
PROCESAMIENTO DE IMÁGENES

Dr. en C. Luis Fernando Gutiérrez Marfileño
Universidad Autónoma de Aguascalientes

Centro de Ciencias Básicas
Departamento de Ciencias de la Computación
Academia de Inteligencia Artificial y Fundamentos Computacionales

12 de Agosto de 2024

Índice general

1. INTRODUCCIÓN	6
1.1. Concepto de imagen	6
1.2. Procesamiento de imágenes	9
1.3. Representación digital	11
1.4. Indexación en arreglos y matrices	14
1.5. Lectura	15
1.6. Clases de imágenes	17
1.7. Escala de grises	21
1.8. Imágenes bi-nivel	23
1.9. Celdas de resolución	25
1.10. Conectividad y vecindad	30
1.11. Componentes conectados	32
2. FILTROS ESPACIALES	38
2.1. Funciones de intensidad	38
2.2. Histograma	40
2.3. Filtros espaciales	45
2.3.1. Ruido	48
3. DOMINIO DE FRECUENCIAS	53
3.1. Transformada de Fourier	53
3.2. Filtrado	59
3.3. Filtro Gaussiano	64
3.4. Frecuencias con filtros espaciales	67
3.5. Convolución	71

Índice de figuras

1.1. Imagen en cueva de Altamira (15,000 aC.)	6
1.2. Imagen producto del reflejo de la luz	7
1.3. Función de imagen	7
1.4. Imágenes por paleta de colores	11
1.5. Digitalización	12
1.6. Imagen digital binaria	16
1.7. Escala de grises	21
1.8. Imagen en modo de escala de grises	21
1.9. Imagen en modo de blanco y negro	22
1.10. Imagen bi-nivel	23
1.11. Imagen bi-nivel con objetos	24
1.12. Resolución de imagen	25
1.13. Comparativa de resolución	26
1.14. Resolución pp1 a ppcm	26
1.15. Tramado digital dpi	27
1.16. Tramado digital ppp	27
1.17. Efecto roseta	28
1.18. Forma del punto	29
1.19. Etiquetado	30
1.20. Elementos de una imagen	30
1.21. Vecindades	31
1.22. Tipos de conectividad	31
1.23. a) Original b) Contorno resaltado c) Aumento de brillo	33
1.24. a) Resta de fondo b) Detección de diferencias	33
1.25. a) Oscuro b) Original c) Brillante	33
1.26. a) Original 1 b) Original 2 c) Resultante	34

1.27. a) Original 1 b) Original 2 c) Resultante	34
1.28. a) Original 1 b) Original 2 c) Resultante	34
1.29. a) Original 1 b) Original 2 c) Resultante	35
1.30. a) Original b) Resultante	35
2.1. a) Imagen b) Histograma	40
2.2. a) Imagen b) Histograma	40
2.3. a) Imagen original b) Bajo contraste oscura c) Bajo contraste clara	41
2.4. a) Imagen Alto contraste	41
2.5. a) Trama de Imagen	41
2.6. Histograma correspondiente a la fig. 2.5	42
2.7. Expansión lineal	43
2.8. Ecualización	44
2.9. Calcular la salida (2, 4) de convolución	47
2.10. Calcular las salidas 2, 4 de la correlación	48
2.11. Función de distribución del ruido gaussiano	49
2.12. Ejemplo de ruido gaussiano	49
2.13. Función de distribución del ruido impulsivo	50
2.14. Ejemplo de ruido sal y pimienta	50
2.15. Ejemplo de ruido de Poisson	50
2.16. Ejemplo de ruido Speckle	51
3.1. Función a expandir	54
3.2. Elementos de la serie de expansión	55
3.3. Transformada de Fourier	55
3.4. Generación de la función senoidal	56
3.5. Trazado de figuras mediante senoidales	56
3.6. Transformadas de figuras simples	56
3.7. Transformada de imagen compleja	57
3.8. Escena agrícola y su Transformada	59
3.9. Escena dunar y su Transformada	60
3.10. Escena oceánica y su Transformada	60
3.11. Filtro aplicado a la Transformada	61
3.12. Transformadas semiperfiles y Resultados	61
3.13. Filtros pasa alto (a) y pasa-bandas (b) y (c)	62
3.14. a) Original b) espectro Fourier c) Filtro binario d) Resultado	62
3.15. Respuesta del filtro Gaussiano	64
3.16. Máscara Gaussiana de 3×3	64
3.17. Ejemplo ruido Gaussiano	65
3.18. Ejemplo filtro Gaussiano	65
3.19. Ejemplo suavizado Gaussiano	66
3.20. a) Original b) Desenfoque gaussiano	66

3.21. Desenfoque debido a varias σ	66
3.22. Etapas del filtrado	67
3.23. Filtro pasa bajos ideal y su ecuación	68
3.24. Ejemplo filtro pasa bajos	68
3.25. Filtro pasa altos ideal y su ecuación	69
3.26. Ejemplo filtro pasa altos	69
3.27. Filtro pasa banda ideal y su ecuación	69
3.28. Ejemplo filtro pasa banda	70
3.29. Convolución	71
3.30. Correlación cruzada (<i>Convolución</i>) de dos funciones, f y g	71
3.31. Proceso de Convolución	72
3.32. Kerneles varios	73
3.33. Elementos de convolución	75
3.34. Máscara	75
3.35. Inicio de una convolución	76
3.36. Kernel de Sobel	76

Índice de cuadros

1.1. Resolución/Lineatura	28
2.1. Tabla de intensidades	42

INTRODUCCIÓN

“Describir y manejar los conceptos fundamentales de imágenes y sus diferentes formatos, manipular archivos y matrices, que le permitan aplicar operaciones sobre imágenes”

Objetivo específico 1 y 2

1.1. Concepto de imagen

EN general una **imagen real** es la **representación visual** de uno o varios objetos en un entorno específico.

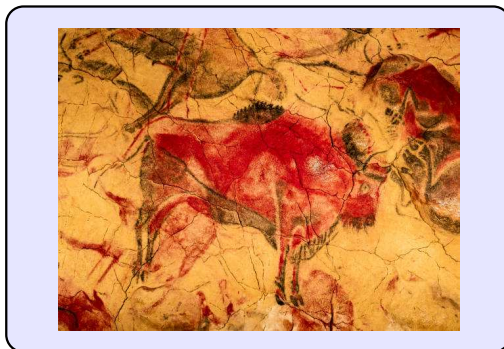


Figura 1.1: Imagen en cueva de Altamira (15,000 aC.)

Actualmente se piensa en una imagen en el sentido de una **representación plana del brillo**, es decir, la cantidad de luz reflejada o transmitida por una escena o un objeto.

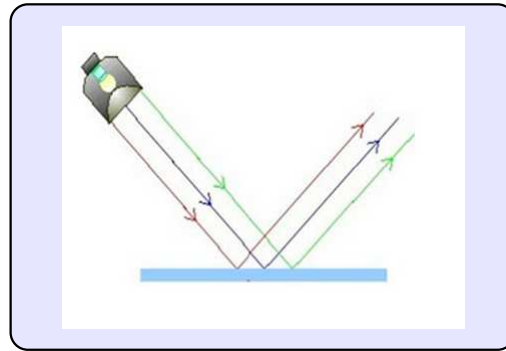


Figura 1.2: Imagen producto del reflejo de la luz

Matemáticamente, una forma de representar una imagen es por medio de una función de dos variables espaciales, por ejemplo, $f[x, y]$, que representa a la función **brillo** f , en la ubicación cartesiana $[x, y]$.

Obviamente, también puede ser graficado en tres dimensiones, con el brillo mostrado en el eje z .

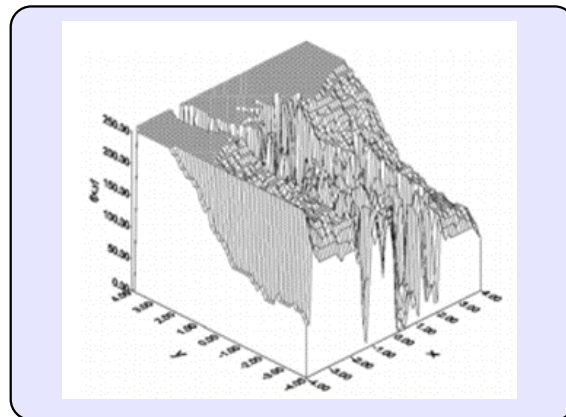


Figura 1.3: Función de imagen

Es cada vez más común tratar con imágenes que tienen más de dos dimensiones coordinadas, por ejemplo:

- $f[x, y, t_d]$: película monocromática, conjunto discreto de imágenes a lo largo del tiempo
- $f[x, y, \lambda]$ imagen espectral con dominio continuo de longitudes de onda
- $f[x, y, \lambda_d]$ imagen multispectral, conjunto discreto de longitudes de onda
- $f[x, y, t]$ imagen monocromática variable en el tiempo en un dominio de tiempo continuo



- $f[x, y, t_m]$ imagen monocromática variable en el tiempo con muestras de tiempo discretas (cine)
- $f[x, y, z]$ imagen monocromática 3D (p. ej., holograma óptico)
- $f[x, y, t_d, \lambda_m]$ muestras discretas en tiempo y longitud de onda, p. ej., película en color
- $f[x, y, z, t, \lambda]$ realidad

Por lo general, es posible *cortar* porciones bidimensionales de estas funciones multidimensionales para crear imágenes, pero las imágenes no tienen por qué ser **pictóricas** (que busquen representar la realidad).

Por ejemplo, considere el 2-D rebanadas *cortadas* de la función 3-D función espacio-temporal $f[x, y, t]$; la rebanada 2-D $f[x, y; t = t_0]$ es pictórica, pero $f[x; y = y_0, t]$ no lo es.

Dicho esto, las unidades de los ejes no tienen ningún efecto sobre los cálculos; es perfectamente factible para una computadora procesar y mostrar $f[x; y = y_0, t]$ para hacer lo mismo con $f[x, y; t = t_0]$.

Una vez convertida la información de la imagen en una matriz de números enteros, la imagen puede ser manipulada, procesada y mostrada por computadora.

El procesamiento informático se utiliza para la mejora de imágenes, restauración, segmentación, descripción, reconocimiento, codificación, reconstrucción o transformación.

En esta función la amplitud de f en cualquier par de coordenadas (x, y) se llama **intensidad** o **nivel de gris** de la imagen en ese punto.

Cuándo x, y , y los valores de intensidad de f son todos cantidades finitas y discretas, se le llama **imagen digital**.

Definición (> Imagen digital)

Es la representación en dos dimensiones de una imagen basada en una matriz numérica binaria.

[Cambridge Dictionary, 2022]

El campo del **procesamiento de imágenes digitales** se refiere a la transformación de imágenes por medio de una computadora digital.

Se debe tener en cuenta que una imagen digital se compone de un número finito de elementos, cada uno de los cuales tiene una ubicación y un valor particulares.

Estos elementos se denominan elementos de imagen, pels o **píxeles**, éste último es el término que se usa más ampliamente para denotar a los elementos de una imagen digital.



1.2. Importancia del procesamiento de imágenes

EL **procesamiento de imágenes** tiene como objetivo mejorar el aspecto de las imágenes y hacer más evidentes en ellas ciertos detalles que se desean hacer notar.

La imagen puede haber sido generada de muchas maneras, por ejemplo, fotográficamente, o electrónicamente, por medio de monitores de televisión.

El procesamiento de las imágenes se puede hacer, en general, por medio de métodos ópticos, o bien por medio de métodos digitales en una computadora.

Procesamiento óptico de imágenes

Los principios del procesamiento óptico de imágenes están bien establecidos desde el siglo pasado, cuando se desarrolló la teoría de la difracción de la luz.

Sin embargo, su aplicación práctica data apenas del principio de la década de los sesenta, cuando se comenzó a disponer del rayo láser.

El procesamiento óptico se basa en el hecho de que la imagen de difracción de Fraunhofer de una transparencia colocada en el plano focal frontal de una lente es una distribución luminosa que representa la distribución de las frecuencias de Fourier que componen la imagen, a la que se le llama técnicamente **transformada de Fourier**.

Procesamiento digital de imágenes

Al igual que en el caso del procesamiento óptico, los principios fundamentales del procesamiento digital de imágenes están establecidos hace muchos años, pero no se llevaban a cabo debido a la falta de computadoras.

Con la aparición de las computadoras de alta capacidad y memoria, era natural que se comenzara a desarrollar este campo.

Uno de los primeros lugares donde se empezó a realizar el procesamiento digital fue en el Jet Propulsion Laboratory, en 1959, con el propósito de mejorar las imágenes enviadas por los cohetes.

Los resultados obtenidos en un tiempo relativamente corto fueron tan impresionantes que muy pronto se extendieron las aplicaciones del método a otros campos.

El procesamiento digital de imágenes se efectúa dividiendo la imagen en un arreglo rectangular de elementos.

Cada elemento de la imagen así dividida se conoce con el nombre de pixel.

El siguiente paso es asignar un valor numérico a la luminosidad promedio de cada pixel.

Así, los valores de la luminosidad de cada pixel, con sus coordenadas que indican su posición, definen completamente la imagen.

Todos estos números se almacenan en la memoria de una computadora.

El tercer paso es alterar los valores de la luminosidad de los pixeles mediante las operaciones o transformaciones matemáticas necesarias, a fin de hacer que resalten los detalles de la imagen que sean convenientes.

El paso final es pasar la representación de estos pixeles a un monitor de alta definición.

La utilidad del procesamiento de imágenes es muy amplia y abarca muchos campos.



Un ejemplo son las imágenes obtenidas con fines de **diagnóstico médico**; este método permite a los profesionales de la salud observar el interior del cuerpo para buscar indicios de una afección médica.

Otro ejemplo son las imágenes aéreas obtenidas para realizar **exámenes del terreno**; mediante este método se pueden analizar los recursos naturales, las fallas geológicas del terreno, etcétera.

Algunas de las transformaciones que se pueden realizar sobre las imágenes digitales son:

Realce

Mejora y correcciones (degradaciones: bajo contraste, ruido, desenfoque (*blur*), orientación, ciertas distorsiones, etc.), especialmente para llevar a intervalos característicos de visión humana (*visualizar*).

En principio no implica modelos del original o de la degradación (mejora empírica, sin que necesariamente coincida con imagen original o condiciones originales, que pueden no ser adecuadas: baja iluminación, o demasiada (por ejemplo una imagen del Sol, datos de ultrasonido, o en rayos infrarrojos, invisibles al ojo humano).

Un posible objetivo final, es el almacenamiento, codificación y transmisión, sin análisis (no necesariamente).

Criterios de calidad de imagen en función del grado de mejora o reducción de la degradación.

Carácter más cualitativo que cuantitativo.

Criterios de apreciación de rasgos e inclusive de estética.

Gran dependencia de la imagen original, poca de la adquirida.

Restauración

Puede implicar mejora y correcciones, pero además, en base a modelos (del objeto o escena original, o de la degradación sufrida), implica recuperar (restaurar) la imagen original (lo mejor posible), o alguna de sus características?

Corrección de distorsiones (geométricas Y/O del atributo): registro (alineación correcta con transformaciones, inclusive no-lineales), normalización, etc..

El objetivo usualmente es menos de visualización y más de recuperación y/o análisis cuantitativo y automático (por eso a veces se le denomina también *pre – procesamiento*, en el sentido de preparar la información, uniformizarla, etc., para su estudio.

Pero se puede desear solo una imagen puramente restaurada (como en las obras de arte).

Utiliza criterios de fidelidad de imagen en función del grado de similitud con el original o alguna referencia, o del modelo de degradación o modelo de cómo debe ser el original.

Es de carácter más cuantitativo que cualitativo.

Tiene mayor dependencia de proceso de adquisición y degradación, menor dependencia de la imagen particular.



1.3. Representación digital

Una **imagen natural** capturada con una cámara, un telescopio, un microscopio o cualquier otro tipo de instrumento óptico presenta una variación de sombras y tonos continua, imágenes de este tipo se conocen como **imágenes analógicas**.

Para que una imagen analógica, pueda ser *manipulada* usando una computadora, primero debe convertirse a un formato adecuado.

Este formato da como resultado a la **imagen digital** correspondiente.

La transformación de una imagen analógica a otra discreta se llama **digitalización** y es el primer paso en cualquier aplicación de procesamiento de imágenes digitales.

Clasificación de imágenes digitales

- Por número de dimensiones:
 - Imágenes 2D
 - Imágenes 3D
- Por paleta de colores:
 - Imágenes binarias
 - Imágenes en escala de grises

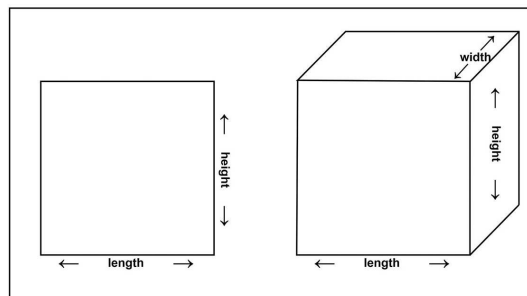


Figura 1.4: Imágenes por paleta de colores



El proceso de digitalización

Digitalizar una imagen significa convertirla en un archivo que puede ser manipulado por la computadora, es decir en un conjunto de bits.

