

UNIVERSIDAD NACIONAL DEL LITORAL
FACULTAD DE INGENIERÍA Y CIENCIAS HÍDRICAS
PROCESAMIENTO DIGITAL DE IMÁGENES

Resumen de Teoría - Parcial I

Autores:

Mario ROSALES
Franco SANTELLÁN
Carlos GENTILE

E-mail:

mariorosales941@gmail.com
fransantellan@gmail.com
csgentile@gmail.com

MAYO, 2017



Índice

1. UNIDAD 1: Introducción	2
1. La visión en las máquinas	2
2. Etapas de un sistema de visión	2
3. Niveles de procesamiento	2
4. Concepto de imagen digital	3
5. Espectro electromagnético	3
6. Física del sensado de imagen	3
7. Física de la imagen	4
8. Muestreo y cuantización de imágenes	5
9. Formatos de las imágenes digitales	6
10. El mecanismo de la visión humana	6
11. Formación de imágenes en el ojo	7
12. Adaptación al brillo y discriminación	7
13. Fenómenos de percepción humana	8
14. Resolución y cuantización en la visión humana	9
2. UNIDAD 2: Operaciones en el Dominio Espacial	10
1. Operaciones Puntuales: Independientes de las características globales - Operaciones en una sola imagen	10
1.1. Definiciones Previas	10
1.2. Operaciones sobre imágenes	10
2. Operaciones Puntuales: Independientes de las características globales - Operaciones entre varias imágenes	14
2.1. Operaciones sobre varias imágenes	14
3. Operaciones Puntuales: Dependiente de las características globales - Manejo de Histogramas	17
3.1. Propiedades estadísticas del Histograma	17
3.2. Manipulación del Histogramas	19
4. Operaciones Locales - Conceptos	20
5. Filtrado Espacial - Conceptos	20
5.1. Filtros Lineales de Suavizado	22
5.2. Filtros Lineales de Acentuado	24
5.3. Filtros No Lineales	26
5.4. Combinación de Métodos	26
6. Procesamiento de COLOR	27
6.1. Fundamentos del Color	27
6.2. Modelos de Color	27
6.3. PDI en Pseudo-Color	29
6.4. Procesamiento a todo color	30
3. UNIDAD 3 - Operaciones en el Dominio Frecuencial	32
1. Introducción	32
1.1. Propiedades de la TFD	33
1.2. Representación Gráfica de la TF	34
1.3. Importancia de la Fase	35
2. Filtrado Frecuencial	37
2.1. Filtros de Suavizado	37
2.2. Filtros de Acentuado	39
2.3. Filtrado de Mascara Difusa	40
2.4. Filtrado de Alta Potencia (high-boost)	41
2.5. Filtrado de Énfasis de Alta Frecuencia	41
2.6. Filtrado Homomórfico	42

1. UNIDAD 1: Introducción

1. La visión en las máquinas

Aunque los seres humanos estemos limitados a la banda visible del espectro electromagnético (EM), las máquinas pueden percibir casi el espectro completo, desde los rayos gamma, a las ondas de radio. Las máquinas también pueden procesar imágenes generadas por fuentes que los humanos no asociamos con imágenes; como es el caso del ultrasonido, la microscopía de electrones, etc.

La visión en los humanos permite percibir e interpretar el mundo que los rodea. La visión en las máquinas intenta imitar la visión humana a través de dispositivos electrónicos (cámaras, sensores, etc.) que capturen las imágenes y procesadores que analicen e interpreten las mismas.

2. Etapas de un sistema de visión

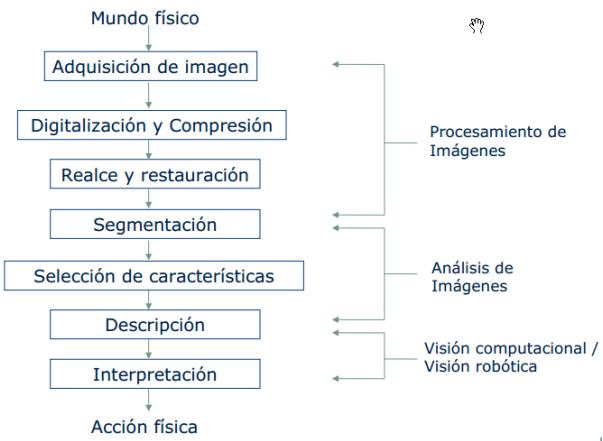
Mundo físico

1. **Adquisición de imagen:** Captura de datos visuales mediante sensores.
2. **Digitalización y Compresión:** Conversión de datos a forma discreta - Almacenamiento y transmisión eficiente.
3. **Realce y restauración:** Mejora de calidad de imagen.
4. **Segmentación:** Partición en objetos o partes constituyentes.
5. **Selección de características:** Extracción de descriptores útiles para diferenciar clases u objetos.
6. **Descripción:** Asignación de etiquetas a los objetos.
7. **Interpretación:** Asignación de significado a los objetos reconocidos.

