1}{4} + \frac{3}{\pi} \sum_{k=0}^{\infty} \frac{\text{sen}(2k+1)t}{2k+1} = \frac{1}{4} + \sum_{k=-\infty}^{\infty} \frac{3}{2\pi(2k+1)} e^{-j\frac{\pi}{2}} e^{j(2k+1)t}$$

siendo $\frac{3}{\pi(2k+1)}$ el módulo y el argumento un retraso de 90°

Para la otra función:

$$f(t) = \frac{1}{4} + \frac{3}{\pi} \sum_{k=0}^{\infty} \frac{\text{sen}(2k+1)(t - \pi/4)}{2k+1} = \frac{1}{4} + \sum_{k=-\infty}^{\infty} \frac{3}{2\pi(2k+1)} e^{-j(\frac{\pi}{2} + (2k+1)\frac{\pi}{4})} e^{j(2k+1)t}$$

y en general, cualquier función periódica puede expresarse como:

$$f(t) = \sum_{k=-\infty}^{\infty} c_k e^{jk\omega_0 t}$$

donde:

$$c_k = \frac{1}{T_0} \int_0^{T_0} f(t) e^{-jk\omega_0 t} dt$$

Necesitando, de nuevo, calcular dos coeficientes: el módulo y el argumento.

4.2.2. La transformada de Fourier.

La transformada de Fourier es una extensión de las series de Fourier a señales no periódicas. Para ello se toma un nuevo coeficiente c'_k que es igual a:

$$c'_k = T_0 c_k = \int_0^{T_0} f(t) e^{-jk\omega_0 t} dt$$

Por otro lado:

$$f(t) = \frac{1}{T_0} \sum_{k=-\infty}^{\infty} c'_k e^{jk\omega_0 t} = \frac{\omega_0}{2\pi} \sum_{k=-\infty}^{\infty} c'_k e^{jk\omega_0 t}$$

ya que: $\omega_0 T_0 = 2\pi$

Como las señales ahora no son periódicas, se supone que su periodo abarca desde el menos infinito al infinito, luego:

$$c'_k = F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{j\omega t} d\omega$$

considerando que:

$$k\omega_0 = \omega$$

$$\omega_0 \rightarrow d\omega \quad \text{cuando} \quad T_0 \rightarrow \infty$$

Para el caso unidimensional discreto se define como:

$$F(\omega) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) e^{-j2\pi\omega x/N}$$

y su antitransformada:

$$f(\omega) = \sum_{\omega=0}^{N-1} F(\omega) e^{-\frac{j2\pi\omega x}{N}}$$

y para el caso bidimensional:

$$F(u,v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) e^{-\frac{j2\pi(ux+vy)}{N}}$$

$$f(x,y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u,v) e^{\frac{j2\pi(ux+vy)}{N}}$$

En la figura 4.8 pueden verse las transformadas de Fourier para una señal que toma un valor constante para un intervalo y cero fuera de él. En la figura 4.9 se ven dos variaciones de este caso. Así, se observa que si el intervalo tiende a infinito (señal continua), la transformada de Fourier es un único valor localizado en cero ya que solo tiene esta frecuencia. Si la señal tiende a la delta de Dirac (señal de amplitud infinita, duración cero y área unidad) se necesitan todas las frecuencias posibles. No se debe olvidar que la transformada de Fourier es un número complejo, y por tanto quedará definido por su módulo y argumento. El primero dará la amplitud de la senoide para una determinada frecuencia y el segundo su desfase respecto al origen. Debido a que generalmente el módulo toma un valor muy alto para frecuencias bajas se suele mostrar su logaritmo:

$$D(u,v) = c \log[1 + |F(u,v)|]$$

Por último, en la figura 4.10 se muestra la transformada para varias figuras geométricas.

4.2.3. Teorema del muestreo.

Supóngase que se quiere captar por la cámara el objeto presente en la figura 4.11-a. Vamos a suponer que se utiliza iluminación a contra luz por lo que el objeto está oscuro y el fondo claro. Si el objeto cae justo debajo de un píxel de la cámara, su valor será cero y 255 el de sus vecinos. Sin embargo es difícil que el objeto se encuentre siempre justo debajo de un píxel; si se toma el caso (b), ahora está justo debajo del punto común de cuatro píxeles. Cada uno de ellos recibe por tanto la luz a través de $\frac{3}{4}$ del píxel y nada por parte del otro cuarto; el nivel de gris resultante será 192. Desde el punto de vista de la imagen, se tendría igual resultado si se quisiera analizar un objeto cuatro veces más grande pero con un nivel de gris en la imagen de 192.

La solución más inmediata a tomar para distinguir estos dos objetos es aumentar la resolución de la imagen, como puede verse en los tres casos siguientes de la misma figura.

Por tanto las dimensiones de la imagen afectan a la apariencia de los objetos en ella, como ya se indicó en el capítulo anterior. Este efecto que se acaba de ver en el dominio del espacio tiene una demostración en el de la frecuencia, que constituye el teorema del muestreo. A la hora de captar una imagen estamos muestreando una señal continua, con lo que se convierte en otra discreta y se puede perder información. El teorema del muestreo indica cuánta información se pierde y lo crítico de tal pérdida.

La transformada discreta de Fourier de una señal abarca el rango de frecuencia entre $-\frac{\pi}{T}$ y $\frac{\pi}{T}$ siendo T el periodo de muestreo (en el caso de imágenes, la inversa de sus dimensiones espaciales N). Es una señal periódica, es decir que se repite su

forma para el resto de frecuencias. La forma es la misma para la señal continua que para la discreta; solamente que esta última tiene un factor de escala $1/T$ (figura 4.12)

Para algunas señales continuas su transformada de Fourier vale cero a partir de una frecuencia determinada ω_0 . Es lo que se denomina señales de banda limitada. Si se muestrean con un periodo que verifique:

$$\frac{2\pi}{T} > 2\omega_0$$

o lo que es lo mismo:

$$T < \frac{\pi}{\omega_0}$$

no se habrá perdido información. Si no ocurriese así, al ser la transformada discreta una señal periódica, se solaparía para algunas frecuencias por lo que sí se produciría pérdida de información (figura 4.13). Por supuesto que siempre se tiene este efecto para aquellas señales cuya transformada de Fourier no es de banda limitada, es decir, para aquellas que nunca toman valores nulos a partir de una frecuencia determinada.

Otra manifestación de lo que se está diciendo pueden ser los patrones de Moiré, en donde se tienen dos patrones transparentes que con un movimiento relativo producen interferencias entre uno y otro (figura 4.14). Este efecto se puede ver por ejemplo en la televisión, cuando aparecen patrones con detalles finos o líneas paralelas muy juntas.

4.2.4. Algunas propiedades de la transformada de Fourier.

Traslación.

Si una imagen $f(x,y)$ sufre un desplazamiento x_0, y_0 , la nueva transformada de Fourier es:

$$f(x + x_0, y + y_0) \Leftrightarrow G(u, v) = F(u, v) e^{\frac{-j2\pi(ux_0 + vy_0)}{N}}$$

Es fácil comprobar que:

$$\frac{G(u, v)F^*(u, v)}{|F(u, v)|^2} = e^{\frac{-j2\pi(ux_0 + vy_0)}{N}}$$

cuya antitransformada es una delta de Dirac en el punto (x_0, y_0) . Aplicando esta propiedad puede obtenerse el movimiento de un objeto dentro de una imagen (figura 4.15).

Rotación.