Para digitalizar una imagen es necesario dividirla en unidades discretas, cada una de las cuales se llama píxel, una vez dividida la imagen, se le asigna un valor a cada uno de los píxeles.

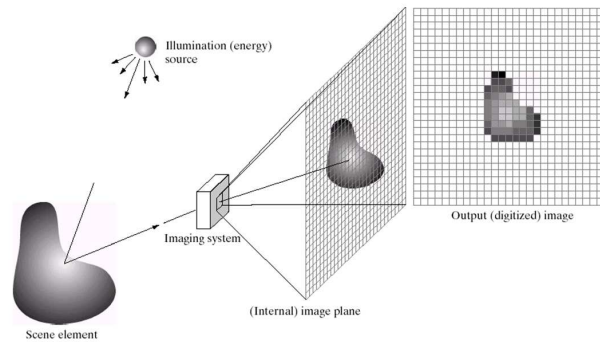
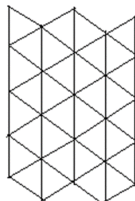


Figura 1.5: Digitalización

El proceso de digitalización consta de dos fases, muestreo y cuantificación.

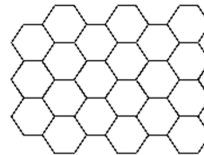
- Un **muestreo** consiste en una subdivisión de la imagen analógica en porciones. Nos centraremos en imágenes 2D. Sólo se estudiarán particiones que envuelven polígonos regulares: triángulos, cuadrados y hexágonos.



Mallado triangular



Mallado cuadrangular



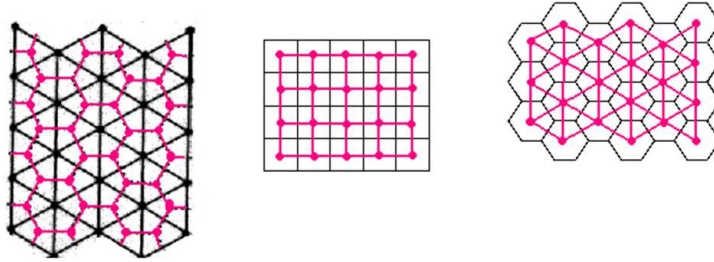
Mallado hexagonal

Éstos polígonos representan sensores sensibles a la intensidad de luz.

En el modelo matemático de una imagen, un píxel se identifica con su centro, pudiendo representar los píxeles como puntos (x,y) del plano, donde (x,y) son las típicas coordenadas cartesianas.

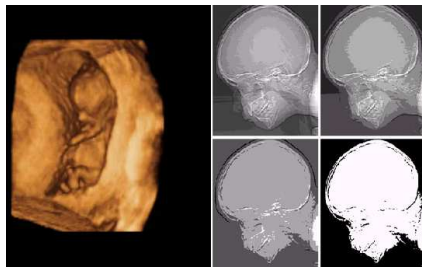
Dependiendo de los distintos tipos de mallado, la distribución de los píxeles es distinta:





Los bordes de las regiones están pintados en negro. Los píxeles están representados por puntos en color rosa. Hemos conectado dos píxeles si las regiones correspondientes comparten un lado común.

- **Cuantificación:** La salida de estos sensores es un valor (amplitud) dentro de una escala (color). La salida pueden ser, o bien un único valor (escala de grises) o bien un vector con tres valores por polígono (RGB) que se corresponden con la intensidad de color rojo (R), verde (G) y azul (B). La escala de colores también tiene un rango discreto (por ejemplo, de 8-bits = 256 valores).



A este proceso de discretización del color se le llama cuantificación.

Un polígono de color constante se le llamará píxel.

La resolución (el grado de detalle discernible) de una imagen depende estrechamente de estos dos parámetros. Cuanto más se incrementan, más se aproxima la imagen digitalizada a la original.

La cantidad de niveles de gris (resolución de intensidad) y la finura del mallado (resolución espacial) escogida, deben producir una imagen digital *aceptable*, en el sentido de que no sea perceptible al ojo humano el paso de un color a otro, entre dos píxeles consecutivos.



1.4. Indexación en arreglos y matrices

Una **matriz** es un arreglo bidimensional de números mediante los cuales es posible representar una imagen; en la computadora una matriz es una zona de almacenamiento continuo que contiene una serie de elementos del mismo tipo (elementos de matriz).

Por lo tanto puede decirse que una matriz es la representación numérica de una imagen.

Por lo tanto el acceso a cada elemento de la matriz imagen es importante al realizar cualquier procesamiento sobre ésta.

Al proceso de acceder a cada uno de éstos elementos se conoce como **indexación**, existen diferentes formas de acceder a un elemento específico de una matriz:

- Por posición
- En forma lineal
- En forma lógica

Indexación por posiciones de elementos

El método más frecuente es especificar de manera explícita los índices de los elementos.

Por ejemplo, para acceder a un único elemento de una matriz, especifique el número de fila seguido del número de columna del elemento.

Indexación en forma lineal

Otro método para acceder a los elementos de un arreglo es utilizar un único índice, independientemente del tamaño o de las dimensiones del arreglo.

Este método se conoce como indexado lineal, aunque algunos softwares muestran arreglos según las formas y los tamaños definidos, en realidad se guardan en la memoria como una única columna de elementos.

Una buena forma de visualizar este concepto es con una matriz.

Indexación en forma lógica

Utilizar indicadores lógicos verdaderos y falsos es otra manera útil de indexar en arreglos, especialmente al trabajar con instrucciones condicionales.

Por ejemplo, suponer que se desea saber si los elementos de una matriz A son inferiores a los elementos correspondientes de otra matriz B.

El operador *menor que* devuelve un arreglo lógico cuyos elementos son 1 cuando un elemento de A es más pequeño que el elemento correspondiente de B.



1.5. Lectura, desplegado y operaciones sobre imágenes

Como ya se señaló, una imagen digital está compuesta por un código numérico que luego la computadora representa en la pantalla mediante píxeles.

Las imágenes digitales se pueden obtener por muchos medios (escáner o cámara digital por ejemplo) y son almacenadas en bits.

El bit

Un bit es la unidad de información que equivale a la elección de 0 ó 1, no pasa corriente o pasa corriente. En la combinación de ellos se basa la informática. ?

Como la información que podemos acumular en 1 bit es muy pequeña, se aunan en Bytes que es un conjunto de 8 bits. Así se obtienen:

- El Kilobyte (k): 1024 bytes
- El Megabyte (MB): 1024 Kb
- El Gigabyte (GB): 1024 MB

La imagen digital es una traducción, la representación bidimensional de una imagen empleando bits.

De acuerdo en cómo son obtenidas hay dos tipos de imágenes digitales:

■ Imágenes Digitales Procesadas (IDP)

Las IDP resultan de lo que se conoce como «procesamiento de imágenes», que consiste en almacenar digitalmente materiales visuales previos - por ejemplo, fotografías, reproducciones fotográficas, películas o vídeos - para trabajar posteriormente sobre ellos y alterarlos/manipularlos de múltiples maneras.

Es decir, son imágenes obtenidas a partir de dispositivos de conversión analógica-digital, como los escáneres o las cámaras digitales y pueden ser manipuladas digitalmente a través de software, pero éste no interviene directamente en su proceso de creación.

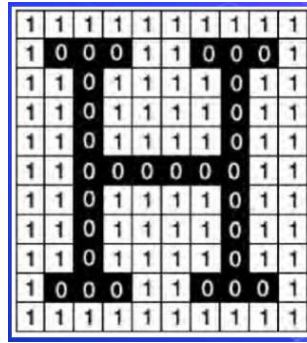
■ Imágenes Digitales de Síntesis (IDS)

Las IDS - también llamadas infografías- se sitúan en el entorno de la *Computación Gráfica*.

Se obtienen a través de software, produciéndolas enteramente a través de algoritmos y sin necesidad de referente real, por ejemplo realizando dibujos con el ratón o por medio de programas de renderización 2D y 3D.

Las IDS, se basan en la introducción de datos matemáticos en la memoria de la computadora, la cual describe o modela, y después almacena las imágenes generadas?





1	1	1	1	1	1	1	1	1	1
1	0	0	0	1	1	0	0	0	1
1	1	0	1	1	1	1	0	1	1
1	1	0	1	1	1	1	0	1	1
1	1	0	1	1	1	1	0	1	1
1	1	0	0	0	0	0	0	1	1
1	1	0	1	1	1	1	0	1	1
1	1	0	1	1	1	1	0	1	1
1	1	0	1	1	1	1	0	1	1
1	0	0	0	1	1	0	0	0	1
1	1	1	1	1	1	1	1	1	1

Figura 1.6: Imagen digital binaria

Estas imágenes pueden manipularse también de diversas maneras.

Las imágenes digitales de síntesis entran en el terreno de la *simulación digital*.

Los bits de cada píxel son almacenados por la computadora en una secuencia, y con frecuencia se los reduce a una representación matemática (comprimida).

Luego la computadora interpreta y lee los bits para producir una versión analógica para su visualización o impresión.



1.6. Clases de imágenes

Existen dos clases de imágenes digitales y cada una de ellas posee características bien diferenciadas.

- **Imágenes entramadas** (*raster*): también llamadas **mapas de bits**, son el tipo de imágenes digitales mas comunes, casi todas las imágenes que se pueden descargar de internet son de este tipo, están formadas por pixeles donde a cada pixel se le asigna un color por ello a mayor resolución mayor calidad de la imagen. Un problema importante en estas imágenes es el llamado **pixelado** (efecto causado por visualizar una imagen o una sección de una imagen a un tamaño en el que los píxeles individuales son visibles al ojo). Este es el tipo de imagen que se utiliza en fotografías, ilustraciones digitales, o gráficos para web. Un editor muy común para este tipo de imágenes es Photoshop.
- **Imágenes vectoriales**: no están formadas por pixeles por lo que son independientes de la resolución lo que significa que al redimensionar la imagen no se produce el pixelado. Los gráficos vectoriales están formados por **puntos de control** que poseen curvas entre ellos y que son definidas gracias a formulas matemáticas que el editor de imágenes se encarga de calcular. Los vectores son muy utilizados en gráficos de baja complejidad como logotipos, iconos composiciones tipográficas. Los editores mas usados para crear y modificar gráficos vectoriales son Adobe Ilustrator y Corel Draw. Una imagen vectorial puede ser convertida a mapa de bits en forma muy fácil, pero es muy complicado obtener la calidad de una imagen rasterizada en vectores ya que siempre se pierden detalles.

A continuación distintos formatos de imágenes raster:

Formato JPEG: Es uno de los formatos más populares para guardar imágenes digitales, este formato fue creado por *Joint Photographic Experts Group*, las siglas de este grupo son las que dan el nombre del formato JPEG.

Se utiliza para trabajos de fotografía, debido a que es el que más colores puede captar.

Una de sus principales ventajas es que se puede reducir el tamaño de la imagen sin que eso influya en su calidad y no admite transparencias.

Utiliza un algoritmo de compresión muy complejo que permite reducir notablemente el tamaño de los archivos, por lo cual es generalmente óptimo para su uso en sitios web, y se utiliza cuando es más importante que el archivo sea pequeño, que el hecho que posee una gran calidad de imagen.

Asimismo, puede mostrar millones de colores, ya que posee un formato de datos de 16 bits.

En general, se recomienda utilizar este tipo de formato en los siguientes escenarios:

- En imágenes fijas
- En fotografías



- En imágenes con muchos colores
- En imágenes con sombreados, con mucha luminosidad y contraste

Formato GIF: Significa *Graphic Interchange Format* (Formato de Intercambio Gráfico) se utilizan para representar videos y animaciones.

El formato fue presentado en 1987, pero ganó gran popularidad en los últimos años debido a las redes sociales.

Sigue la técnica de compresión sin pérdida LZW. Significa que se mantiene la calidad original de los datos.

Sin embargo, los GIFs sólo permiten pixeles de 8 bits, es decir, sólo hay 256 posibles combinaciones de colores en el formato.

Puesto que los GIFs almacenan *fotos en movimiento*, el tamaño del archivo es generalmente más grande que los JPEGs.

Los GIFs se utilizan para expresar emociones y para entretenimiento, incluso para motivos educativos debido a su naturaleza interactiva.

Sin embargo, la práctica común de usar GIFs es en plataformas sociales, como WhatsApp, Messenger, Tumblr, Twitter, etc.

Tiene un formato ligero, lo que permite su reproducción en cualquier dispositivo.

Cuentan con una capacidad alta de llamar la atención y permite la comprensión rápida y sin esfuerzo de lo que intentan transmitir.

Generan notoriedad y engagement, es una forma de relacionarse rápidamente a un concepto y llegar a un público objetivo.

Son un formato que cuenta con una gran capacidad de viralización.

Formato PNG: El formato *Portable Network Graphics* es un formato de archivos de gráficos de mapa de bits (una trama).

Fue desarrollado en 1995 como una alternativa gratuita al formato GIF, cuyos derechos pertenecen a Unisys (propietario del algoritmo de compresión LZW).

Permite almacenar imágenes en blanco y negro (una profundidad de color de 16 bits por píxel) y en color real (una profundidad de color de 48 bits por píxel).

Así como también imágenes indexadas, utilizando una paleta de 256 colores.

Además, soporta la transparencia de canal alfa, es decir, la posibilidad de definir 256 niveles de transparencia, mientras que el formato GIF permite que se defina como transparente solo un color de la paleta.

También posee una función de entrelazado que permite mostrar la imagen de forma gradual.

Un archivo PNG comprende una firma, que permite indicar que se trata de un archivo PNG, seguida de una serie de elementos denominados fragmentos.

Cada fragmento comprende 4 partes: el tamaño, un entero de 4 bytes no firmado, que describe el tamaño del fragmento; el tipo de fragmento, un código de 4 caracteres (4 bytes) comprendido por caracteres alfanuméricos ASCII (A-Z, a-z, 65 a 90 y 97 a 122) que permite establecer la naturaleza del fragmento.



Los datos del fragmento; la CRC (comprobación de redundancia cíclica), un código de corrección de 4 bytes que permite comprobar la integridad.

Los fragmentos pueden aparecer en cualquier orden excepto por el hecho de que el fragmento de encabezado debe ir adelante (fragmento IHDR) y el de cierre (fragmento IEND) detrás.

Los principales fragmentos (denominados fragmentos críticos) son IHDR Encabezado de información de mapa de bits, PLTE la paleta, IDAT los datos de la imagen, IEND el cierre de la imagen.

Los otros fragmentos (denominados fragmentos secundarios) son los siguientes: bKGD el color de fondo, cHRM las cromaticidades primarias y el punto blanco, gAMA la gama de la imagen, hIST el histograma de la imagen, pHYS.

Las dimensiones del píxel físico, sBIT los bits importantes, tEXt los datos de texto, tIME la hora de la última modificación, tRNS la transparencia y zTXt los datos de texto comprimidos.

Formato TIFF: El formato *Tagged Image File Format* es un formato de archivos de gráficos de mapa de bits (una trama).

TIFF es un formato flexible y extensible que es compatible con una amplia variedad de plataformas y aplicaciones de procesamiento de imágenes.

Los archivos TIFF pueden almacenar imágenes con un número arbitrario de bits por píxel y pueden emplear diversos algoritmos de compresión.

Se pueden almacenar varias imágenes en un único archivo TIFF de varias páginas.

La información relacionada con la imagen (marca del analizador, equipo host, tipo de compresión, orientación, muestras por píxel, entre otros) se puede almacenar en el archivo y organizarse mediante el uso de etiquetas.

El formato TIFF se puede extender según sea necesario mediante la aprobación y adición de nuevas etiquetas.

Formato BMP: El formato *Bit Map* propio del sistema operativo Microsoft Windows.

Puede guardar imágenes de 24 bits (16,7 millones de colores), 8 bits (256 tonos de gris) y menos.

Dependiendo de la profundidad de color que tenga la imagen cada píxel puede ocupar 1 o varios bytes.

En esta clase de archivos puede seleccionarse una compresión RLE (*Run Length Encoding*) sin pérdida de calidad.

El uso más común de este formato consiste en generar imágenes de poco peso y no se aconseja utilizarlo en imágenes recién captadas, sino en imágenes una vez reducidas a los 24 bits.

Se utiliza mucho para crear fondos para el escritorio de Windows, aunque, con el auge de Internet, ha ido perdiendo importancia en favor del formato JPEG, que produce archivos más pequeños y más rápidos viajando por la red.



Formato AI: *Adobe Illustrator* es uno de los formatos vectoriales más usados y la opción por defecto del programa Adobe Illustrator.

Entre las ventajas que ofrece se encuentra la de que, además de que las imágenes se pueden escalar sin pérdida de calidad de imagen como ocurre con todos los vectores, el formato guarda la transparencia del diseño original.

Suele utilizarse para crear logotipos, infografías, gráficos digitales y diseños para impresión.

Formato EPS: el formato *Encapsulated PostScript* y es un formato antiguo de archivos vectoriales.

Uno de los grandes beneficios de usar este formato vectorial es que los archivos .EPS son compatibles con versiones nuevas y antiguas de Adobe Illustrator y otros programas de edición vectorial; sin embargo, no es el más indicado para diseños con transparencias.

Muchos impresores prefieren el formato .EPS para asegurarse de que es compatible con su versión de los softwares de edición.

Formato PDF: el formato *Portable Document File* no es el formato vectorial por excelencia, un archivo .PDF también puede usarse como vectorial.

La gran ventaja es que se trata del formato más universal y puede abrirse con software específico de edición con Adobe Illustrator, pero también con programas gratuitos de solo lectura o con un navegador web.

Además, es muy útil a la hora de enviar archivos a imprimir, ya que contiene toda la información necesaria del diseño y no suele provocar problemas de compatibilidad.

Formato SVG: el formato *Scalable Vector Graphics* es un formato basado en XML.

Su uso está muy extendido en el diseño web gracias a que está optimizado como lenguaje de programación y puede ser indexado, por eso suele utilizarse para logotipos, botones y otros elementos web.



1.7. Escala de grises

SE denomina **escala de grises** a un modo de color en el que a cada color se le asigna un valor en una graduación de gris, de tal manera que las imágenes están compuestas por una serie de grises, desde el negro hasta el blanco variando gradualmente en intensidad de grises.

En las imágenes binarias, por el contrario, los colores se representan sin tonalidades de gris entre ambos, únicamente en blanco o en negro.



Figura 1.7: Escala de grises

La escala de grises es un modo de color o estado pictórico, que permite representar una imagen digital mediante distintos tonos de gris (fig. 1.7).

En esta escala cada pixel de la imagen toma un color que es igual a un tono de gris.

Al estar representada en escala de grises, la imagen toma los colores que comúnmente se conocen como *en blanco y negro* (B/N), aunque estrictamente hablando deberíamos utilizar el término *escala de grises*, ya que no sólo se utilizan el blanco y el negro, sino también los tonos grises intermedios (fig. 1.8).



Figura 1.8: Imagen en modo de escala de grises

En el ámbito profesional existen grandes diferencias en cuanto a lo que se denomina una imagen *en blanco y negro* y una imagen en *escala de grises*, ya que las imágenes en *en blanco y negro* son monocromas y no emplean grises, únicamente el negro.

Así pues, en una foto en modo de color de *escala de grises*, todos los colores son negros, blancos o una graduación entre ambos, es decir, grises con diferente tonalidad.

En cambio, en una foto en modo de color *en blanco y negro* (monocromo), se representan los colores sin escala intermedia (fig. 1.9).





Figura 1.9: Imagen en modo de blanco y negro

Las imágenes que se encuentran en escala de grises se componen de diferentes tonos de gris, desde el negro de las zonas más oscuras, hasta el blanco de las partes más claras, variando gradualmente en intensidad de grises.

Generalmente las imágenes digitales en escala de grises emplean 8 bits para representar cada pixel, lo que sólo permite una escala con 256 intensidades de gris o tonos de color diferentes.

Por el contrario, las imágenes digitales en modo de color *en blanco y negro* (monocromo) utilizan sólo 1 bit para representar cada pixel, ya que sólo hay dos opciones, blanco o negro.

Aplicación de la escala de grises en diferentes dispositivos

Aunque en la actualidad se pueden representar imágenes con millones de colores en diferentes dispositivos, la escala de grises sigue siendo una opción muy utilizada, sobre todo para los amantes de la fotografía y los que comúnmente usan un dispositivo electrónico para capturar imágenes.

En la mayoría de aparatos y software relacionados con el manejo de imágenes, existe la posibilidad de capturar o transformar una imagen en color a escala de grises.

Así, la escala de grises aparece como una opción de captura en la mayoría de cámaras digitales y dispositivos que incorporan una cámara de fotos. Al transformar una imagen a escala de grises, cada color de la imagen se convierte a su equivalente en gris.

En la actualidad, muchos fotógrafos profesionales hacen sus trabajos en escala de grises, ya que les ayuda mejor a expresar lo que pretenden a través de sus imágenes.

También en el uso popular de las imágenes es común la aplicación de la escala de grises en las imágenes.

Aunque actualmente se cuenta con tecnología para hacer imágenes en color, la fotografía en blanco y negro se continúa utilizando principalmente porque la ausencia de color permite al fotógrafo concentrarse en lo fotografiado en mayor medida, sin la influencia de los colores.



1.8. Imágenes bi-nivel

Una imagen binaria es una imagen digital que tiene únicamente dos valores posibles para cada píxel. Normalmente, los colores utilizados para su representación son negro y blanco, aunque puede usarse cualquier pareja de colores (fig.1.10).

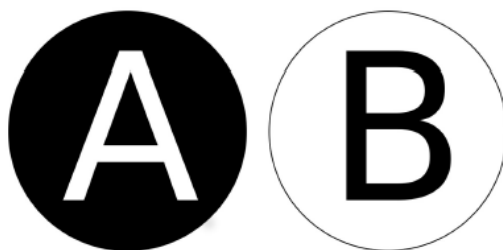


Figura 1.10: Imagen bi-nivel

Uno de los colores se emplea como fondo y el otro para los objetos que aparecen en la imagen.

Los píxeles simplemente toman los valores 0 o 1, tales píxeles constituyen una imagen binaria.

Usualmente, los valores 1 y 0 representan regiones claras y oscuras o el objeto y su fondo. Para obtener una cuantificación más fina de la intensidad de la luz de las imágenes de video, generalmente se usa un byte por píxel, lo que lleva a valores enteros que varían de 0 (negro) a 255 (blanco).

La **binarización** de una imagen consiste en un proceso de reducción de la información de la misma, en la que sólo persisten dos valores: verdadero y falso. En una imagen digital, estos valores, verdadero y falso, pueden representarse por los valores 0 y 1 o, más frecuentemente, por los colores negro (valor de gris 0) y blanco (valor de gris 255).

En el proceso y análisis de imagen, la binarización se emplea para separar las regiones u objetos de interés en una imagen del resto. Las imágenes binarias se usan en operaciones booleanas o lógicas para identificar individualmente objetos de interés o para crear máscaras sobre regiones.

En muchos casos, una imagen binaria es el resultado de una segmentación por niveles de gris o de una segmentación por selección de un rango de color determinado.

En otros casos, una imagen binaria es simplemente el resultado de una selección interactiva de regiones de interés, las cuales se utilizarán como máscaras de comparación o referencia.

Objetos binarios

La formación de objetos binarios en una imagen es una consecuencia directa del proceso de binarización de la misma.

Un objeto binario es una porción de la imagen formada por un conjunto de píxeles con



valor verdadero (p. ej. blancos) conectados entre sí y totalmente rodeados por píxeles con valor falso (p. ej. negro).

En la imagen de la fig. existen tres objetos binarios: un triángulo, un rectángulo y un círculo.

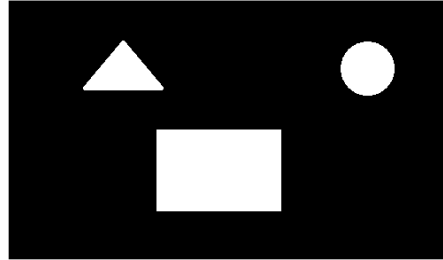


Figura 1.11: Imagen bi-nivel con objetos

Sólo existen dos niveles de intensidad: blanco (nivel de gris 255), que corresponde a verdadero, y negro (nivel de gris cero), que corresponde al falso en la terminología binaria.



1.9. Celdas de resolución

EN una imagen digital la **resolución** viene determinada por la cantidad de elementos independientes por unidad de medida lineal.

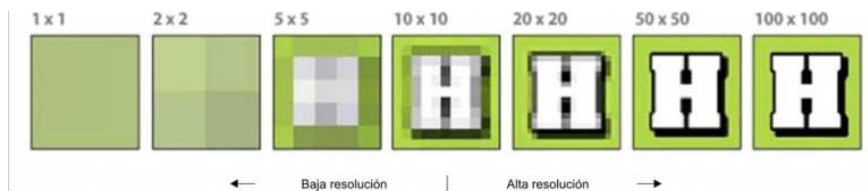


Figura 1.12: Resolución de imagen

Para entenderlo en un lenguaje más coloquial, la cantidad de celdas negras por unidad de medida lineal.

La unidad de medida lineal utilizada para la resolución son los píxeles por pulgada (ppi), que indican los píxeles o celdas por pulgada lineal.

La resolución de la imagen se divide en cuatro tipos distintos:

1. **Resolución Espacial:** Indica la capacidad del sistema para distinguir el objeto más pequeño sobre una imagen. Depende de la distancia de captura al objeto y del tamaño del píxel (ppp o número de píxeles por pulgada).
2. **Resolución Geométrica:** Es la diferencia que hay entre la posición teórica de un píxel y la real.
3. **Resolución Espectral:** Se trata del número y rango de longitud de onda del espectro electromagnético registrado en cada banda de la imagen. En el caso de los sistemas fotográficos podemos encontrar películas pancromáticas, color natural (RGB), infrarrojo b/n o infrarrojo color.
4. **Resolución Radiométrica:** Conocida también como profundidad o contraste radiométrico, indica el número de niveles de gris que son recogidos por cada banda. En imágenes digitales lo más habitual es que contengan 256 niveles por píxel (de 0 a 255).

La **resolución** es una unidad de densidad de imagen, es decir, de concentración de píxeles en línea recta.

Así, una imagen a 72 PPI tendrá 72 celdas distribuidas a lo largo de una línea de 2,54 cm., mientras que una imagen a 300 PPI contendrá 300 celdas a lo largo de una línea de la misma longitud.

Como puede observarse, la longitud de la línea sigue siendo la misma, lo que cambia es el número de celdas o píxeles que se introducen en ese espacio.



Dicho de otro modo, en una imagen a 300 PPI hay mayor concentración de píxeles en una misma unidad de longitud que en una imagen a 72 PPI.

Cuanto mayor sea la resolución mayor será la calidad de la imagen, pues habrá mayor número de celdas para poder definir la forma.

En la imagen se puede ver una simulación de cuadrícula de menor a mayor resolución sobre una letra, y en negro los pixeles que llevarían cada una.

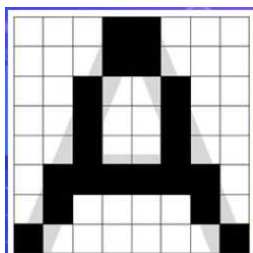


Figura 1.13: Comparativa de resolución

La resolución también se puede expresar en pixels por centímetro (ppcm).

Para calcular la equivalencia entre ppi y ppcm, se tiene que partir del conocimiento que una pulgada es igual a 2,54 cm, así que dependiendo de la medida que se tenga será dividir entre 2,54 (ppi a ppcm) y multiplicar por 2,54 (ppcm a ppi).

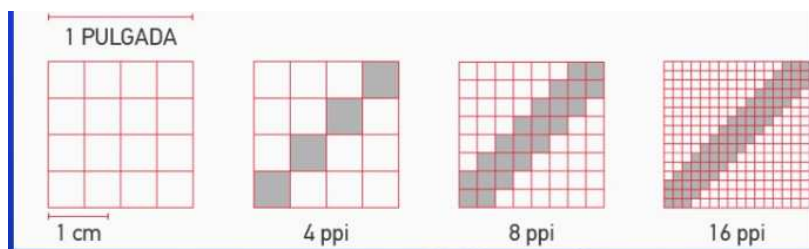


Figura 1.14: Resolución ppi a ppcm

Como ya se ha dicho, la resolución indica la calidad de la imagen, a mayor resolución mayor calidad de la imagen, pero también mayor volumen de información ocupará. La resolución óptima de las imágenes para su posterior impresión es de 300 ppi o 120 ppcm. En caso de tener imágenes con menor resolución y si se quiere que salga con calidad en el impreso, deberán ir reducidas, una vez importadas al 50 %.

Tramado digital

El tramar las imágenes es una necesidad por la limitación de la impresión, para poder conseguir el efecto óptico de una imagen de tono continuo.



Para entenderlo de otra forma, se hace creer ver al ojo un tono continuo donde sólo hay una sucesión de puntos.

Antes se hacía de forma convencional interponiendo una trama entre la imagen original y el material sensible, ahora eso se hace mediante el **tramado digital**.

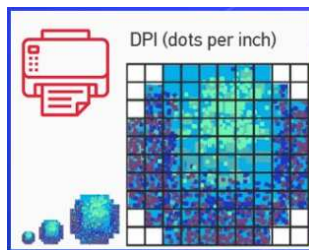


Figura 1.15: Tramado digital dpi

Con los periféricos de salida, impresora láser y filmadora, se puede **tramar** la imagen.

En el tramado de imágenes se debe tener en cuenta la resolución y lineatura de salida, con estos dos parámetros se controla la calidad del trabajo.

Para entender la resolución de salida el concepto es parecido al anteriormente visto, es la cantidad de puntos que puede poner en una pulgada.

A mayor resolución de salida más puntos pondrá y mayor definición se obtiene, las más habituales son: 1200, 2540, 3600 y 4000 píxeles/pulgada.

La **lineatura de salida** es el número de puntos por unidad de longitud (pulgadas o centímetros).

La conversión de líneas por pulgadas a líneas por centímetro y viceversa es igual que lo visto en el apartado anterior.

Las lineaturas más habituales son: 50, 65, 80, 100, 120, 133, 150, 175 y 200 líneas/pulgada.

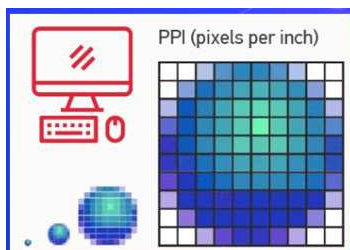


Figura 1.16: Tramado digital ppp

Existe una pequeña relación entre la resolución y la lineatura a la hora de combinarlas, limitada por los periféricos de salida:

Dentro del tramado, digital y convencional, hay otro concepto que es la frecuencia o ángulo de trama, que se aplica a cada uno de los colores de la cuatricomía: cian (15° ó 105°), magenta (75° ó 165°), amarillo (0° ó 90°) y negro (45° ó 135°).



Resolución	Lineatura
1200	50, 65, 80 ,100
2540	120,133
3600	150, 175
4000	200

Cuadro 1.1: longcaption

Cuando se trabaja con imágenes con dos colores se utilizan las inclinaciones del negro y el magenta, en caso de tres negro, cián y magenta.

El amarillo no se utiliza por que debido al ángulo, siempre dará muaré (que es una interferencia visual; un efecto geométrico de distorsión ocasionado por la interacción de dos patrones de trama, situados uno encima del otro).

Al aplicar diferentes ángulos a cada color se evita que se produzca un empastamiento de los puntos de la selección de color, y se crea el efecto óptico llamado roseta.

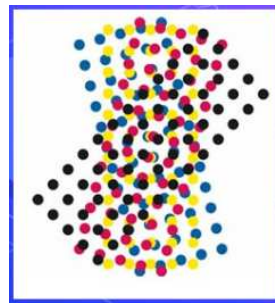


Figura 1.17: Efecto roseta

Esto se puede ver bien en los carteles del metro y en los displays de los autobuses.

En la figura se ve que los ángulos de la lineatura están formados por círculos, a esto se le denomina forma del punto.

Habitualmente se utilizan tres tipos diferentes:

Redondo, cuadrado y elíptico, habiendo también algunos especiales muy poco usados.

Los redondos dan menor visibilidad y se tocan alrededor del 75 % de punto, se suelen utilizar en trabajos a baja lineatura.

Los cuadrados dan mayor sensación de contraste y se tocan alrededor del 50 % de punto.

Los elípticos para imágenes más suaves y mayor sensación de continuidad, sus extremos se tocan en porcentajes a partir de 30 % de punto.



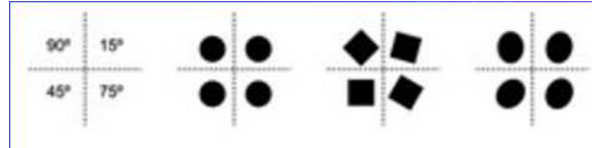


Figura 1.18: Forma del punto

Relación entre resolución de imagen y resolución de impresión

La resolución de la impresora es distinta de la resolución de imagen, pero sí están relacionadas.

Al imprimir una fotografía de gran calidad en impresoras de inyección de tinta, una resolución de imagen de al menos 220 PPI suele ofrecer resultados aceptables.

Asimismo, existe una clasificación estandarizada de las resoluciones óptimas más frecuentes según el sistema de impresión que se puede usar como guía de referencia:

- Impresión offset: 300 PPI (imágenes monocromas: 600 PPI)
- Impresión huecograbado: 300 PPI
- Impresión flexografía: de 150 a 300 PPI
- Impresión serigrafía: de 100 a 200 PPI
- Impresión digital de gran formato: de 72 a 150 PPI
- Impresión digital de pequeño formato: de 200 a 300 PPI



1.10. Conectividad y vecindad

EL propósito de la conectividad es separar los objetos de la escena:

- Del fondo
- Unos de otros
- De los agujeros

Se trata de etiquetar cada región de píxeles continuos con un valor diferente.

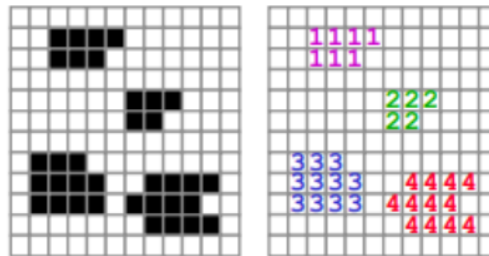


Figura 1.19: Etiquetado

El problema consiste de determinar cuántos objetos, agujeros y fondo existen en una imagen.