Si se cambia la notación de una imagen de las coordenadas cartesianas $f(x,y)$ a las polares $f(r, \theta)$:

$$\begin{aligned} x &= r \cos \theta & y &= r \sin \theta \\ u &= \varpi \cos \phi & v &= \varpi \sin \phi \end{aligned}$$

por lo que obtienen $f(r, \theta)$ y $F(\varpi, \phi)$. Si ahora se rota la imagen un cierto ángulo los cambios en Fourier son exactamente los mismos (figura 4.15).

$$f(r, \theta + \varphi) \Leftrightarrow F(\varpi, \phi + \varphi)$$

Separabilidad.

La transformada discreta de Fourier puede expresarse como:

$$F(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} e^{-j2\pi ux/N} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi vy/N}$$

$$f(x, y) = \sum_{u=0}^{N-1} e^{j2\pi ux/N} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi vy/N}$$

que lo que indica es que se puede realizar la transformada unidimensional primero para las columnas y luego para las filas

Propiedad distributiva

De la definición puede comprobarse fácilmente que:

$$F(f(x, y) + g(x, y)) = F(f(x, y)) + F(g(x, y))$$

Escalado

$$a f(x, y) \Leftrightarrow a F(u, v)$$

$$f(ax, by) \Leftrightarrow \frac{1}{|ab|} F(u/a, v/b)$$

Valor medio

Un valor muy utilizado en el preprocesamiento de imágenes es el valor medio del entorno de un punto.

$$\bar{f}(x, y) = \frac{1}{N^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y)$$

como el valor de la TDF para el origen es

$$F(0,0) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y)$$

se tiene que

$$\bar{f}(x, y) = \frac{1}{N} F(0,0)$$

4.2.5. La transformada rápida de Fourier.

Como su nombre indica esta transformada tiene como finalidad reducir el número de operaciones necesarias para realizar el paso al dominio de la frecuencia. La restricción es que la dimensión debe ser potencia de 2. Así se tiene que:

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) W_N^{ux}$$

donde:

$$W_N = e^{-j2\pi/N}$$

$$N = 2^n = 2M$$

por lo que se puede reescribir:

$$F(u) = \frac{1}{2M} \sum_{x=0}^{N-1} f(x) W_N^{ux} = \frac{1}{2} \left[\frac{1}{M} \sum_{x=0}^{M-1} f(2x) W_{2N}^{u2x} + \frac{1}{M} \sum_{x=0}^{M-1} f(2x+1) W_{2N}^{u(2x+1)} \right]$$

y tal como se ha definido W_N

$$W_{2M}^{2ux} = W_M^{ux}$$

por lo que:

$$F(u) = \frac{1}{2} \left[\frac{1}{M} \sum_{x=0}^{M-1} f(2x) W_N^{ux} + \frac{1}{M} \sum_{x=0}^{M-1} f(2x+1) W_N^{ux} W_{2N}^u \right]$$

denominando:

$$F_{par}(u) = \frac{1}{M} \sum_{x=0}^{M-1} f(2x) W_N^{ux}$$

$$F_{impar}(u) = \frac{1}{M} \sum_{x=0}^{M-1} f(2x+1) W_N^{ux}$$

se tiene que:

$$F(u) = \frac{1}{2} \left[F_{par}(u) + F_{impar}(u) W_{2N}^u \right]$$

y ya que:

$$W_M^{u+M} = W_M^u$$

$$W_{2M}^u = -W_{2M}^u$$

$$F(u+M) = \frac{1}{2} \left[F_{par}(u) - F_{impar}(u) W_{2N}^u \right]$$

por lo tanto el cálculo de la primera mitad de los puntos de N, F(u), se obtiene con la primera expresión, y la otra mitad con la segunda, de los que ya se conoce Fimpar y Fpar. Como se ha comentado al principio, se puso la restricción de que las dimensiones fueran potencia de 2. Esto es porque el mismo razonamiento se puede hacer ahora para Fimpar y Fpar, que se obtendrían entonces a partir de otras dos transformaciones. Lo mismo se haría con estas, con lo que al final se realizaría un algoritmo recursivo. Para terminar de acelerar la transformada se realiza primero por columnas y después por filas.

4.3. Relaciones entre las transformadas.

Las relaciones entre convolución y Fourier son:

$$f(x,y) * g(x,y) \Leftrightarrow F(u,v) \cdot G(x,y)$$

$$f(x,y) \cdot g(x,y) \Leftrightarrow F(u,v) * G(x,y)$$

La relación entre correlación y Fourier:

$$f(x,y) \circ g(x,y) \Leftrightarrow F^*(u,v) \cdot G(x,y)$$

$$f^*(x,y) \cdot g(x,y) \Leftrightarrow F(u,v) \circ G(x,y)$$

4.4 Operaciones Aritméticas, lógicas y geométricas.

4.4.1. Operaciones aritméticas y lógicas.

Las operaciones aritméticas y lógicas entre píxeles son usadas en el procesamiento de imágenes. Las operaciones aritméticas entre dos píxeles p y q son las siguientes:

- Suma
- Resta
- Multiplicación
- División

Si a la imagen de la figura 4.16-a, una pirámide que va desde el valor de gris mínimo al máximo, se le suma una imagen de nivel de gris constante 127 se obtiene la imagen de la figura 4.16-b, y si se le resta, la de la 4.16-b. Los resultados no son los esperados. Si a una imagen se le suma un valor de gris constante se espera que sea más clara y si se le resta más oscura. Observando las dos imágenes resultado se ve que no se parecen a la imagen original y además son la misma. El que el sumar o restar un valor distinto de cero produzca el mismo resultado indica que hay algo que se está realizando mal. Esto es debido al efecto *pixel overflow* o desbordamiento del píxel. Las imágenes tienen una resolución que suele ser de 8 bits por lo que el valor más alto es 255. Si por efecto de sumar un valor se sobrepasa ese máximo, el píxel tendrá, como nivel de gris, el valor menos 255. Así se ve que la imagen es una pirámide que va aumentando su valor. Al llegar a 255 se para a cero y se sigue aumentando el valor. Con esto se explica por qué la imagen 4.16-b no representa realmente la suma, pero ¿cómo la imagen de la suma y la resta son iguales? Esto es debido a que al trabajar con imágenes se supone que todos los píxeles tiene valores positivos, ya que todos los puntos reflejan más o menos luz pero ninguno la absorbe. Sin embargo al restar sí que se pueden alcanzar valores negativos. Y el problema radica en que al representarlos siempre se entienden como positivos. De todo esto se deduce que, después de una operación y antes de escribir los valores en la imagen, hay que normalizar los valores entre 0 y 255 (si el tamaño de los píxeles es de 8 bits.) tal como se ve en la figura 4.16-e. Otra opción es realizar las operaciones teniendo en cuenta la posible saturación del píxel, es decir, que no puede ser menor que cero o mayor que 255 (4.16-c y 4.16-d).

Las principales operaciones lógicas utilizadas en el procesamiento de imágenes son:

- AND
- OR
- NOT

Es preciso tener en cuenta que las operaciones lógicas sólo se pueden aplicar sobre imágenes binarizadas, lo cual no ocurre con las operaciones aritméticas. Un uso de las operaciones lógicas se verá en la segmentación por color

4.4.2. Transformaciones geométricas.

Las transformaciones geométricas modifican las relaciones espaciales entre los píxeles. Se necesitan dos tipos distintos de algoritmos. El primero es el que determina la relación entre las coordenadas de la imagen original y la imagen resultante. Como caso

general, las transformaciones de números enteros no tienen por qué dar enteros, así que será necesario un algoritmo de interpolación que determine el nivel de gris de la imagen final a partir de uno o varios píxeles de la imagen original.

Transformaciones

Traslación.

Si se quiere trasladar el origen de la imagen se aplica la ecuación:

$$xf = xi + xo$$

$$yf = yi + yo$$

Que en coordenadas homogéneas es:

$$\begin{bmatrix} xf \\ yf \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & xo \\ 0 & 1 & yo \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} xi \\ yi \\ 1 \end{bmatrix}$$

Magnificación

$$xf = \frac{xi}{a}$$

$$yf = \frac{yi}{b}$$

o lo que es lo mismo:

$$\begin{bmatrix} xf \\ yf \\ 1 \end{bmatrix} = \begin{bmatrix} 1/a & 0 & 0 \\ 0 & 1/b & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} xi \\ yi \\ 1 \end{bmatrix}$$

Rotación respecto al origen.

$$\begin{bmatrix} xf \\ yf \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\text{sen}\theta & 0 \\ \text{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} xi \\ yi \\ 1 \end{bmatrix}$$

Rotación respecto a un punto cualquiera.

Aquí se manifiesta la ventaja de la notación matricial, ya que es igual a trasladar la imagen al nuevo centro, rotar y deshacer la traslación:

$$\begin{bmatrix} xf \\ yf \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & xo \\ 0 & 1 & yo \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\text{sen}\theta & 0 \\ \text{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -xo \\ 0 & 1 & -yo \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} xi \\ yi \\ 1 \end{bmatrix}$$

Interpolación.

La interpolación es necesaria porque, en el caso general para un píxel, no se va a obtener de la imagen original un valor entero en la imagen destino. De hecho, a varios píxeles de la imagen original les corresponde el mismo en la imagen destino. Una mayor rapidez o exactitud dependerá del método de interpolación que se elija. Con carácter general, la transformación no se realiza de la imagen inicial a la final, sino al

revés; dado un píxel de la imagen final se localizan los píxeles de la inicial, que se utilizan para calcular el nuevo nivel de gris. La interpolación puede entenderse como una nueva digitalización de la imagen continua para obtener la discreta.

$$f(x, y) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} g(i, j) h(x - i, y - j)$$

Donde:

- $g(x,y)$ es la imagen continua.
- $f(x,y)$ es la imagen discreta.
- $h(x,y)$ representa la interpolación.

Si $h(x,y)$ fuera la delta de Dirac, la imagen inicial y la final serían la misma.

Vecino más próximo.

Es el método más sencillo. A partir de la transformación inversa se redondean los valores. Esto es lo mismo que ver cuál es el píxel más cercano. La función $h(x,y)$ sería:

$$h(x) = \begin{cases} 1 & \text{si } 0 < |x| < 0.5 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

Interpolación Bilineal.

Se denomina así porque se supone que la variación entre un píxel y el siguiente sigue un plano (figura 4.17):

$$f(x,y) = ax+by+cxy+d$$

Si una de las coordenadas x o y permanece constante, la ecuación anterior es la de una recta. Los coeficientes a, b, c, d deben fijarse de manera que los valores de la función coincidan con los cuatro píxeles vecinos. Las ecuaciones de las rectas son:

$$f(x,0) = f(0,0) + (f(1,0) - f(0,0)) x$$

$$f(x,1) = f(0,1) + (f(1,1) - f(0,1)) x$$

con estos dos puntos el valor respecto al otro eje será:

$$f(x,y) = f(x,0) + (f(x,1) - f(x,0)) y$$

y despejando obtenemos la ecuación final:

$$f(x,y) = (f(1,0) - f(0,0)) x + (f(0,1) - f(0,0)) y + (f(1,1) + f(0,0) - f(0,1) - f(1,0)) xy + f(0,0)$$

Los resultados son mejores que con el método anterior aunque, al usar una recta, se produce un desdibujado de los objetos. Si se usase la fórmula de la convolución la máscara vendría definida por:

$$h(x) = \begin{cases} 1 - |x| & \text{si } 0 < |x| < 1 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

Interpolación bicúbica.

Para evitar el inconveniente anterior se utiliza la interpolación cúbica, en la que se toman en consideración 16 puntos de la imagen. La convolución viene definida por:

$$h(x) = \begin{cases} 1 - 2|x|^2 + |x|^3 & \text{si } 0 < |x| < 1 \\ 4 - 8|x| + 5|x|^2 - |x|^3 & \text{si } 1 < |x| < 2 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

Caso general de magnificación.

Cuando se quiera magnificar una imagen $f(x,y)$ de dimensiones $N \times M$, para obtener otra $g(x,y)$ de dimensiones $pN \times pM$, se puede aplicar el siguiente algoritmo.

Primero se obtiene la imagen $g(x,y)$ utilizando la fórmula:

$$g(x, y) = \begin{cases} f(i, j) & \text{si } \frac{x}{p} = i \text{ y } \frac{y}{p} = j \\ 0 & \text{en cualquier otro caso} \end{cases}$$

similar a la del vecino más próximo; después se convoluciona $g(x,y)$ p -veces con la máscara:

$$\begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{bmatrix}$$

Los resultados pueden verse en la figura 4.18

4.5 Bibliografía.

- Brigham, E. O. *The Fast Fourier Transform*. Prentice Hall, 1974.
- Castleman K. R. *Digital Image Processing*. Prentice Hall 1996.
- Foley, J. D., Van Dam, A., Hughes, J.F. *Computer Graphics: Principles and Practice*. Addison-Wesley, 1990.
- Kraniuskas, P. *Transforms in signal and systems*. Addison-Wesley. 1992.
- Nussbaumer, H. J. *Fast Fourier Transform and Convolution Algorithms*. Springer Verlag, 1981.