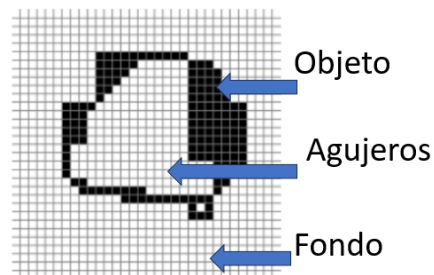


Figura 1.20: Elementos de una imagen

Todo depende de lo que se considere adyacente a un píxel, también de lo que se considere interior, exterior, etc..

Definición (Conectividad)

Es la proximidad espacial entre píxeles de la imagen binaria.



Vecindades

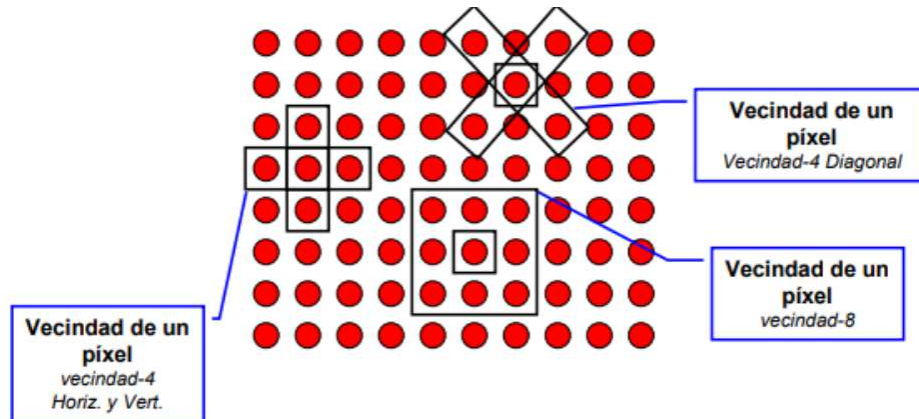


Figura 1.21: Vecindades

Pixel p con coordenadas (x, y)

Vecindad 4

$$V_4(p) = \{(x+1, y), (x-1, y), (x, y+1), (x, y-1)\}$$

$$V_D(p) = \{(x+1, y+1), (x-1, y-1), (x-1, y+1), (x+1, y-1)\}$$

Vecindad 8

$$V_8(p) = \{V_4(p) \cup V_D(p)\}$$

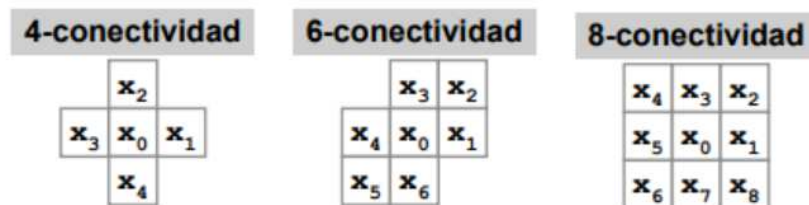


Figura 1.22: Tipos de conectividad



1.11. Componentes conectados

Cuando en una imagen digital se haga referencia a un pixel en particular se emplearán letras minúsculas como p y q .

Y, a un subconjunto de pixeles de una imagen mediante la letra mayúscula S .

A continuación algunas definiciones:

- Un pixel p es adyacente de un pixel q si están **conectados**.
- Dos subconjuntos S_1 y S_2 , son **adyacentes** si algún pixel adyacente de S_1 es adyacente a un pixel de S_2 .
- Un camino desde un pixel p de coordenadas (x, y) a un pixel p de coordenadas (a, b) es una sucesión de diversos pixeles de coordenadas $(x_0, y_0); (x_1, y_1); \dots; (x_n, y_n)$ donde $(x_0, y_0) = (x, y)$ y $(x_n, y_n) = (a, b)$;
- Si p y q son pixeles de un subconjunto S específico de una imagen, se dice que p está conectado a q dentro de S si existe un camino desde p hasta p que consista totalmente de pixeles de S .

Medidas de distancia

Para los pixeles p , q y z , de coordenadas (x, y) , (a, b) y (u, v) respectivamente, D es una función distancia si:

1. $D(p, q) \geq 0$ ($D(p, q) = 0$ si y sólo si $p = q$)
2. $D(p, q) = D(q, p)$
3. $D(p, z) \leq D(p, q) + D(q, z)$

La distancia euclídea entre p y q está definida por:

$$D_e(p, q) = \sqrt{(x - a)^2 + (y - b)^2}$$

La distancia D_4 (distancia city block) entre dos puntos p y q se define como:

$$D_4(p, q) = |(x - a)| + |(y - b)|$$

La distancia D_8 (distancia tablero de ajedrez) entre dos puntos p y q se define como:

$$D_8(p, q) = \max(|x - a|, |(y - b)|)$$



Operaciones aritmético lógicas sobre píxeles

Las operaciones aritmético-lógicas se emplean con mucha frecuencia en el procesamiento de imágenes, a continuación algunas de ellas:

Adición: $p + q$

La suma de píxeles puede usarse para resaltar los contornos de una imagen o, si se suma un valor constante a todos los píxeles de una imagen el efecto es aumentar su brillo.

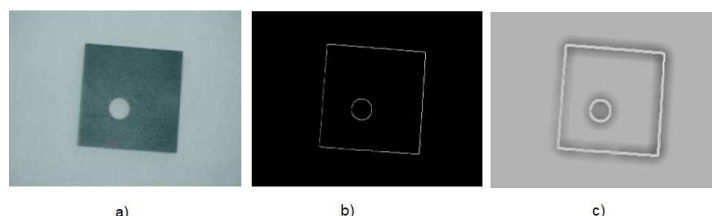


Figura 1.23: a) Original b) Contorno resaltado c) Aumento de brillo

Substracción: $p - q$

El uso común de esta técnica es restar variaciones de iluminación de fondos así como en la substracción absoluta para detectar cambios.



Figura 1.24: a) Resta de fondo b) Detección de diferencias

Multipliación (escalado): $p \times q$

En general un factor de escala mayor a 1 hará la imagen más brillante, y si es menor a 1 la hará más oscura, esta técnica produce mejores resultados que la suma.

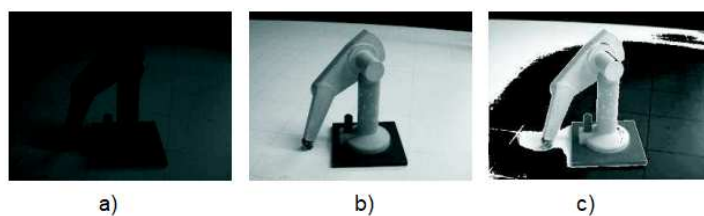


Figura 1.25: a) Oscuro b) Original c) Brillante



División: p/q

Uno de los usos más importantes de la división es la detección de movimiento (cambios), en lugar de obtener el cambio absoluto se obtiene la razón entre los píxeles correspondientes.

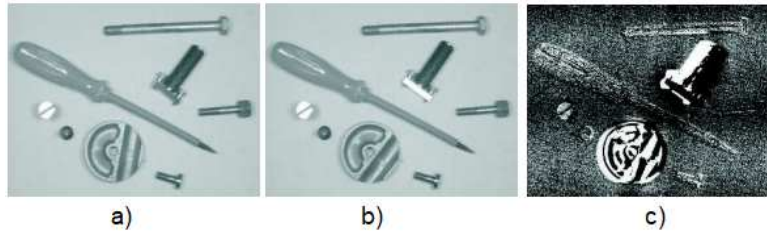


Figura 1.26: a) Original 1 b) Original 2 c) Resultante

AND: $p \& q$

El uso más común de esta operación es calcular la intersección entre dos imágenes, por ejemplo para detectar aquellos objetos en una escena que no han tenido movimiento.

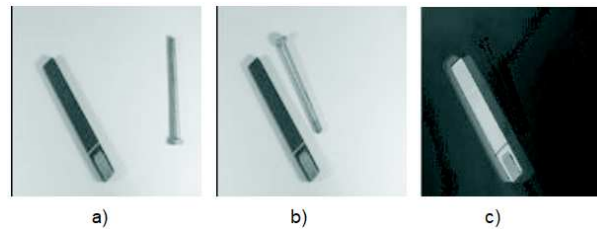


Figura 1.27: a) Original 1 b) Original 2 c) Resultante

OR: $p || q$

Mediante esta operación se obtiene la unión de dos imágenes.

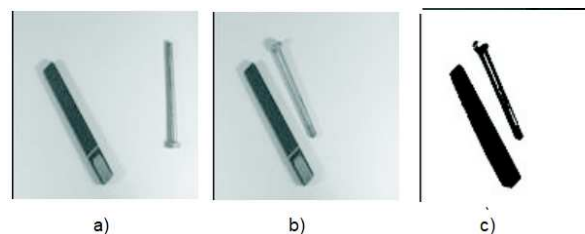


Figura 1.28: a) Original 1 b) Original 2 c) Resultante



XOR: $p \oplus q$

Se utiliza para detectar cambios en imágenes binarias preferentemente.

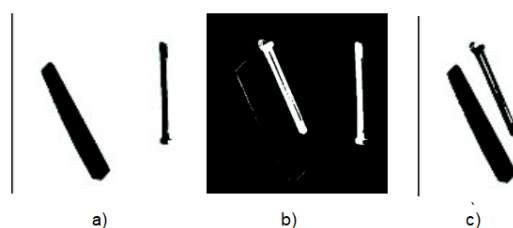


Figura 1.29: a) Original 1 b) Original 2 c) Resultante

NOT: $\sim p$

Esta operación cambia la polaridad de la imagen.

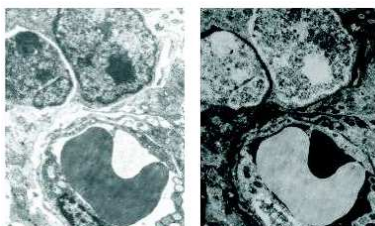


Figura 1.30: a) Original b) Resultante

Imágenes RGB (Truecolor)

Una imagen RGB, a veces denominada imagen de color verdadero, se almacena en MATLAB como una matriz de datos $m \times n \times 3$ que define los componentes de color rojo, verde y azul para cada píxel individual.

Las imágenes RGB no utilizan paleta, el color de cada píxel está determinado por la combinación de las intensidades de rojo, verde y azul almacenadas en cada plano de color en la ubicación del píxel.

Los formatos de archivo de gráficos almacenan imágenes RGB como imágenes de 24 bits, donde los componentes rojo, verde y azul son de 8 bits cada uno.

Esto produce un potencial de 16 millones de colores.

La precisión con la que se puede replicar una imagen de la vida real ha dado lugar al sobrenombre de imagen en color verdadero.

Una matriz RGB MATLAB puede ser de clase double, uint8 o uint16.

En una matriz RGB de clase double, cada componente de color es un valor entre 0 y 1.

Un píxel cuyos componentes de color son (0,0,0) se muestra como negro y un píxel cuyos componentes de color son (1,1,1) se muestra en blanco.



Los tres componentes de color para cada píxel se almacenan a lo largo de la tercera dimensión de la matriz de datos.

Para guardar la imagen con un formato específico:

Mostrar imágenes Truecolor

Las imágenes Truecolor (también llamadas imágenes RGB), representan los valores de color directamente, en lugar de a través de un mapa de colores.

Una imagen de color verdadero es una matriz $m \times n \times 3$.

Para cada píxel (r, c) en la imagen, el color está representado por el triplete (r, c, 1: 3).

Para mostrar una imagen en color verdadero, llamar a la función `imshow` o abrir la aplicación Visor de imágenes.

Por ejemplo, este código lee una imagen de color verdadero en el espacio de trabajo de MATLAB y luego muestra la imagen.

Esta documentación utiliza el nombre de variable `RGB` para representar una imagen de color verdadero en el espacio de trabajo

```
>RGB = imread('peppers.png');  
>imshow(RGB)
```

El tipo de dato matriz, que contendrá una imagen puede ser de varios tipos (según el tipo de dato de cada píxel):

- `double`: Doble precisión, números en punto flotante que varían en un rango aproximado de -10308 a 10308 (8 bytes por elemento)
- `uint8`: Enteros de 8 bits en el rango de [0,255] (1 byte por elemento)
- `uint16`: Enteros de 16 bits en el rango de [0, 65535] (2 bytes por elemento)
- `uint32`: Enteros de 32 bits en el rango de [0, 4294967295] (4 bytes por elemento)
- `int8`: Enteros de 8 bits en el rango de [-128, 127] (1 byte por elemento)
- `int16`: Enteros de 16 bits en el rango de [-32768, 32767] (2 bytes por elemento)
- `int32`: Enteros de 32 bits en el rango de [-2147483648,2147483647] (4 bytes por elemento)
- `logical`: Los valores son 0 ó 1 (1 bit por elemento)

Los sistemas que utilizan 24 bits por píxel de pantalla pueden mostrar imágenes en color verdadero directamente, porque asignan 8 bits (256 niveles) a cada uno de los planos de color rojo, verde y azul.

En sistemas con menos colores, `imshow` muestra la imagen usando una combinación de aproximación de color y difuminado.

Nota Si muestra una imagen en color y aparece en blanco y negro, compruebe si la imagen es una imagen indexada.



Con imágenes indexadas, debe especificar el mapa de color asociado con la imagen.

Agregar barra de color a la imagen desplegada

Este ejemplo muestra cómo mostrar una imagen en escala de grises con una barra de color que indica la asignación de valores de datos a colores.

Ver la correspondencia entre los valores de los datos y los colores mostrados mediante el uso de una barra de colores es especialmente útil si está mostrando datos de rango no convencionales como una imagen, la función es:

```
colorbar
```



FILTROS ESPACIALES

“Utilizar las funciones y matrices para representar una imagen y buscar parámetros propios de la imagen a través de su histograma.”

Objetivo específico 1 y 2

2.1. Funciones de intensidad

EN general una imagen es la representación visual de uno o varios objetos en un entorno específico.

Los **filtros** son operaciones que se aplican a los píxeles de una imagen para optimizarla, enfatizar cierta información o conseguir un efecto especial en ella.

Los principales objetivos que se persiguen con la aplicación de filtros son:

- **Suavizar la imagen:** reducir la cantidad de variaciones de intensidad entre píxeles vecinos.
- **Eliminar ruido:** eliminar aquellos píxeles cuyo nivel de intensidad es muy diferente al de sus vecinos y cuyo origen puede estar tanto en el proceso de adquisición de la imagen como en el de transmisión.
- **Realzar bordes:** destacar los bordes que se localizan en una imagen.
- **Detectar bordes:** detectar los píxeles donde se produce un cambio brusco en la función intensidad.

El proceso de filtrado consiste en la aplicación a cada uno de los pixels de la imagen de una matriz de filtrado de tamaño $N \times N$ (generalmente de 3×3 aunque puede ser mayor) compuesta por números enteros y que genera un nuevo valor mediante una función del valor original y los de los pixels circundantes. El resultado final se divide entre un escalar, generalmente la suma de los coeficientes de ponderación.



2.2. Histograma

Una de las primeras herramientas que proporcionan información sobre las intensidades de una imagen (o colores) es su **histograma**.

El histograma es una representación de la frecuencia relativa de cada color de una imagen.

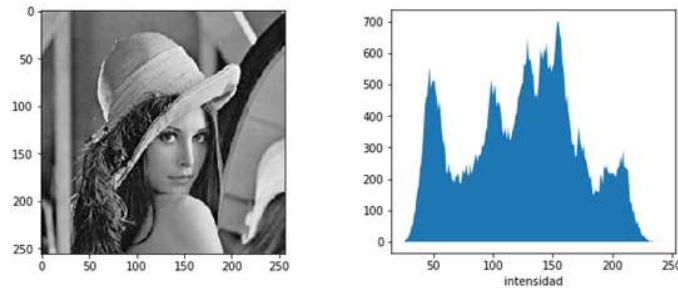


Figura 2.1: a) Imagen b) Histograma

Histograma de una imagen a color

Las imágenes a color poseen tres matrices para representar los canales de RGB, es posible extraer esos canales y obtener el histograma de cada uno de manera independiente.

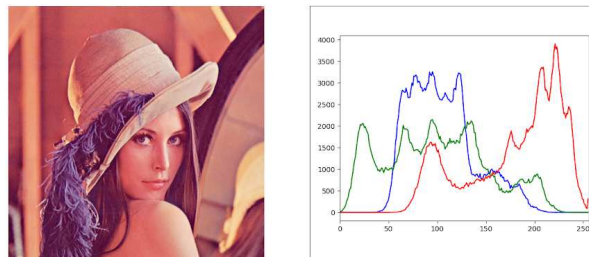


Figura 2.2: a) Imagen b) Histograma

Según su histograma, podemos clasificar las imágenes como:

De **bajo contraste** (*low-contrast*): Los valores están claramente agrupados en una región del histograma.

Oscuras (*dark*): La mayoría de los valores están en la parte baja del histograma.

Claros (*bright*): La mayoría de los valores están en la parte alta del histograma.

De **alto contraste** (*high-contrast*): Los valores están extendidos a lo largo de todo el histograma (histograma uniforme).