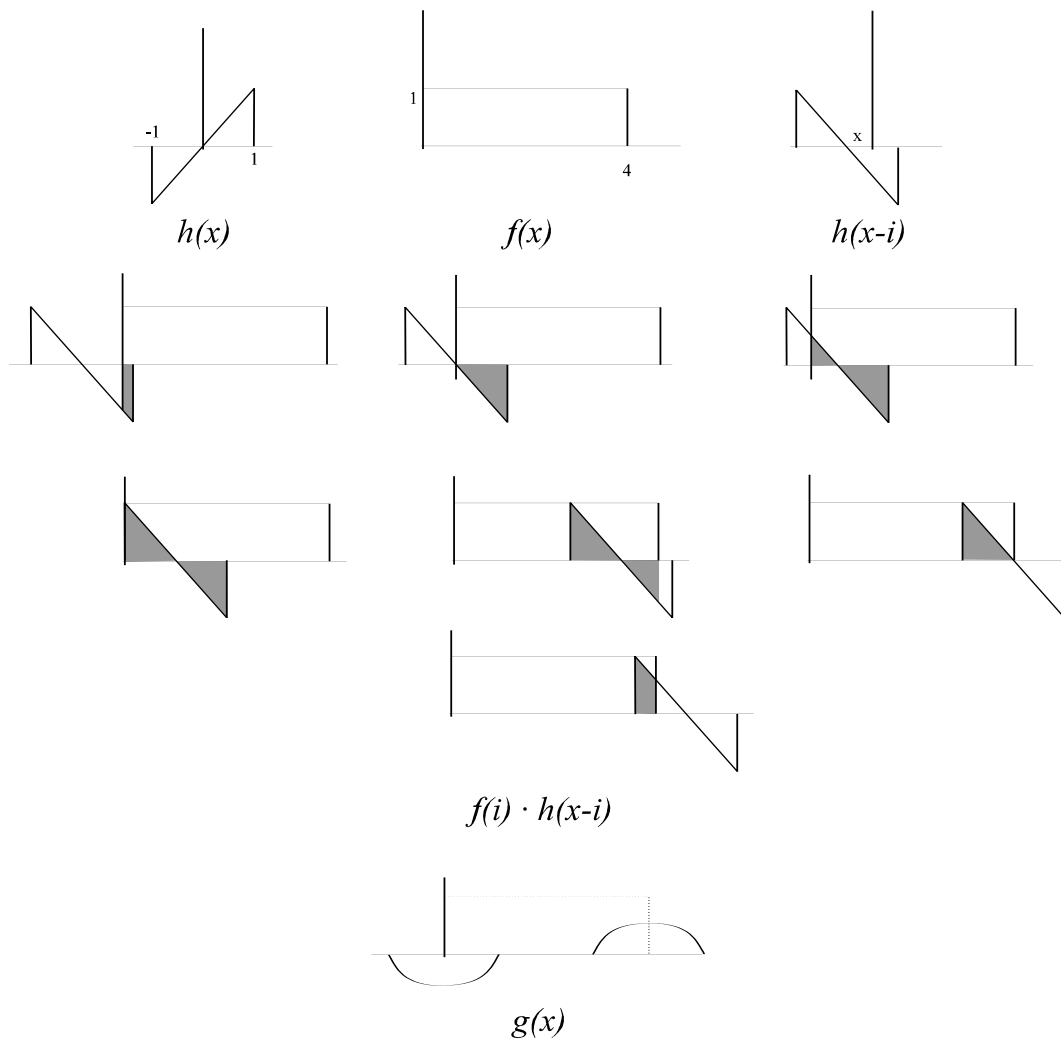


Figura 4.1. Convolución unidimensional.

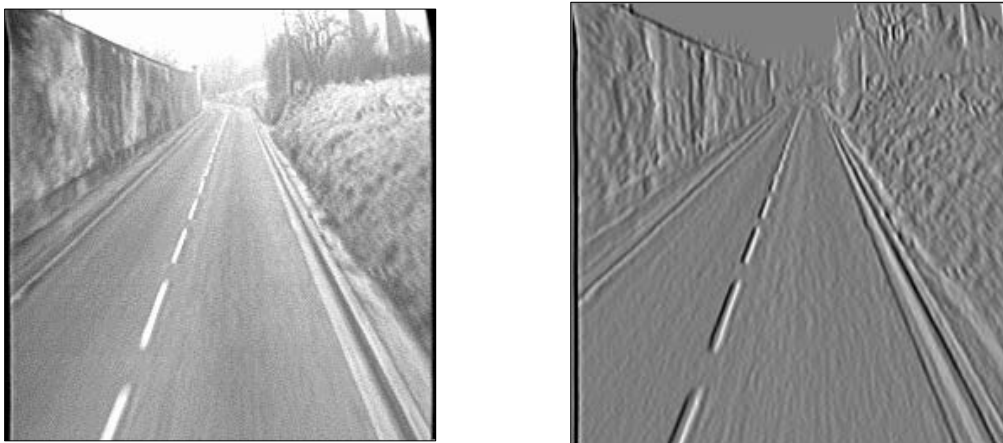


Figura 4.2. Convolución de una imagen digital.

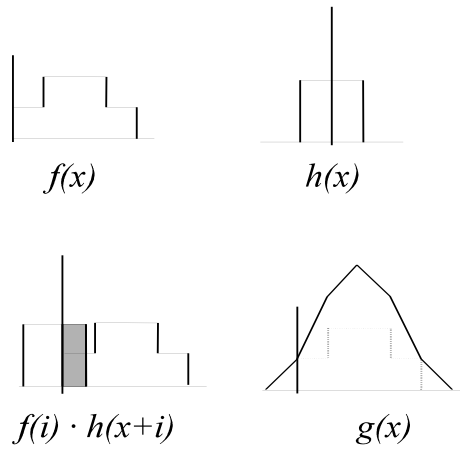


Figura 4.3. Correlación.

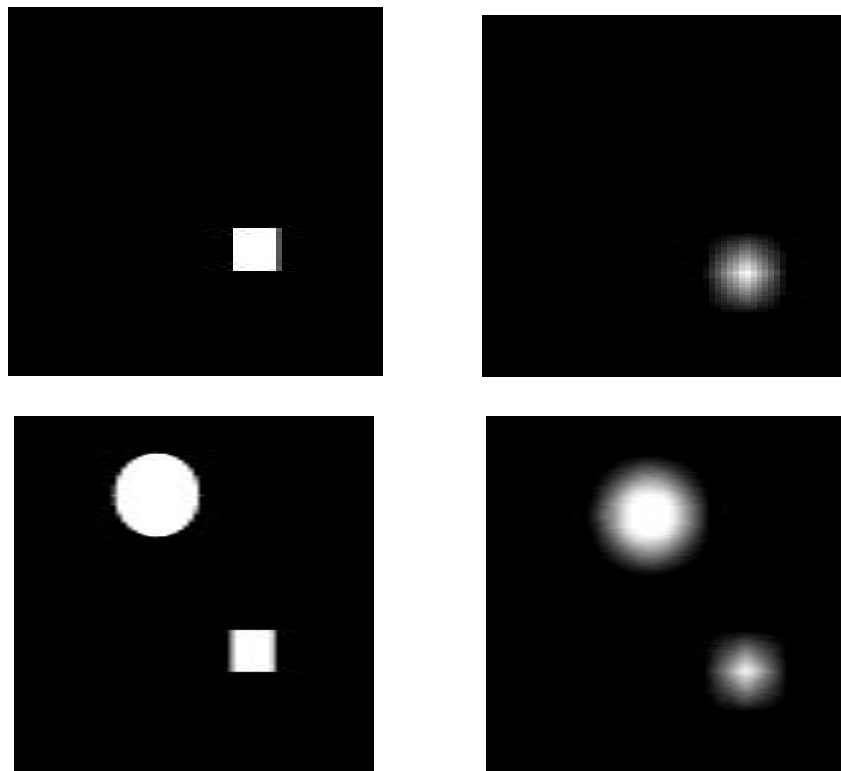


Figura 4.4 Detección de un cuadrado mediante correlación.

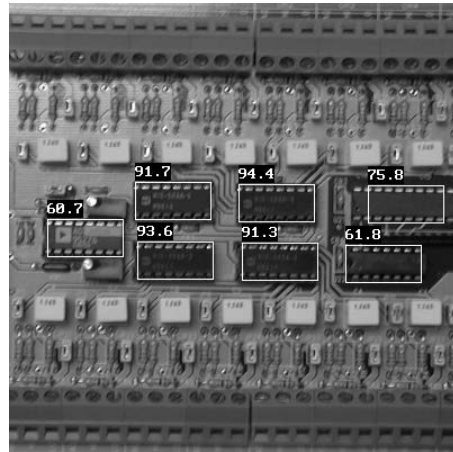
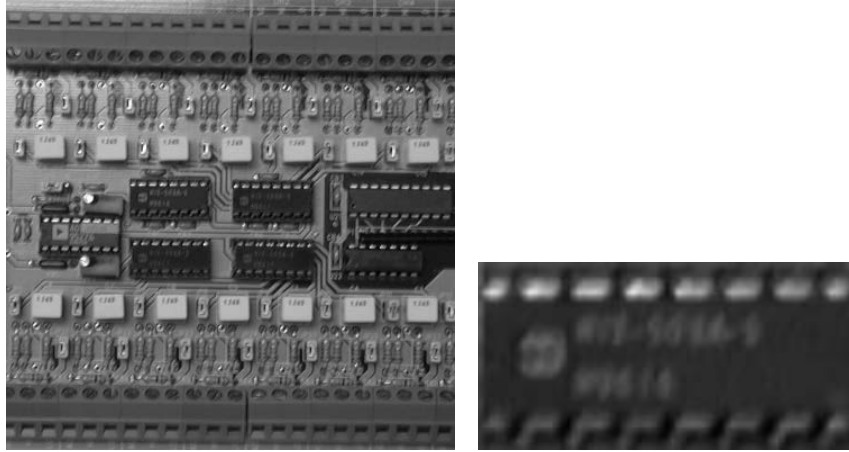


Figura 4.5. Correlación normalizada. (a) imagen (b) modelo ampliado (c) resultados de la correlación normalizada

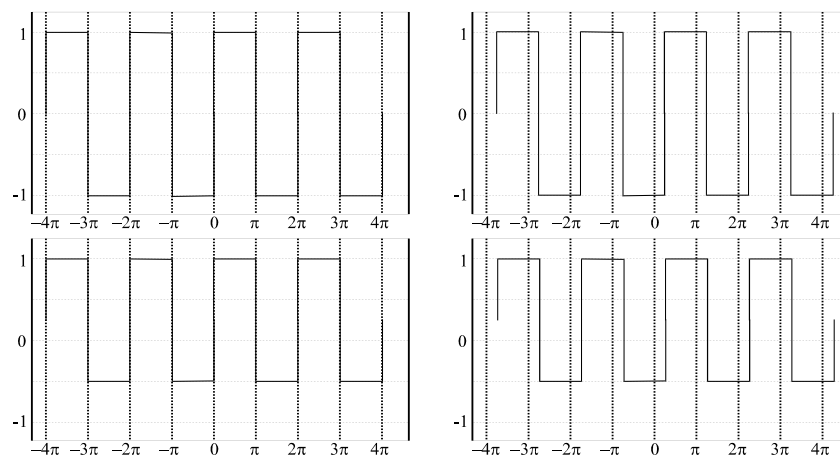


Figura 4.6: Diversas señales periódicas.

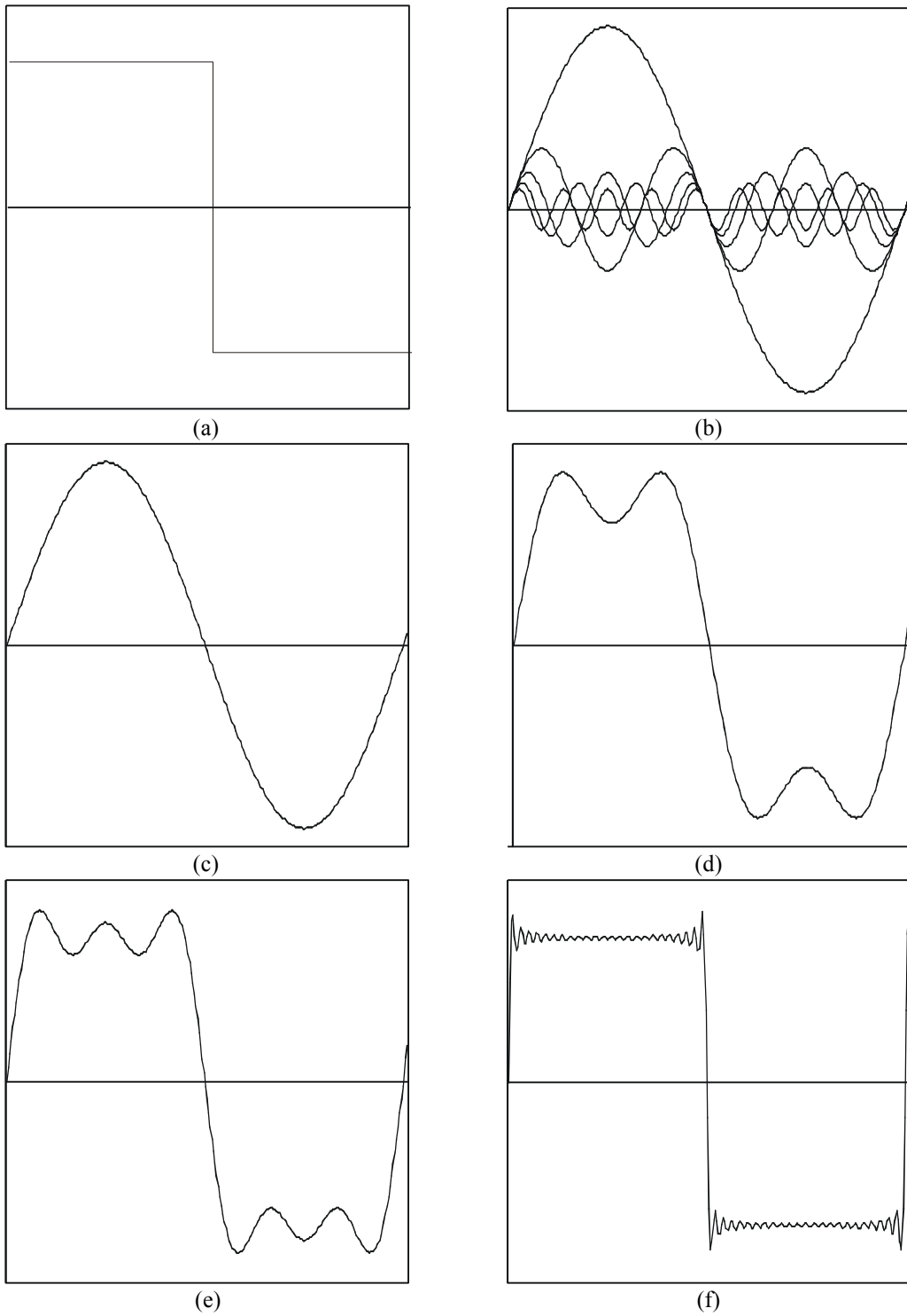


Figura 4.7 : Descomposición de una señal periódica en sus series de Fourier. (a) señal original (b) primeros cinco coeficientes (c) (d) (e) (f) suma de uno, dos, tres y veinticinco coeficientes

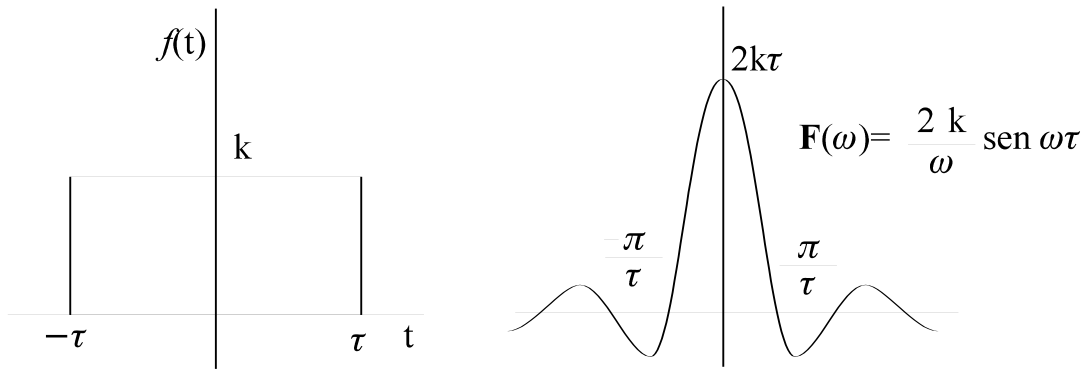


Figura 4.8. Transformada de Fourier.