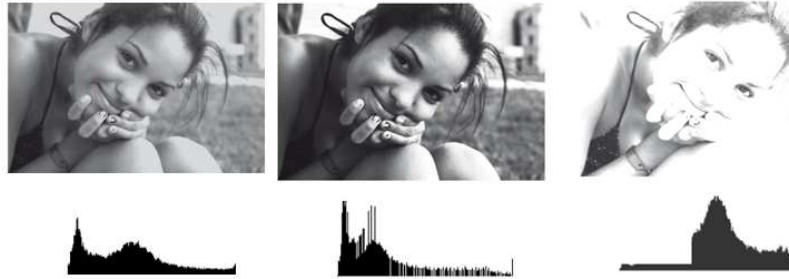


Figura 2.3: a) Imagen original b) Bajo contraste oscura c) Bajo contraste clara

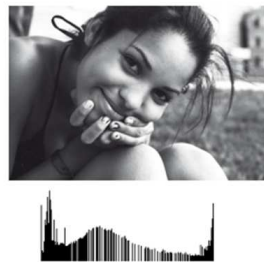


Figura 2.4: a) Imagen Alto contraste

Generación del Histograma

Se va a usar la imagen que representa la fig.2.5 para comprender como se genera un histograma.

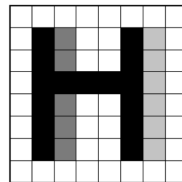


Figura 2.5: a) Trama de Imagen

Se trata de una imagen en niveles de gris muy simple, de 8x8 píxeles de tamaño (se han señalado los límites entre píxeles para facilitar su identificación). Sólo son posibles 4 niveles de gris, porque se van a usar 2 bits para codificar el brillo de cada píxel.

De la forma habitual, los niveles de gris se numeran del 0 al 3, correspondiendo un brillo mayor a los valores más altos.

La tabla 2.1 resume esto:



Nivel de gris	Brillo
0	Negro
1	Gris oscuro
2	Gris claro
3	Blanco

Cuadro 2.1: Intensidades

La gráfica que aparece en la fig.2.6 es el histograma correspondiente a la imagen anterior.

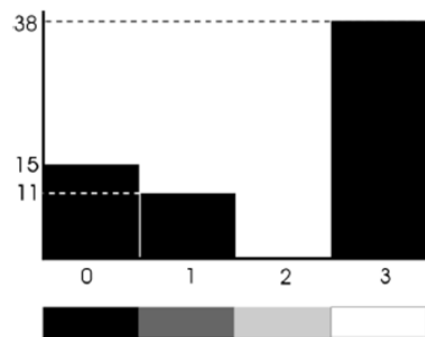


Figura 2.6: Histograma correspondiente a la fig. 2.5

Los números que aparecen el eje horizontal representan los niveles de gris que pueden aparecer en la imagen: a la izquierda está el valor más oscuro (negro) y en el extremo derecho el más claro (blanco).

El resto de niveles se distribuyen uniformemente. Se ha puesto una escala con los tonos de gris correspondientes para facilitar la comprensión.

En un histograma real habitualmente no encontrará numerado el eje vertical, ni la escala de tonos para el eje horizontal.

La altura de cada barra representa el número de píxeles de la imagen que presentan ese nivel de gris concreto. Se puede deducir entonces que la imagen tiene 15 píxeles completamente negros (con nivel 0), 11 de tono gris oscuro (nivel 1) y 38 píxeles completamente blancos (nivel 3). No hay ningún píxel en la imagen con un nivel de gris 2.

¿Cuánto deberán sumar las alturas de todas las barras? Efectivamente, 64, que es el número total de píxeles que tiene la imagen.

Evidentemente, con sólo mirar el histograma se puede deducir algunas cosas interesantes sobre la imagen, lo que demuestra su utilidad:

La mayor parte de los píxeles son blancos, así que probablemente se aprecie un fondo blanco uniforme.



Hay un número significativo de píxeles totalmente blancos y totalmente negros, por lo que presentará un aspecto bien contrastado.

Operaciones

El histograma se puede modificar utilizando funciones de mapeo:

- Expansión Lineal (Stretch)
- Compresión (Shrink)
- Desplazamiento (Slide)
- Ecuación

Expansión lineal (stretch)

Consiste en hacer que el mínimo y el máximo de los valores presentes en la imagen, correspondan con dos puntos seleccionados del histograma (los extremos: 0 y $2^B - 1$).

$$\text{stretch}((I)(x, y)) = \left[\frac{I(x, y)}{\max - \min} \right] [MAX - MIN] + MIN$$

Es una forma de *estirar* el histograma escalando los valores intermedios sobre todo el rango de valores disponibles

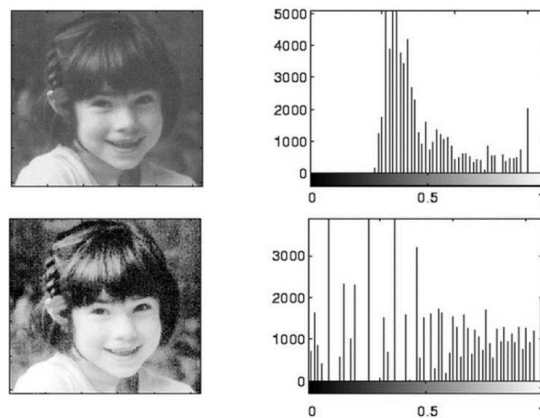


Figura 2.7: Expansión lineal

Compresión (shrink)

$$\text{shrink}((I)(x, y)) = \left[\frac{\text{shrink}_{\max} - \text{shinrk}_{\min}}{\max - \min} \right] [I(x, y) - \min] + \text{shrink}_{\min}$$



Desplazamiento (slide)

$$slide((I)(x, y)) = I(x, y) + OFFSET$$

Ecualización del histograma

Busca producir una imagen con un histograma uniforme, lo que resulta en un aumento del contraste.

El objetivo es generar una imagen con valores proporcionales a su valor y a su frecuencia en la imagen original.

Además de considerar los extremos del histograma y también considera su forma.

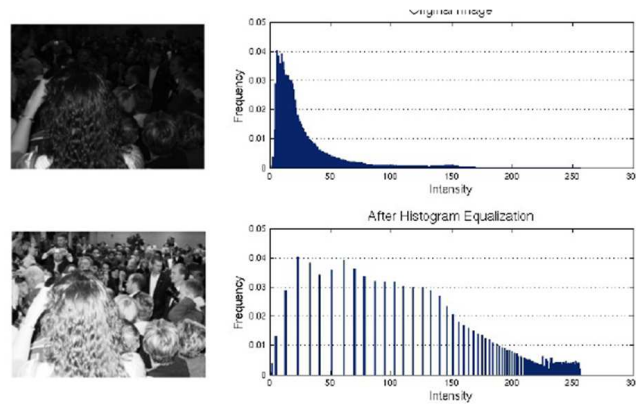


Figura 2.8: Ecualización



2.3. Filtros espaciales

EL **filtrado** es una operación de vecindad, en la que el valor de cualquier píxel dado en la imagen de salida se determina aplicando algún algoritmo a los valores de los píxeles en la vecindad del píxel de entrada correspondiente.

Los filtros espaciales pueden clasificarse basándose en su linealidad en filtros lineales y en filtros no lineales.

Filtros lineales

Los filtros lineales pueden ser clasificados según las frecuencias que dejen pasar: los filtros paso bajo atenúan o eliminan las componentes de alta frecuencia a la vez que dejan inalteradas las bajas frecuencias; los filtros paso altas atenúan o eliminan las componentes de baja frecuencia con lo que agudizan las componentes de alta frecuencia; los filtros paso banda eliminan regiones elegidas de frecuencias intermedias.

A continuación se describe el uso de los diferentes filtros:

- **Filtros paso bajas:** son utilizados en la reducción de ruido; suavizan y aplanan un poco las imágenes y como consecuencia se reduce o se pierde la nitidez, en inglés son conocidos como *Smoothing Spatial Filters*.
- **Filtros paso altas:** estos filtros son utilizados para detectar cambios de luminosidad. Son utilizados en la detección de patrones como bordes o para resaltar detalles de una imagen. En inglés son conocidos como *Sharpening Spatial Filters*. Los filtros *Unsharp Masking* son filtros paso altas usados en el mejoramiento de la nitidez o de la calidad visual de una imagen.
- **Filtros paso banda:** son utilizados para detectar patrones de ruido. Ya que un filtro paso banda generalmente elimina demasiado contenido de una imagen casi no son usados, sin embargo, los filtros paso banda son útiles para aislar los efectos de ciertas bandas de frecuencias seleccionadas sobre una imagen. De esta manera, estos filtros ayudan a simplificar el análisis de ruido, razonablemente independiente del contenido de la imagen.

Filtros lineales en Matlab

Hay dos niveles de frecuencia presentes en una imagen, los bordes y el ruido de una imagen representan componentes de alta frecuencia y las áreas suaves representan componentes de baja frecuencia.

Los componentes de alta frecuencia tienen un nivel de alta intensidad o alto contraste, y los componentes de baja frecuencia tienen un nivel de baja intensidad o bajo contraste.

Suponer que una imagen tiene bordes afilados o ruido y se quiere suavizarla. En ese caso, se tiene que crear un filtro de paso bajo que permita que los componentes o píxeles de baja frecuencia pasen y detengan los componentes o píxeles de alta frecuencia, bajando el nivel de intensidad de los píxeles en la imagen.



Filtros no lineales

Funcionan ordenando los valores en la vecindad de cada punto de menor a mayor, y obteniendo algún valor a partir de la lista ordenada.

Los filtros no lineales se clasifican en:

- **Filtro de Máximo:** Selecciona el mayor valor dentro de una ventana ordenada de valores de nivel de gris; elimina el ruido pimienta (puntos negros), sólo funciona en este caso y como consecuencia tiende a aclarar la imagen.
- **Filtro de Mínimo:** Selecciona el menor valor dentro de una ventana ordenada de valores de nivel de gris; elimina el ruido sal (puntos blancos), sólo funciona en este caso y como consecuencia tiende a obscurecer la imagen.
- **Filtro de la Mediana:** Selecciona el valor en la posición intermedia, se suele usar para eliminar ruido en la imagen.

Convolución

Para realizar el filtrado lineal de una imagen, se debe llevar a cabo una operación denominada **convolución**.

La convolución es una operación de entorno en la que cada píxel de salida es la suma ponderada de los píxeles de entrada del entorno.

La matriz de las ponderaciones se denomina el núcleo de convolución, también conocido como el **filtro**.

Un núcleo de convolución es un núcleo de correlación que se ha girado 180 grados.

Por ejemplo, suponer una imagen representada por la matriz:

$$A = \begin{bmatrix} 17 & 24 & 1 & 8 & 15 \\ 23 & 5 & 7 & 14 & 16 \\ 4 & 6 & 13 & 20 & 22 \\ 10 & 12 & 19 & 21 & 3 \\ 11 & 18 & 25 & 2 & 9 \end{bmatrix}$$

y el núcleo de correlación es:

$$h = \begin{bmatrix} 8 & 1 & 6 \\ 3 & 5 & 7 \\ 4 & 9 & 2 \end{bmatrix}$$

Debería usar los pasos siguientes para calcular el píxel de salida en la posición (2, 4):

1. Girar el núcleo de correlación 180 grados a partir del elemento central para crear un núcleo de convolución.
2. Deslizar el elemento central del núcleo de convolución de modo que quede encima del elemento (2, 4) de A.



3. Multiplicar cada ponderación del núcleo de convolución girado por el píxel de A de debajo.
4. Sumar los productos individuales del paso 3.

Así pues, el píxel de salida (2, 4) es:

$$1 * 2 + 8 * 9 + 15 * 4 + 7 * 7 + 14 * 5 + 16 * 3 + 13 * 6 + 20 * 1 + 22 * 8 = 575$$

En la figura 2.9 se muestra el cálculo.

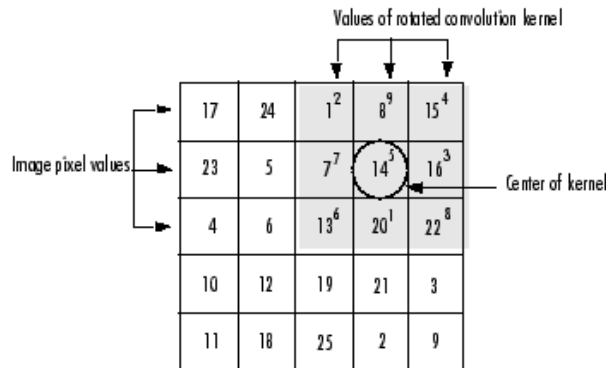


Figura 2.9: Calcular la salida (2, 4) de convolución

Correlación

La operación llamada correlación está estrechamente relacionada con la convolución.

En correlación, el valor de un píxel de salida también se calcula como una suma ponderada de los píxeles vecinos.

La diferencia es que la matriz de ponderaciones, en este caso llamada núcleo de correlación, no se rota durante el cálculo.

La figura 2.10 muestra cómo calcular el píxel de salida (2,4) de la correlación de A, asumiendo que has un kernel de correlación en lugar de un kernel de convolución, siguiendo estos pasos:

1. Deslizar el elemento central del núcleo de correlación de modo que quede encima del elemento (2,4) de A.
2. Multiplicar cada peso en el núcleo de correlación por el píxel de Abajo
3. Suma los productos individuales.

El píxel de salida (2,4) de la correlación es:

$$1 * 8 + 8 * 1 + 15 * 6 + 7 * 3 + 14 * 5 + 16 * 7 + 13 * 4 + 20 * 9 + 22 * 2 = 585$$



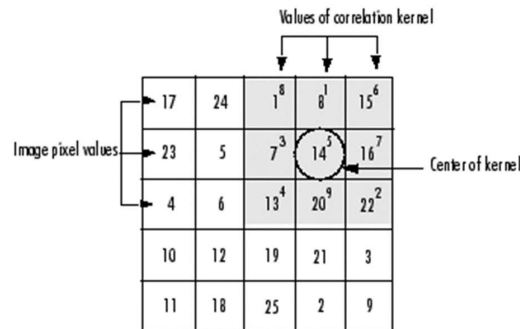


Figura 2.10: Calcular las salidas 2, 4 de la correlación

2.3.1. Ruido

Las imágenes digitales son propensas a varios tipos de ruido. El ruido es el resultado de errores en el proceso de adquisición de la imagen que dan lugar a valores de los píxeles que no reflejan las verdaderas intensidades de la escena real. Hay varias formas en las que se puede introducir ruido en una imagen, dependiendo de cómo se cree la imagen. Por ejemplo:

Si la imagen se escanea a partir de una fotografía hecha en película, el grano de la película es una fuente de ruido. El ruido también puede ser el resultado de daños en la película, o que lo introduzca el propio escáner.

Si la imagen se adquiere directamente en formato digital, el mecanismo de recogida de datos (como un detector CCD¹) puede introducir ruido.

La transmisión electrónica de los datos de la imagen puede introducir ruido.

Definición (Ruido digital)

Es la aparición aleatoria de señales que no pertenecen a la imagen original, es decir, es la aparición de píxeles o conjunto de píxeles de un color que no corresponde con el de la imagen.

Tipos de ruido

Las principales fuentes de ruido surgen en la adquisición y/o transmisión de las imágenes.

Ruido Gaussiano

En el ruido de tipo Gaussiano, todos los píxeles que componen la imagen cambian su valor en base a una distribución normal o gaussiana. La distribución mostrada en la figura 2.11, el ruido está centrado en el valor Z (que se corresponde con el valor de intensidad del píxel) y tiene una desviación estándar de sigma.

¹ CCD proviene del inglés *Charge-Coupled Device*



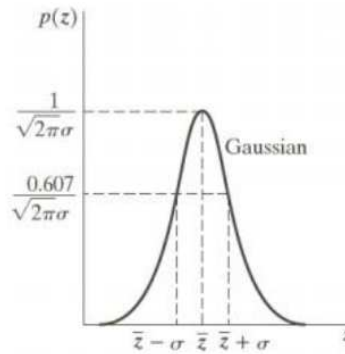


Figura 2.11: Función de distribución del ruido gaussiano

El efecto de este ruido es que en general se generan valores aleatorios que variarán poco respecto del valor original del píxel. La mayor parte de los valores establecidos quedaran similares al original pero con alguna pequeña modificación tal y como puede verse en la figura 2.12, donde la imagen de la izquierda muestra la imagen original y a su derecha la imagen con ruido gaussiano.

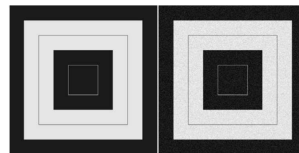


Figura 2.12: Ejemplo de ruido gaussiano

Ruido Impulsivo (sal y pimienta)

El ruido Sal y pimienta es un caso particular del ruido impulsivo. El valor que toma el píxel no tiene relación con el valor original, sino que toma valores muy altos o muy bajos, es decir, casi blanco o negro.

Matemáticamente este ruido se modela con una función de distribución no gaussiana o escalón como la mostrada en la figura 2.13.

El efecto final de este ruido es que ciertos píxeles de forma aleatoria cambian a el valor máximo (sal=blanco) o el mínimo (pimienta=negro), tal y como puede verse en la figura 2.14, donde la imagen de la izquierda muestra la imagen original y a su derecha la imagen con ruido impulsivo Sal y Pimienta.