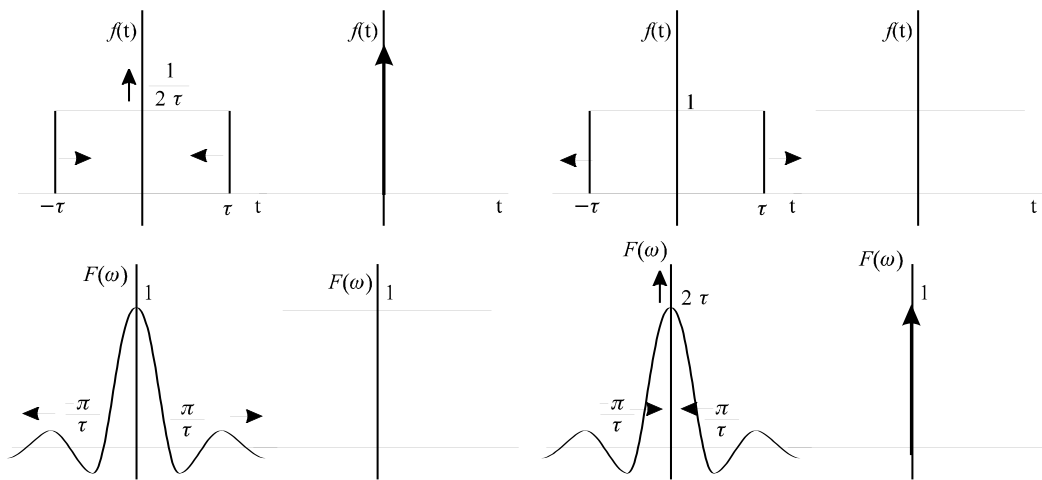


Figura 4.9 Transformada de Fourier de una señal continua y de la delta de Dirac.

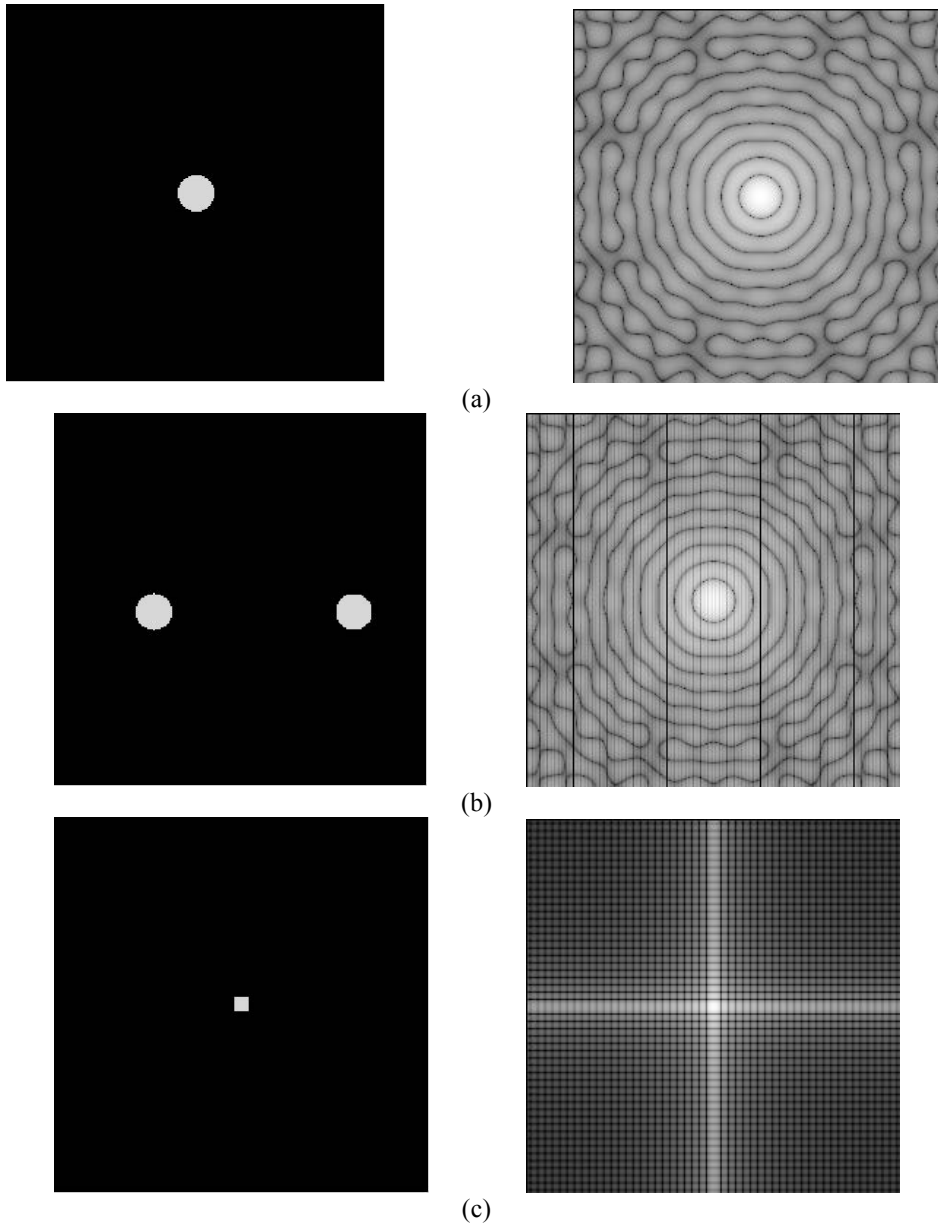


Figura 4.10 Módulo de la transformada de Fourier de (a) un círculo (b) dos círculos (c) un cuadrado.