Ruido de Poisson

El ruido de Poisson (o de disparo) aparece cuando hay mucha diferencia entre la luminosidad de una zona con respecto a la luminosidad media de la toma.



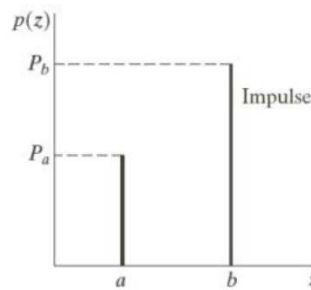


Figura 2.13: Función de distribución del ruido impulsivo



Figura 2.14: Ejemplo de ruido sal y pimienta

Es decir, cuando el sensor capta mucha menos luz en una zona concreta con respecto a la luz media que tiene en el resto, ahí es cuando aparece este tipo de ruido. El ruido de disparo suele aparecer en condiciones de escasa iluminación.



Figura 2.15: Ejemplo de ruido de Poisson

Ruido Speckle

El ruido Speckle (o moteado) aparece en imágenes con iluminación coherente, como las de un scanner ultrasónico, sonar y radar de abertura sintética (SAR).

Este ruido se desvía del modelo clásico, en el que se supone ruido Gaussiano independiente de la señal y adicionado al verdadero valor.



Es multiplicativo y no Gaussiano (en los formatos intensidad y amplitud), y dificulta la interpretación de las imágenes porque el efecto sal y pimienta corrompe la información.

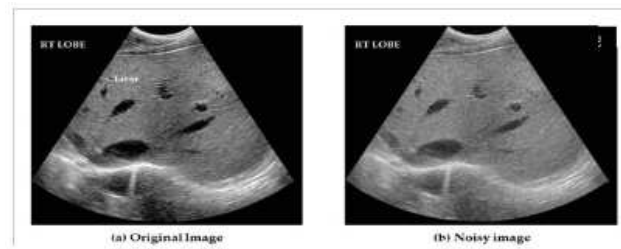


Figura 2.16: Ejemplo de ruido Speckle

Funciones para generar ruido en Matlab

Ruido gaussiano

```
I = imnoise(I, 'gaussian')
```

Ruido Poisson

```
I = imnoise(I, 'poisson')
```

Ruido Sal & Pimienta

```
I = imnoise(I, 'salt&pepper')
```

Ruido Speckle

```
I = imnoise(I, 'speckle')
```

Eliminación de ruido

Una forma de eliminar el ruido de una imagen es mediante el suavizado de imágenes, por ejemplo el ruido gaussiano puede ser eliminado mediante un **filtro lineal pasa-bajos**.

Una posible desventaja con el uso de filtros pasa-bajo para eliminar el ruido de imágenes consiste en que los bordes de los objetos pueden volverse borrosos.

En matlab la función **imfilter** sirve para aplicar una mascara especifica que filtra en este caso las frecuencias altas.

La siguiente máscara, aplicada mediante convolución sirve para eliminar el ruido gaussiano:

$$h = \begin{bmatrix} 0 & 1/8 & 0 \\ 1/8 & 1/2 & 1/8 \\ 0 & 1/8 & 0 \end{bmatrix}$$



En el caso del **ruido de tipo poisson**, se puede utilizar la siguiente máscara:

$$h = \begin{bmatrix} 1/16 & 1/8 & 1/16 \\ 1/8 & 1/16 & 1/8 \\ 1/16 & 1/8 & 1/16 \end{bmatrix}$$



DOMINIO DE FRECUENCIAS

“Comprender que la representación de cualquier señal puede representarse en términos de la serie de Fourier; conocer el límite de las series de Fourier para deducir la Transformada de Fourier y aplicarla al caso de imágenes y utilizar el filtrado de imágenes para mejorarlas o restaurarlas..”

Objetivos específicos 1, 2 y 3

3.1. Transformada de Fourier

Algunos problemas de análisis de imágenes son más fáciles de resolver transformando la imagen del dominio espacial a otro dominio, realizar dichas tareas en ese nuevo dominio, y aplicar una transformación inversa para regresar al dominio espacial con el problema resuelto.

La **transformada de Fourier** se utiliza en un amplio rango de aplicaciones tales como análisis, filtrado, reconstrucción y compresión de imágenes.

La transformada de Fourier es una operación matemática usada para transformar funciones entre el dominio del tiempo o del espacio al dominio de la frecuencia y viceversa.

El concepto de *Transformada de Fourier* se refiere a varios elementos de forma simultánea que son:

- Operación de transformación de una función
- Función resultado de la operación
- Espectro de frecuencias de una función

La función original suele recibir el nombre de $x(t)$ siendo muy común que t sea el tiempo, mientras que la función de la transformada suele recibir el nombre de $X(f)$ en mayúscula, siendo la f la frecuencia.

Por ejemplo, si se tiene la función $p(t)$, donde p es la potencia de una señal acústica y t el tiempo, $P(f)$ es una transformada de Fourier que informa de cómo se distribuye $p(t)$ en función de la frecuencia de potencia de la señal.

En otro ejemplo, si $e(s)$ informa sobre la energía de una señal en función del espacio, $E(f)$ es la transformada de Fourier que informa cómo se distribuye $e(s)$ en función de la frecuencia de energía de la señal.

Es importante destacar que aunque el tiempo, el espacio y la frecuencia son valores reales, tanto $x(t)$ y $x(s)$ como sus respectivas transformadas $X(f)$ no tienen porqué tomar valores reales.

En los ejemplos de potencia y energía es posible que las magnitudes tengan **elementos complejos**.

A pesar de que calcular la transformada de Fourier de una función es bastante sencillo, es necesario aprender a **integrar múltiplos de funciones complejas** con anterioridad.

Si $x(t)$ es la función original su transformada $X(f)$ será:

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt$$

Para obtener la transformada es importante tener presente varios puntos importantes:

- **La integral actúa sobre 't'** ('s' si la función original depende del espacio) de modo que la frecuencia 'f' debe ser tratada como si fuera una constante, de hecho, el exponente $-j2\pi f$ es una constante.
- Si la función original existe entre dos valores reales a y b, la **primitiva de la integral** se resuelve entre esos dos valores reales.

Ejemplo

Expandir la siguiente función en una serie de Fourier.

$$f(x) = \begin{cases} 0 & \text{para } -\pi < x < 0 \\ \pi - x & \text{para } 0 \leq x < \pi \end{cases}$$

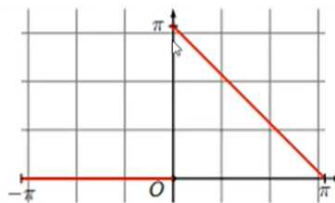


Figura 3.1: Función a expandir



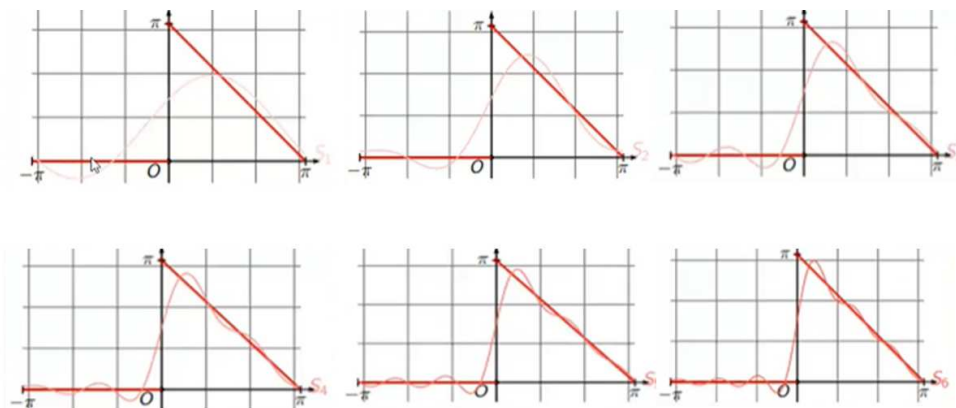


Figura 3.2: Elementos de la serie de expansión

La transformada de Fourier juega un papel muy importante en el procesamiento de imágenes, ya que es una herramienta que permite obtener la representación de información del espacio de frecuencias y aplicando un operador en este dominio, se puede operar sobre la imagen para detectar y realzar bordes, eliminación de ruido etc..

La transformada de Fourier es una representación de una imagen como una suma de exponenciales complejas de diferentes magnitudes, frecuencias y fases.

La transformada de Fourier es fundamental en una amplia gama de aplicaciones de procesamiento de imágenes incluida la mejora, el análisis, la restauración y la compresión.

La transformada de Fourier es útil para descomponer una señal en el tiempo en sus componentes de frecuencia (armónicos).

En la figura 3.3 la curva roja representa una señal en el tiempo y sus componentes de frecuencia se proyectan en las barras azules.

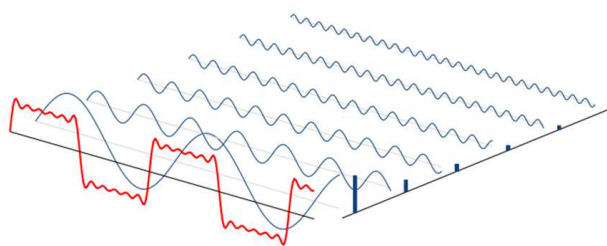


Figura 3.3: Transformada de Fourier

La integración de todas las señales senoidales reconstruye la señal en el tiempo.

En pocas palabras, la transformada de Fourier es una forma de dividir algo en un montón de ondas sinusoidales.



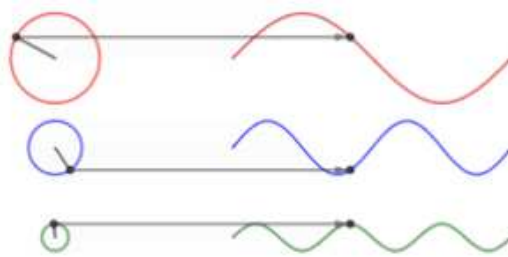


Figura 3.4: Generación de la función senoidal

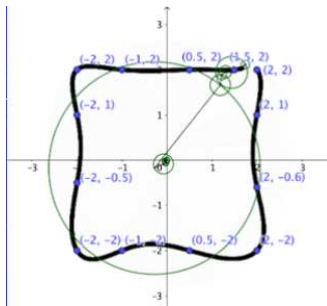


Figura 3.5: Trazado de figuras mediante senoidales

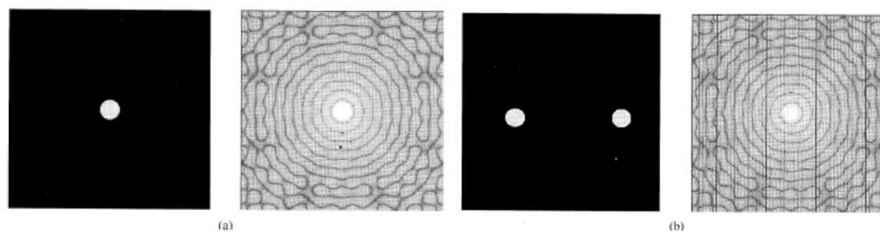


Figura 3.6: Transformadas de figuras simples



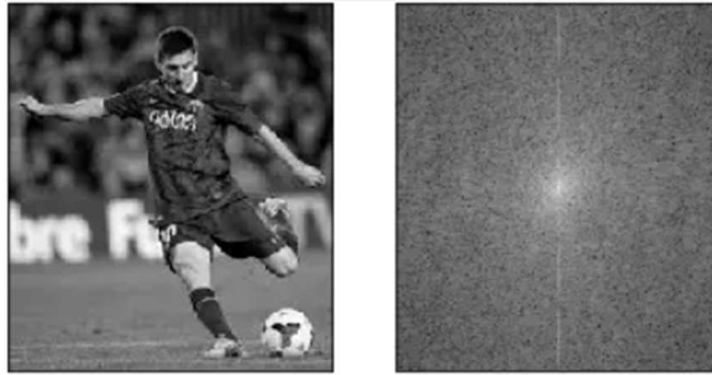


Figura 3.7: Transformada de imagen compleja

Transformada Discreta de Fourier

Suponiendo funciones bidimensionales discretas (como es el caso de las imágenes digitales), la transformada discreta de Fourier o **DFT** (*Discrete Fourier Transform*) se obtiene empleando sumatorias en lugar de integrales:

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

Donde M representa el número de columnas (píxeles en la dirección X) y N número de filas (píxeles en la dirección Y).

Transformada inversa

La transformación inversa viene dada por la expresión:

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

El espectro de Fourier, la fase y el espectro de energía se calculan de la misma forma que en el caso de una función continua, con la única diferencia de que las variables independientes son ahora discretas.

El componente de fase es relevante en cuanto a la estructura espacial de la imagen y contiene información sobre la posición relativa de las características u objetos de la imagen.

Propiedades de la transformada discreta

Algunas propiedades importantes de la transformada discreta de Fourier son:

- La **separabilidad** o posibilidad de escribir la transformada de la forma (suponiendo una imagen cuadrada de dimensión NxN):



$$F(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} e^{-j\frac{2\pi ux}{N}} * \sum_{y=0}^{N-1} f(x, y) * e^{-j\frac{2\pi vy}{N}}$$

Cuya principal ventaja es que, tanto $f(x,y)$ como $F(u,v)$ pueden obtenerse en dos pasos sucesivos, aplicando la transformada de Fourier unidimensional o su inversa para cada fila y después para cada columna.

- La **traslación** o desplazamiento de la función $f(x)$ no afecta a la magnitud de su transformada de Fourier. Resulta interesante debido a que el análisis visual de la transformada suele limitarse al estudio de su magnitud o módulo.
- **Periodicidad** la DFT tiene un período de longitud N . La magnitud de la transformada está centrada en el origen. Para visualizar un período completo, basta con desplazar el origen de la transformada al centro de la imagen ($U=N/2$, $V=N/2$). De esta forma, al trasladar $F(0,0)$ al centro de la imagen, el pixel central en la imagen de la transformada de Fourier representa el valor de la intensidad media en la imagen.
- La **rotación** de $f(x,y)$ en un ángulo θ , genera una rotación de $F(u,v)$ en el mismo ángulo y viceversa.
- La transformada y su inversa son distributivas en la suma pero no en la multiplicación.
- **Valor medio** el primer componente de la transformada $F(0,0)$, representa el valor medio de la función discreta original, tal y como se ha comentado anteriormente.



3.2. Filtrado en el dominio de frecuencias

Los valores de los píxeles en la imagen de la transformada de Fourier, $F(u,v)$, representan los componentes de frecuencia espacial, verticales y horizontales, de la imagen original, análogamente a la frecuencia de una señal en el tiempo.

Así, una señal sinusoidal con alta frecuencia, por ejemplo, varía muy rápidamente, mientras que otra con una frecuencia baja varía lentamente en el tiempo.

De la misma forma, una imagen con una elevada frecuencia espacial en la dirección horizontal, contiene cambios frecuentes de intensidad en la dirección horizontal

Los pasos a seguir para realizar un filtrado en el dominio frecuencial son:

1. Transformar la imagen original al dominio de frecuencias aplicando la FFT (*Fast Fourier Transform*) o transformada rápida de Fourier
2. Aplicar el filtrado en el dominio de frecuencias.
3. Aplicar la transformada de Fourier inversa para devolver la imagen filtrada al dominio espacial.

La ventaja de filtrar en el dominio frecuencial es que es matemáticamente más sencillo, debido a que la transformada de Fourier de la convolución de dos funciones es igual a la suma de las transformadas de Fourier de cada una de ellas.

Es decir, en el dominio de frecuencias basta con sumar las transformadas de la imagen y del filtro, mientras que en el dominio espacial hay que aplicar una convolución.

En las figuras se muestran ejemplos de tres imágenes con sus respectivas transformadas. En ellas se pueden ver tres distribuciones espaciales distintas y el efecto que tienen sobre la representación en forma de imagen de sus transformadas, incluyendo el claro componente direccional de las frecuencias.

La primera de ellas representa una escena agrícola con profusión de pequeñas parcelas de distintos cultivos, lo que conlleva a una gran cantidad de altas frecuencias en varias direcciones, generando una apariencia de estrella en su transformada, cuyos brazos son perpendiculares a las orientaciones dominantes.

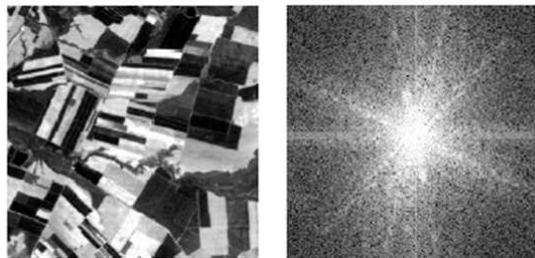


Figura 3.8: Escena agrícola y su Transformada



La segunda imagen representa un paisaje dunar, con una orientación dominante (NE-SW) debida a la acción de los vientos en la zona.

Su transformada de Fourier se muestra en forma de nube, debido a que existe poca definición en la orientación dominante, alargada según la dirección perpendicular a la orientación de las dunas, es decir, según la dirección con mayor variación de la frecuencia espacial.

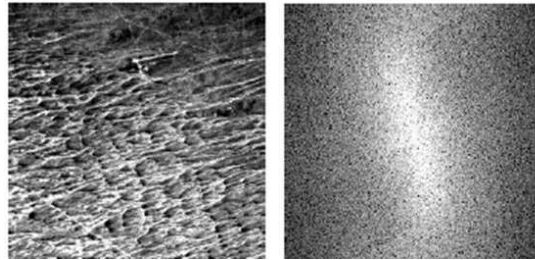


Figura 3.9: Escena dunar y su Transformada

Por último, la tercera imagen representa una zona oceánica en la que se aprecian corrientes en forma helicoidal, lo cual se transforma en el espacio de Fourier, en una serie de pequeñas manchas próximas al origen (bajas frecuencias) y poco definidas linealmente (falta de periodicidad).

Nótese que, en todas ellas, existe una línea vertical de elevada amplitud que pasa por el origen, debida al patrón periódico generado por el efecto de barrido transversal de los sensores, el cual tan solo se aprecia visualmente en la imagen inferior

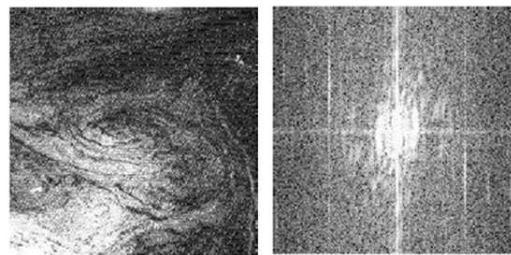


Figura 3.10: Escena oceánica y su Transformada

Un filtro paso-bajo consiste en la eliminación o atenuación de las altas frecuencias de la imagen, para ello, bastará definir una máscara binaria de forma que al aplicarla sobre el espectro de frecuencias mantenga los valores bajos, situados en el centro de la imagen, y elimine los altos.

Esto se consigue mediante círculos binarios de valor uno en el interior y cero en el exterior, cuanto menor sea el radio del círculo, mayor será el efecto del filtro, ya que el intervalo de altas frecuencias que se elimina en la multiplicación ($F(u, v) \times H(u, v)$) será mayor.



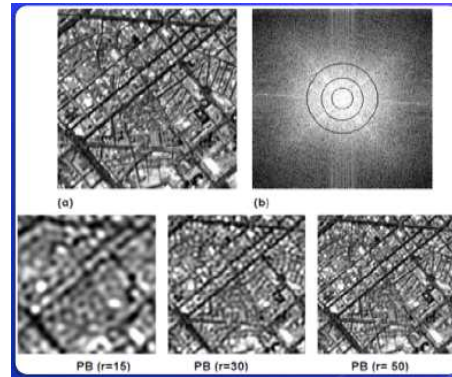


Figura 3.11: Filtro aplicado a la Transformada

Al contrario, cuanto mayor sea el radio del círculo, menor es el efecto del filtro. También puede determinarse la discriminación del filtro, entendida como la suavidad con la que pasa de la zona de corte a la de paso.

Así, si en vez de definirse un círculo binario, el filtro se diseña como una función gaussiana con valores de 0 a 1, la elección de la desviación típica de la misma permite controlar la pendiente del corte de las frecuencias del filtro.

En la siguiente figura se muestran dos ejemplos, uno con una función gaussiana con un valor de $\sigma = 0,2$, y en el otro con $\sigma = 0,8$.

Tal y como se aprecia en los gráficos de los perfiles de los filtros, el primero realiza un filtrado de las altas frecuencias mucho más severo que el segundo.

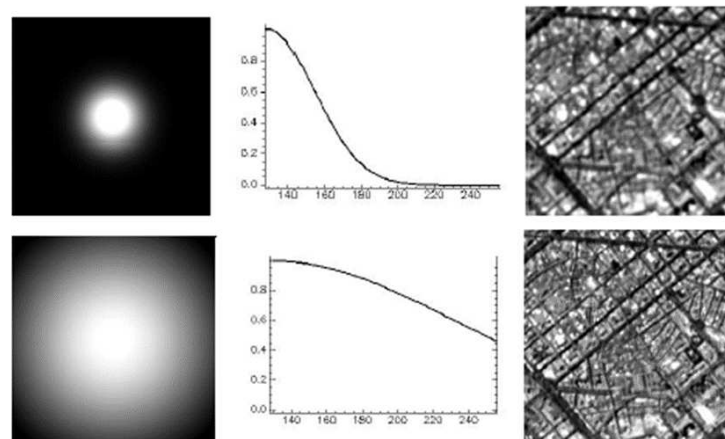


Figura 3.12: Transformadas semiperfiles y Resultados

Análogamente, un filtro paso-alto podrá diseñarse de forma sencilla mediante una máscara binaria circular, tal que dentro del círculo tenga como valor cero, y fuera la unidad.



Cuanto mayor sea el radio del círculo, mayor será el intervalo de bajas frecuencias que se elimina y, por lo tanto, más severo será el efecto del filtrado, al contrario de los que ocurría en el diseño de filtros paso-bajo.

Si el filtro se define mediante una máscara binaria en forma de anillo, con un radio interno y otro externo, el efecto que se consigue es el filtrado de las frecuencias intermedias o extremas. Estos filtros se denominan paso-banda.

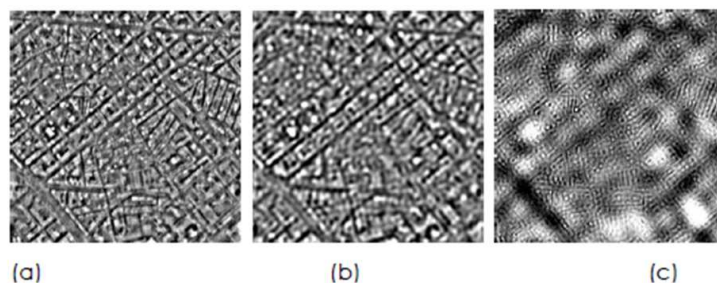


Figura 3.13: Filtros pasa alto (a) y pasa-bandas (b) y (c)

El diseño de filtros específicos en el espacio de Fourier es especialmente relevante cuando se trata de eliminar el ruido periódico en la imagen.

Por ejemplo, supongamos la imagen de la siguiente figura en la que se representa un detalle del mar Mediterráneo.

En ella se aprecia claramente el bandeo horizontal provocado por el sucesivo barrido transversal del sensor TM de Landsat¹.

El espectro de frecuencias refleja este efecto en forma de líneas verticales con diferentes frecuencias.

El diseño de un filtro binario que enmascare o elimine estas líneas, y su multiplicación sobre el espectro, permite obtener la imagen filtrada en la que el patrón de bandeo horizontal o ruido periódico ha desaparecido.

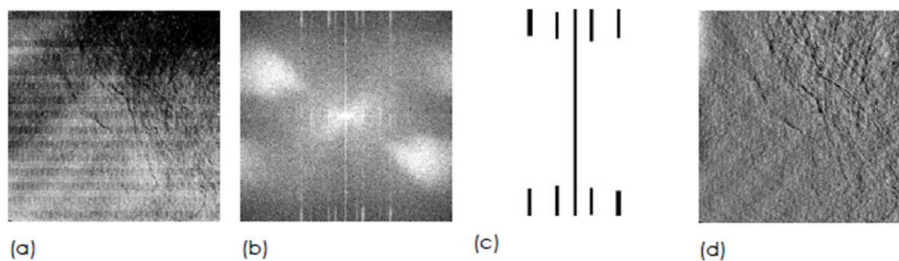


Figura 3.14: a) Original b) espectro Fourier c) Filtro binario d) Resultado

¹TM es un tipo de sensor en la plataforma de satélites LANDSAT que genera tomas de la tierra



Nótese además cómo en el resultado aparecen ligeramente más realzadas las altas frecuencias que en la imagen original, debido a que el filtro elimina una pequeña parte de las bajas frecuencias correspondientes a la zona central (origen) del espectro.



3.3. Filtro Gaussiano

EL filtro gaussiano es un filtro de suavizado lineal que puede eliminar el ruido gaussiano y se usa ampliamente en el proceso de eliminación de ruido en el área de procesamiento de imágenes.

El filtrado gaussiano es el proceso de promedio ponderado de toda la imagen.

El valor de cada píxel se obtiene mediante el promedio ponderado de sí mismo y otros valores de píxeles en la vecindad.

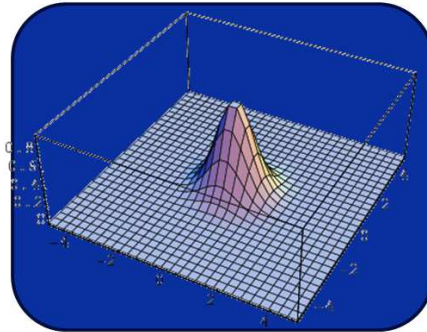


Figura 3.15: Respuesta del filtro Gaussiano

La operación específica del filtrado gaussiano es usar una plantilla (o convolución, máscara) para escanear cada píxel en la imagen y usar la escala de grises promedio ponderada de los píxeles en la vecindad determinada por la plantilla para reemplazar el valor del píxel central de la plantilla.

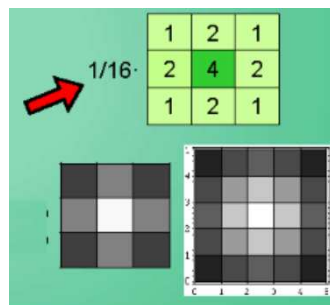


Figura 3.16: Máscara Gaussiana de 3×3

Ruido Gaussiano

Entre los diferentes tipos de ruido en imágenes, el más común es el **ruido gaussiano**, que aparece en el momento de la adquisición de la imagen por un dispositivo en malas condiciones, mala iluminación o altas temperaturas y tiene como consecuencia el *emborronamiento* de todos los píxeles de la imagen.



En el ruido de tipo Gaussiano, la intensidad de todos y cada uno de los píxeles que componen la imagen se ven afectados y cambian su valor, de acuerdo con una distribución normal.

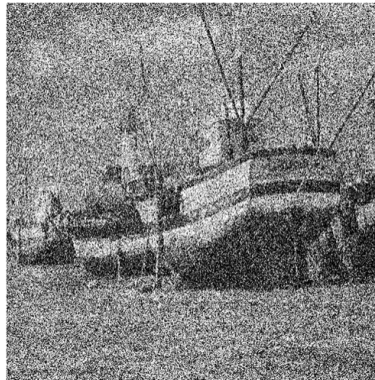


Figura 3.17: Ejemplo ruido Gaussiano

Su función de densidad de probabilidad viene dada en términos de la media μ y la varianza σ^2 de una variable aleatoria x , definida por la siguiente ecuación:

$$p_q(x) = (2\pi\sigma^2)^{-1/2} e^{-(x-\mu)^2/2\sigma^2}$$

El efecto visual de la imagen generada por la tecnología de desenfoque gaussiano es como observar la imagen a través de una pantalla semitransparente, lo cual es obvio por el efecto de imagen desenfocado de la lente y el efecto a la sombra de la iluminación ordinaria diferente.



Figura 3.18: Ejemplo filtro Gaussiano

El **suavizado gaussiano** también se utiliza en la etapa de preprocesamiento de los algoritmos de visión por computadora para mejorar el efecto de imagen de las imágenes a diferentes escalas.





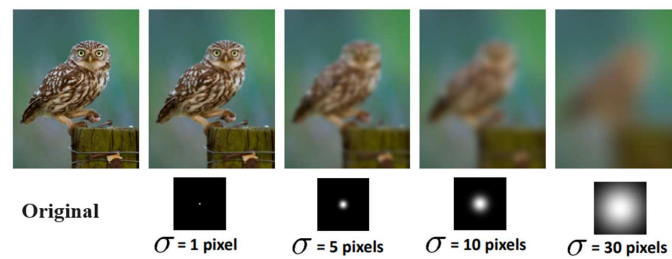
Figura 3.19: Ejemplo suavizado Gaussiano

Desde un punto de vista matemático, el proceso de **desenfoque gaussiano** de una imagen es la convolución de la imagen y la distribución normal.



Figura 3.20: a) Original b) Desenfoque gaussiano

Y la desviación estándar (σ) de la función define la cantidad de desenfoque.

Figura 3.21: Desenfoque debido a varias σ

La función para el filtro gaussiano en matlab es:

`imgaussfilt(I, σ)`

Donde: Valor por default de $\sigma = 0,5$.



3.4. Obtención del dominio de frecuencias a partir de filtros espaciales

UN **filtro digital** es un sistema que, dependiendo de las variaciones de las señales de entrada en el tiempo y amplitud, se realiza un procesamiento matemático sobre dicha señal, obteniéndose en la salida el resultado del procesamiento matemático o la señal de salida.

En el caso de las imágenes, la señal de entrada depende de las coordenadas (x,y) y de la amplitud (o intensidad) de cada píxel.

Se consideran los filtros como operaciones que se aplican a los píxeles de una imagen digital para optimizarla, enfatizar cierta información o conseguir un efecto especial en ella.

El proceso de filtrado puede llevarse a cabo tanto en el dominio de frecuencia como en el espacial.

Aquí se describe la técnica de filtrado en el dominio de la frecuencia, el cuál consta de:

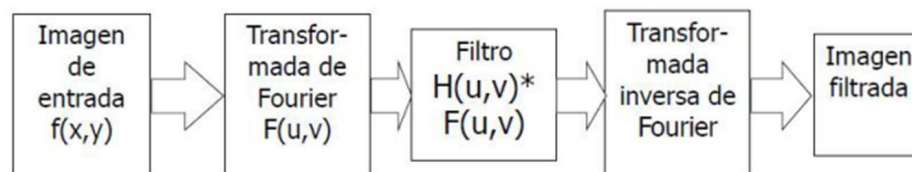


Figura 3.22: Etapas del filtrado

El filtrado en frecuencia se lleva a cabo siguiendo la serie de pasos que se muestran en la fig. 3.22.

Inicialmente, se toma la función $f(x,y)$ de la imagen de entrada (imagen a modificar) y se le aplica la DFT, obteniendo una función en el dominio de la frecuencia $F(u,v)$.

Luego, se realiza la convolución entre F y una función filtro $H(u,v)$ acorde al resultado que se espera del filtrado.

A partir de (3), se puede concluir que la antitransformada de Fourier de esta convolución, resulta ser el producto de las antitransformadas de Fourier de $F(u,v)$ y $H(u,v)$.

Por lo que finalmente, se aplica antitransformada de Fourier a la convolución para obtener la imagen filtrada.

Tipos de filtro

Existen básicamente tres tipos distintos de filtros que pueden aplicarse:

- **Filtro pasa bajo:** atenúa las frecuencias altas y mantiene sin variaciones las bajas. El resultado en el dominio espacial es equivalente al de un filtro de suavizado, donde las altas frecuencias que son filtradas se corresponden con los cambios fuertes



de intensidad. Consigue reducir el ruido suavizando las transiciones existentes. La Figura 2 muestra el gráfico de un filtro pasa bajo ideal y su ecuación.

- **Filtro pasa alto:** atenúa las frecuencias bajas manteniendo invariables las frecuencias altas. Puesto que las altas frecuencias corresponden en las imágenes a cambios bruscos de densidad, este tipo de filtros es usado, porque entre otras ventajas, ofrece mejoras en la detección de bordes en el dominio espacial, ya que estos contienen gran cantidad de dichas frecuencias. Refuerza los contrastes que se encuentran en la imagen. En la figura 3 se presenta la función de un filtro pasa alto y su ecuación.
- **Filtro pasa banda:** atenúa frecuencias muy altas o muy bajas manteniendo una banda de rango medio.