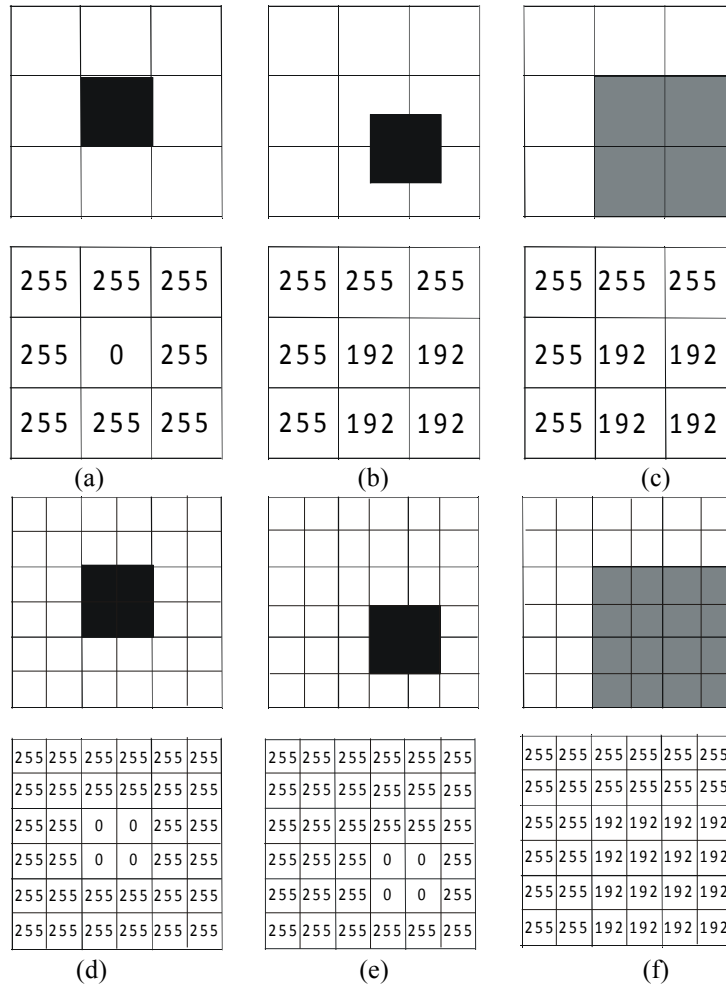


Figura 4.11. Digitalización de dos objetos variando su posición y el muestreo

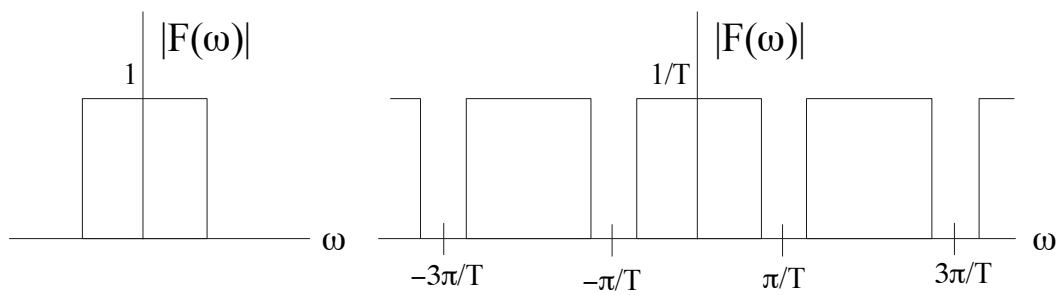


Figura 4.12. Transformada de Fourier de una señal continua y su equivalente para la señal muestreada

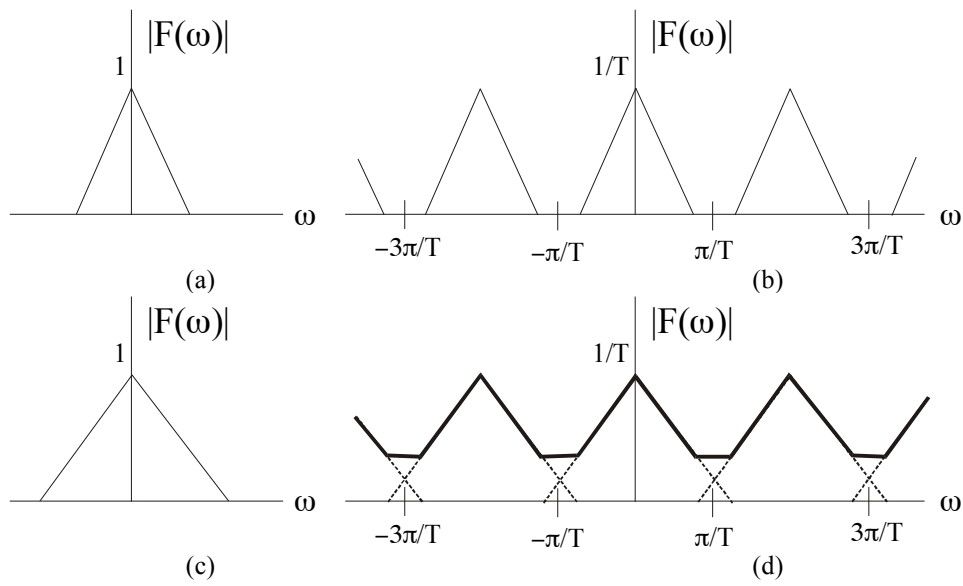


Figura 4.13. Influencia del tiempo de muestreo. (a) Transformada de Fourier de una señal continua (b su equivalente para la señal muestreada (c)(d) lo mismo pero para una señal que no cumple el teorema del muestreo

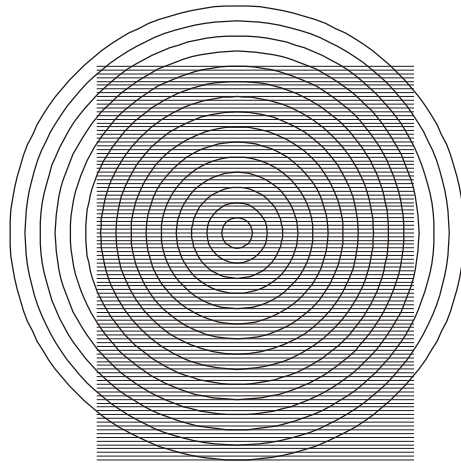


Figura 4.14. Patrones de Moiré

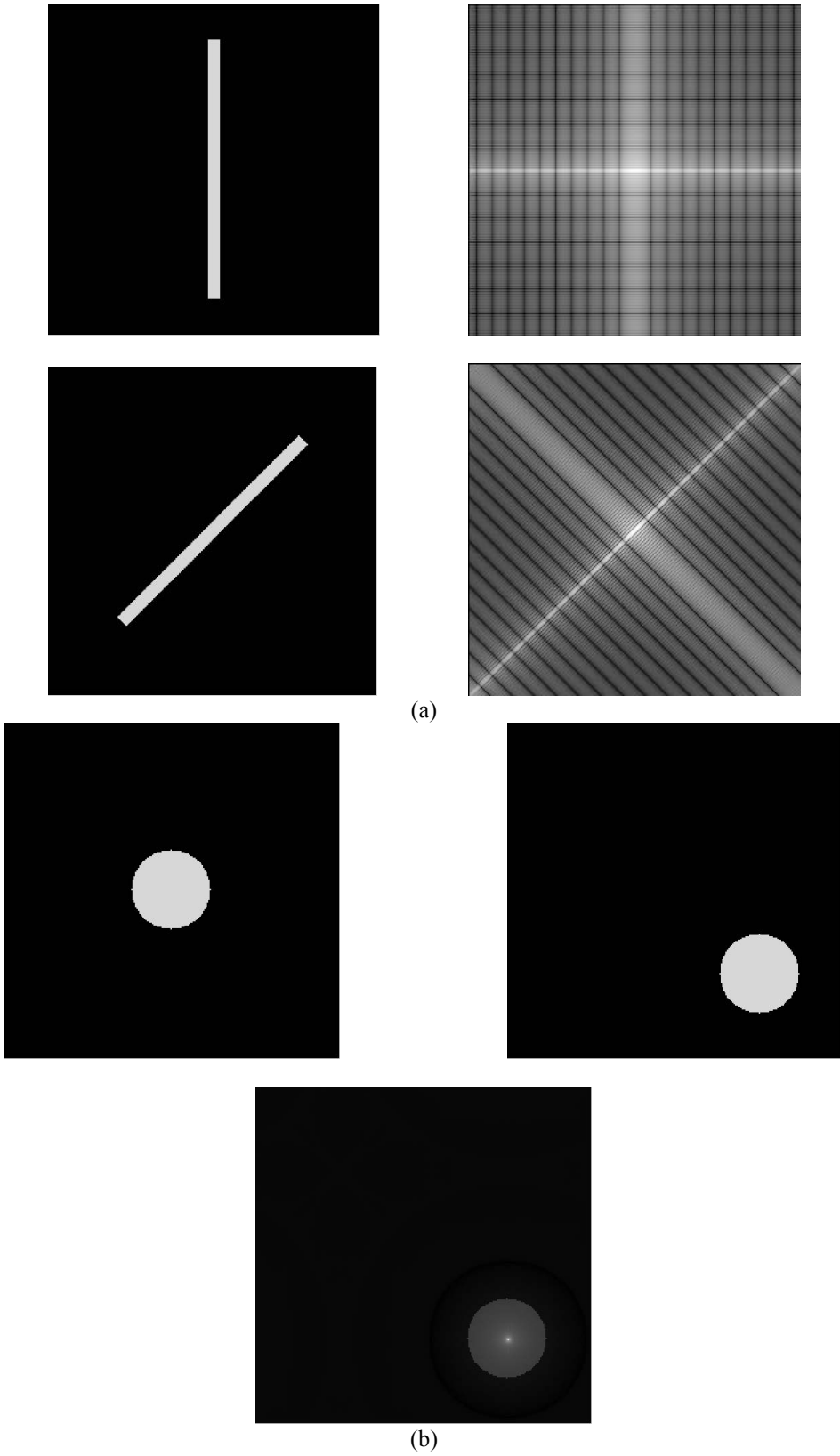


Figura 4.15. Algunas propiedades de la transformada de Fourier. (a) rotación (b) traslación