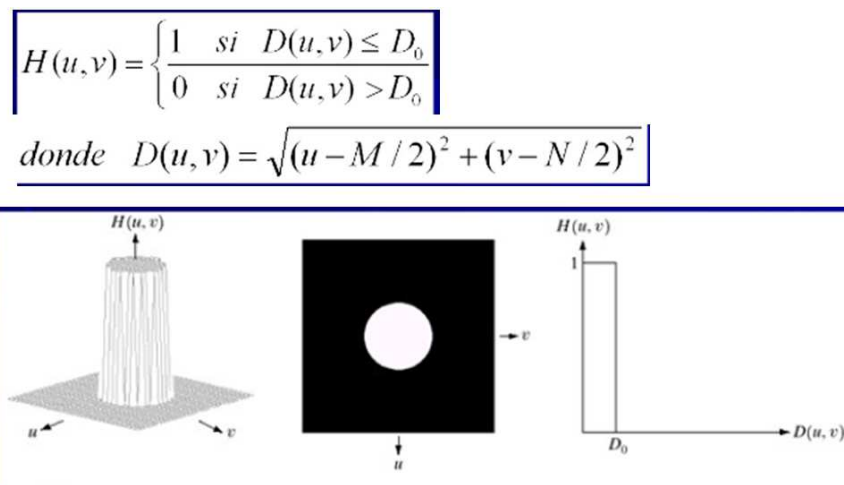


Figura 3.23: Filtro pasa bajos ideal y su ecuación

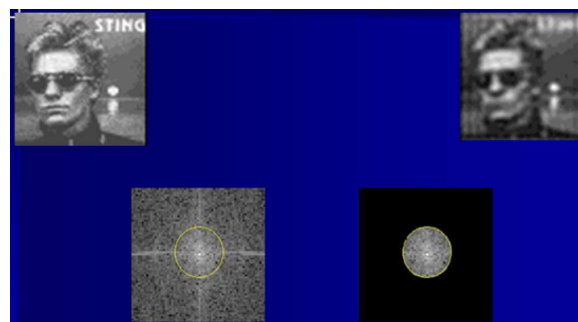


Figura 3.24: Ejemplo filtro pasa bajos



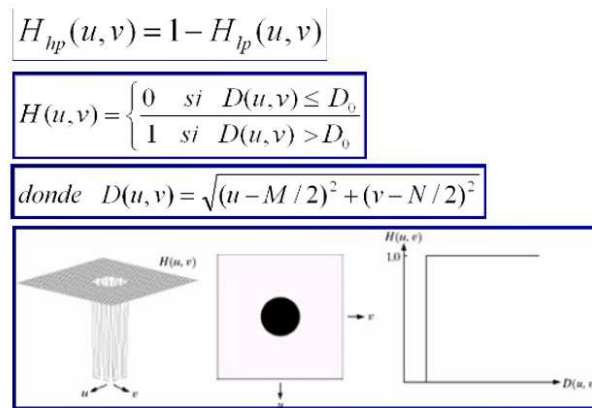


Figura 3.25: Filtro pasa altos ideal y su ecuación

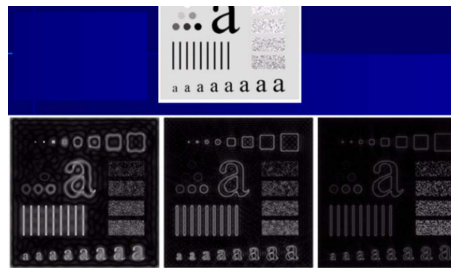


Figura 3.26: Ejemplo filtro pasa altos

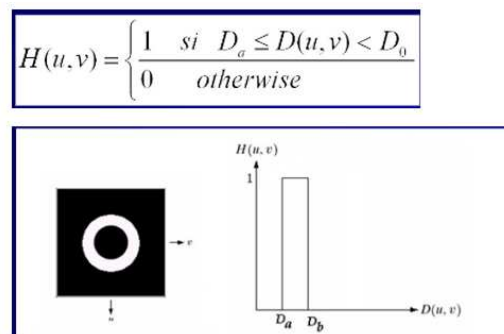


Figura 3.27: Filtro pasa banda ideal y su ecuación



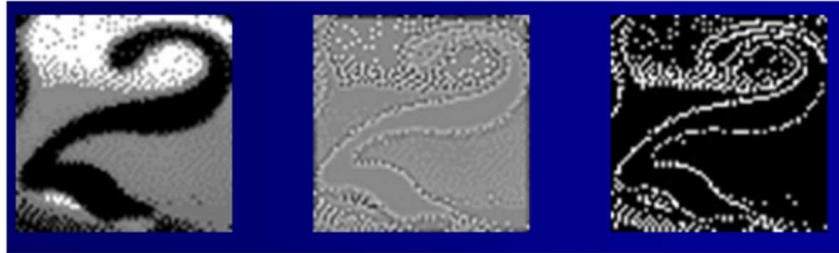


Figura 3.28: Ejemplo filtro pasa banda



3.5. Convolución

LA **convolución** es la operación de modificar los píxeles de una imagen basándose en alguna función sobre el vecindario local de cada píxel

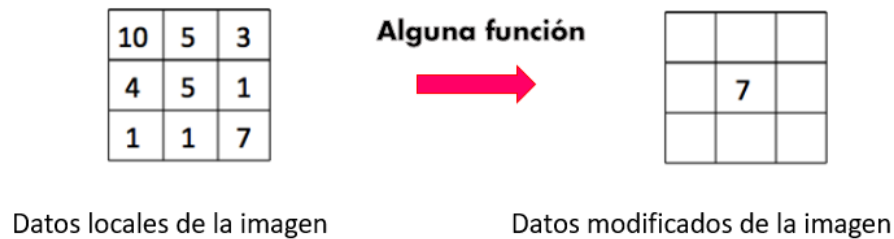


Figura 3.29: Convolución

En términos matemáticos puros, una convolución representa la mezcla de dos funciones, $f(x)$ y $g(x)$, mientras una se desliza sobre la otra.

Por cada pequeño desplazamiento deslizante (dx), los puntos correspondientes de la primera función $f(x)$ y la imagen especular de la segunda función $g(t-x)$ se multiplican juntos y luego se suman.

El resultado es la convolución de las dos funciones, representada por la expresión $[f * g](t)$.

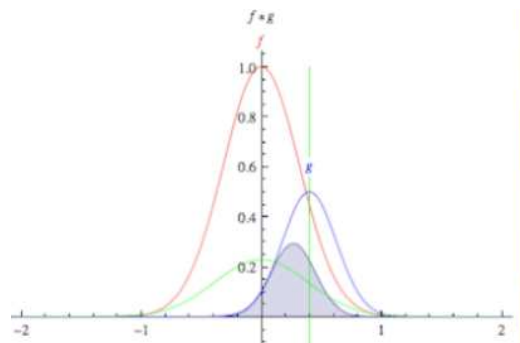


Figura 3.30: Correlación cruzada (*Convolución*) de dos funciones, f y g

Definición general de la convolución

$$f(x) * g(x) = \int_{-\infty}^{\infty} f(u)g(x - u)du$$

Convolución para imágenes digitales

$$F(x, y) = f(x, y)g(x, y) = \sum_i \sum_j f(i, j)g(x = 1, y = j)$$



Donde g es la máscara de convolución espacial.

La transformada de Fourier de la convolución de dos funciones es el producto de sus transformadas de Fourier.

$$F[g * h] = F[g]F[h]$$

La transformada inversa de Fourier del producto de dos transformadas de Fourier es la convolución de las dos transformadas inversas de Fourier.

$$F^{-1}[gh] = F^{-1}[g] * F^{-1}[h]$$

La convolución en el dominio espacial es equivalente a multiplicación en el dominio de la frecuencia.

En el procesamiento de imágenes, la convolución se realiza deslizando una pequeña matriz de números, típicamente una matriz de tamaño $[3 \times 3]$ o $[5 \times 5]$, secuencialmente sobre diferentes partes de la imagen.

Esta **Matriz de convolución** también se conoce como Filtro de convolución o núcleo.

Para cada posición de la matriz de convolución, los valores de píxeles correspondientes se multiplican y se suman para reemplazar el píxel central original.

De esta manera, los valores de los píxeles vecinos se mezclan con los del píxel central para crear un Matriz de características convolucionada.

Los valores relativos de los elementos de la matriz determinan cómo afectará esa fusión a la imagen transformada.

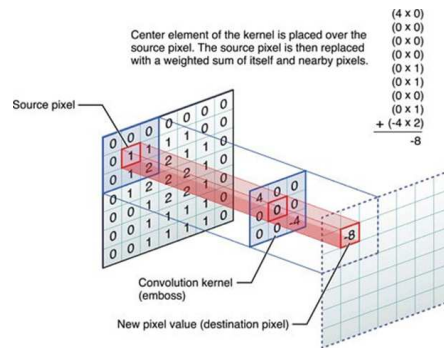


Figura 3.31: Proceso de Convolución

Tamaño del kernel

Esto se refiere al tamaño de la matriz deslizante que pasa sobre la imagen. Por lo general, el tamaño es $[3 \times 3]$, $[5 \times 5]$ o $[7 \times 7]$.

Los tamaños de kernel más pequeños proporcionan una mejor resolución de las características más pequeñas de la imagen y generalmente se prefieren para el procesamiento de resonancia magnética / radiología.



El compromiso es una mayor carga computacional de las muchas circunvoluciones más pequeñas producidas en cada paso.

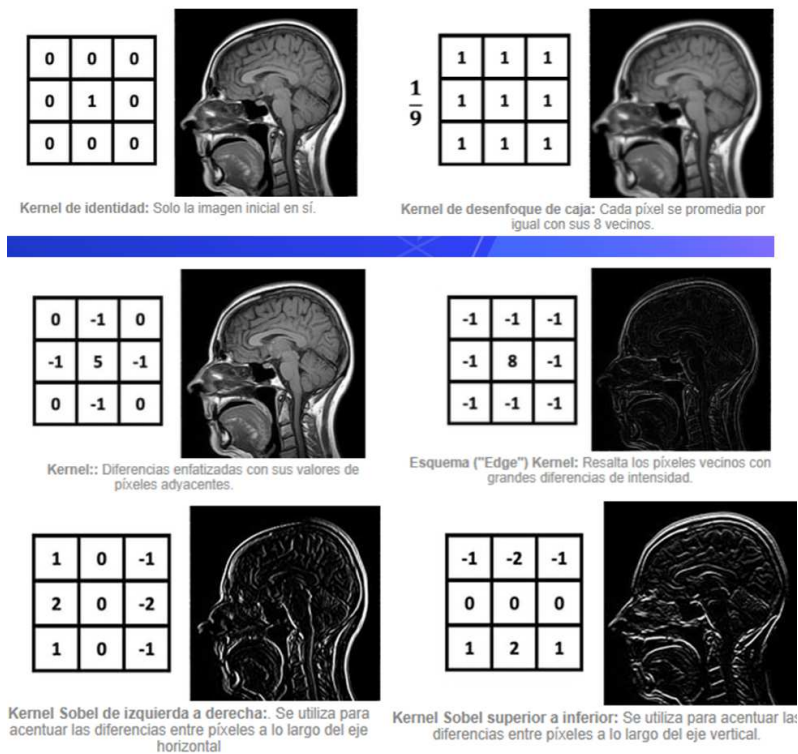


Figura 3.32: Kerneles varios

En el procesamiento de imágenes, la convolución requiere 3 componentes:

1. Una imagen de entrada.
2. Una matriz de kernel que vamos a aplicar a la imagen de entrada.
3. Una imagen de salida para almacenar la salida de la imagen de entrada convolucionada con el kernel.

¿Cómo se realiza la convolución

La convolución es simplemente la suma de la multiplicación de matrices de elementos entre el núcleo y la vecindad que cubre el núcleo de la imagen de entrada.

1. Seleccionar una coordenada (x, y) de la imagen original.
2. Colocar el centro del núcleo en esta coordenada (x, y).



3. Multiplicar cada valor del kernel por el valor de píxel de imagen de entrada correspondiente y, a continuación, tomar la suma de todas las operaciones de multiplicación. (En pocas palabras, estamos tomando la multiplicación de elementos de la región de la imagen de entrada y el núcleo, luego sumando los valores de todas estas multiplicaciones en un solo valor. La suma de estas multiplicaciones se denomina salida del núcleo).
4. Utilizar la misma coordenada (x, y) del paso 1, pero esta vez almacenar la salida del kernel en la misma ubicación (x, y) que la imagen de salida.

Características de la convolución

- La convolución es una operación lineal
- La convolución es la operación más general que se puede aplicar a una imagen
- A veces también nos referimos a la convolución como filtrado
- Filtrado espacial: Directamente aplicable mediante una máscara
- Las máscaras de convolución, o filtros, no suelen ser grandes (3x3, 5x5) para ahorrar tiempo de cómputo

Propiedades de la convolución

Sean a, b y c, y d imágenes continuas o discretas

- La convolución es conmutativa

$$c = a \otimes b = b \otimes a$$
- La convolución es asociativa

$$c = a \otimes (b \otimes c) = (a \otimes b) \otimes c = a \otimes b \otimes c$$
- La convolución es distributiva

La convolución provee una manera de multiplicar dos arreglos de números, generalmente de tamaños diferentes, pero de la misma dimensionalidad, para producir un tercer arreglo de números de la misma dimensionalidad.

Esto puede ser utilizado para implementar operadores cuyos valores de píxeles de salida sean simples combinaciones lineales de ciertos valores de píxeles de entrada

Uno de los arreglos de entrada es usualmente una imagen en niveles de gris en 2D.

El segundo arreglo es usualmente mucho más pequeño y también es 2D (aunque puede ser de 1x1, i.e. un píxel)



I ₁₁	I ₁₂	I ₁₃	I ₁₄	I ₁₅	I ₁₆	I ₁₇	I ₁₈	I ₁₉
I ₂₁	I ₂₂	I ₂₃	I ₂₄	I ₂₅	I ₂₆	I ₂₇	I ₂₈	I ₂₉
I ₃₁	I ₃₂	I ₃₃	I ₃₄	I ₃₅	I ₃₆	I ₃₇	I ₃₈	I ₃₉
I ₄₁	I ₄₂	I ₄₃	I ₄₄	I ₄₅	I ₄₆	I ₄₇	I ₄₈	I ₄₉
I ₅₁	I ₅₂	I ₅₃	I ₅₄	I ₅₅	I ₅₆	I ₅₇	I ₅₈	I ₅₉
I ₆₁	I ₆₂	I ₆₃	I ₆₄	I ₆₅	I ₆₆	I ₆₇	I ₆₈	I ₆₉

K ₁₁	K ₁₂	K ₁₃
K ₂₁	K ₂₂	K ₂₃

Figura 3.33: Elementos de convolución

Una **máscara** es una matriz (usualmente) más pequeña, máscaras de diferentes tamaños conteniendo diferentes patrones de números dan resultados diferente.

1	1	1
1	1	1
1	1	1

Set of coordinate points =

{ (-1, -1), (0, -1), (1, -1),
 (-1, 0), (0, 0), (1, 0),
 (-1, 1), (0, 1), (1, 1) }

Figura 3.34: Máscara

La implementación se hace deslizando la máscara sobre la imagen, generalmente comenzando de la esquina superior izquierda, y moviéndolo a todas las posiciones donde quede completamente dentro de los límites de la imagen.

Si la imagen es de $M \times N$ y la máscara de $m \times n$, el tamaño de la imagen de salida será $(M - m + 1) \times (N - n + 1)$.

Cada posición de la máscara corresponde a un solo pixel de salida, cuyo valor se calcula como sigue:

$$O(i, j) = \sum_{k=1}^m \sum_{l=1}^n I(i + k - 1, j + l - 1) K(k, l)$$

donde:

$$i = 1 \dots M - m + 1, j = 1 \dots N - n + 1$$



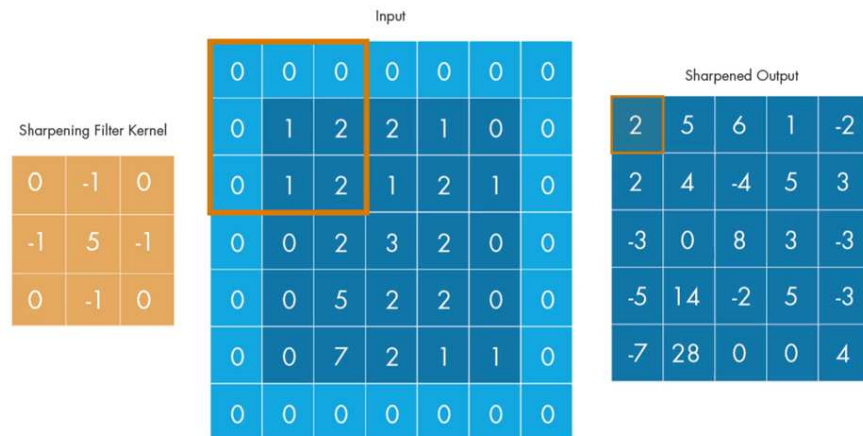


Figura 3.35: Inicio de una convolución

Un ejemplo de uso de la convolución es para hacer una realce de bordes de una imagen aplicando un Filtro de Sobel.

-1	-2	-1
0	0	0
1	2	1

Figura 3.36: Kernel de Sobel

La convolución y la deconvolución de la imagen se pueden realizar rápidamente en el dominio de la frecuencia.

En el dominio de la frecuencia, la multiplicación puntual es la convolución, mientras que la división puntual es la deconvolución.

Teorema de la convolución

El teorema de convolución se puede enunciar desde el dominio de la frecuencia de la siguiente manera: *La transformada de Fourier de la convolución de dos funciones es igual al producto de sus transformadas de Fourier individuales.*

Esto significa que si se tienen dos funciones, $f(x)$ y $g(x)$, y las convolucionamos, podemos encontrar sus transformadas de Fourier individualmente, multiplicarlas juntas, y luego tomar la transformada de Fourier inversa para obtener la convolución de $f(x)$ y $g(x)$.



Procedimiento

1. Convertir la imagen de entrada en el dominio de la frecuencia, a través de la transformada rápida de Fourier.
2. Rellenar el kernel de convolución con ceros, hasta que alcance el mismo tamaño que la imagen de entrada.
3. Convertir el núcleo de convolución en el dominio de la frecuencia, a través de la transformada rápida de Fourier.
4. Multiplicar puntualmente la imagen de entrada del dominio de frecuencia y el núcleo de convolución del dominio de frecuencia juntos.
5. Convertir el producto de la multiplicación puntual de nuevo en el dominio espacial a través de la transformada de Fourier rápida inversa.



Bibliografía

[Diccionario Larousse, 2003] “*Diccionario Larousse, Edición Premium*”, EDICIONES LAROUSSE MÉXICO y SPS EDITORIAL BARCELONA, (2003).

[Cambridge Dictionary, 2022] “*Cambridge Dictionary on-line*”, <https://dictionary.cambridge.org/>, (2022).

[DRALE] “*Diccionario de la Real Academia de la Lengua Española*”, versión on-line (2017).