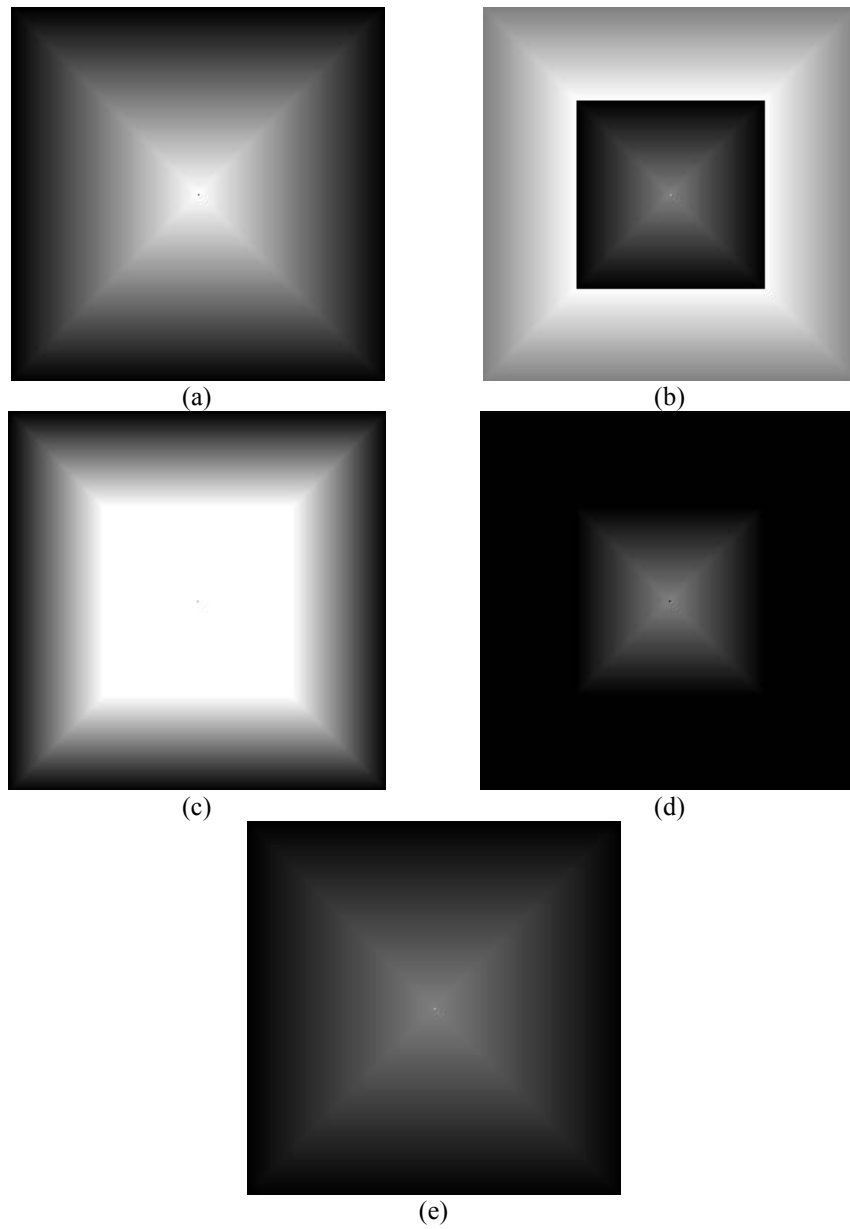


Figura 4.16. Desbordamiento del pixel. (a) imagen original (b) desbordamiento de la suma y la resta (c) suma con saturación (d) resta con saturación (e) normalización de resultados

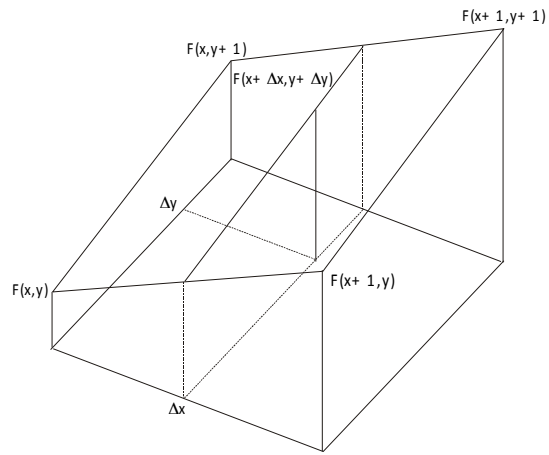
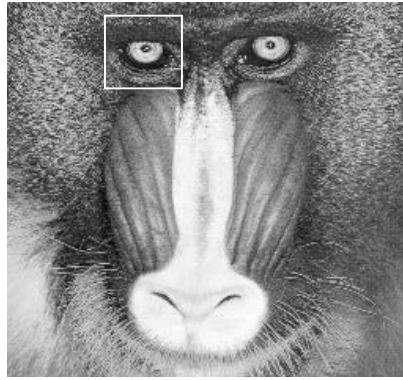
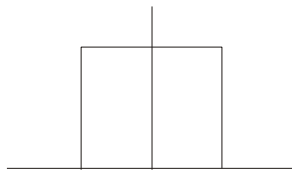
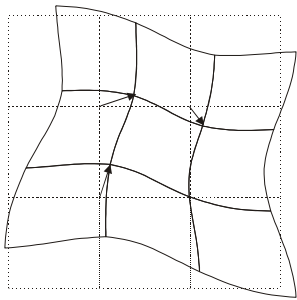


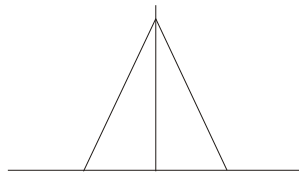
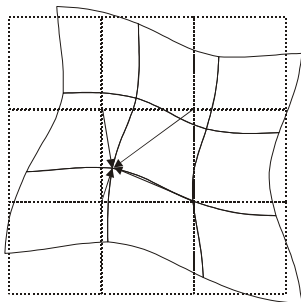
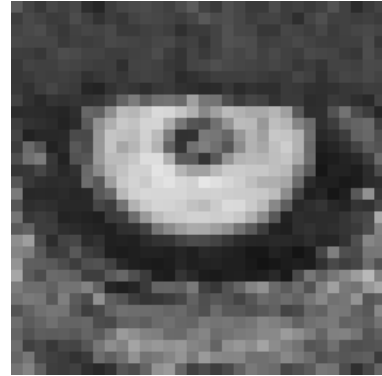
Figura 4.17. Interpolación bilineal.



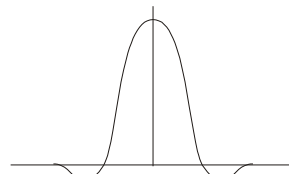
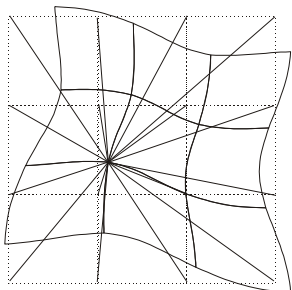
(a)



(b)



(c)



(d)



Figura 4.18. Interpolación mediante convoluciones. (a) imagen original (b) vecino más próximo (c) bilineal (d) bicúbica