Acción física

3. Niveles de procesamiento

Los Niveles de Procesamientos hacen referencia al comienzo y fin de los ámbitos de estudio de cada área. En PDI tenemos que:



Procesos de Bajo Nivel: [Imagen → Imagen] Utilizan operaciones como el preprocesamiento de imagen para reducir el ruido, mejora del contraste, y filtros de enfoque. Se caracterizan por que sus entradas son imágenes y sus salidas también.

Procesos de Nivel Medio: [Imagen → Atributos] Operaciones como segmentación y clasificación de objetos individuales. Se caracterizan por que sus entradas son generalmente imágenes, pero sus salidas son atributos extraídos de esas imágenes (contornos, bordes, identidad de objetos individuales).

Procesos de Alto Nivel: [Atributos → Decisión] Implica el obtener algún significado de un conjunto de objetos reconocidos – análisis de imágenes – y, finalmente, realizar las funciones cognitivas asociadas con la vista. (ejemplo de símbolos de tráfico)

4. Concepto de imagen digital

¿Qué entendemos por imagen?

Una imagen puede ser definida matemáticamente como una función bidimensional, $f(x,y)$, donde x e y son coordenadas espaciales (en un plano), y $f(x,y)$ es la intensidad o nivel de gris de la imagen en esa coordenada.

¿Qué entendemos por imagen digital?

Cuando x , y y $f(x,y)$ son todas cantidades finitas y discretas, decimos que la imagen es una imagen digital. Una imagen digital se compone de un número finito de elementos, cada uno con un lugar y valor específicos. Estos elementos son llamados pixels.

En cuanto a la dimensión del valor en el punto (x,y) podemos decir que:

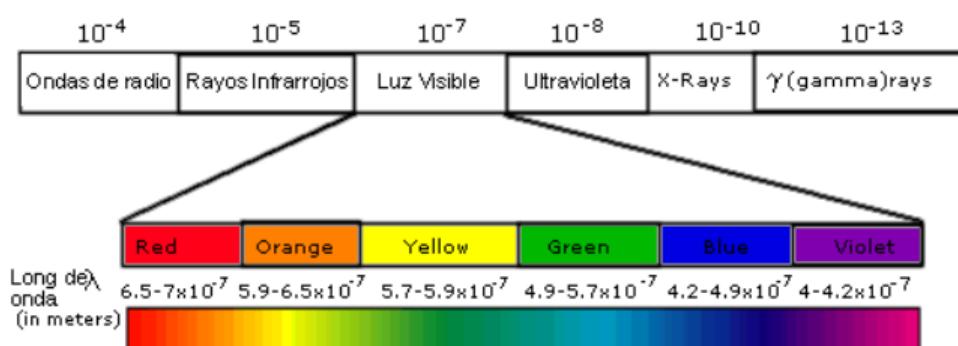
- Si $f \in \mathbb{R}^1$: imagen en intensidades de gris
- Si $f \in \mathbb{R}^3$: imagen en color RGB, CMY, HSV, ...
- Si $f \in \mathbb{R}^n$: imagen multiespectral (satelital, RMN, ...)

5. Espectro electromagnético

La principal fuente de energía de las imágenes es el espectro electromagnético (otras fuentes son la acústica, la ultrasónica, y la electrónica). Las imágenes basadas en la radiación del espectro EM son las más familiares, especialmente las imágenes de rayos X y las bandas visuales del espectro.

Las ondas electromagnéticas pueden ser conceptualizadas como ondas sinusoidales de longitudes de onda variadas, o como una corriente de partículas sin masas viajando en un patrón de ondas y moviéndose a la velocidad de la luz. Su unidad es el fotón. Agrupando las bandas espectrales de acuerdo a su energía por fotón, obtenemos el espectro de la figura, desde los rayos gamma (con mayor energía) hasta las ondas de radio.

Relación entre λ (longitud de onda) y ν (frecuencia): $\lambda = c/\nu$, con c : velocidad de la luz ($3 \cdot 10^8 m/s$)



6. Física del sensado de imagen

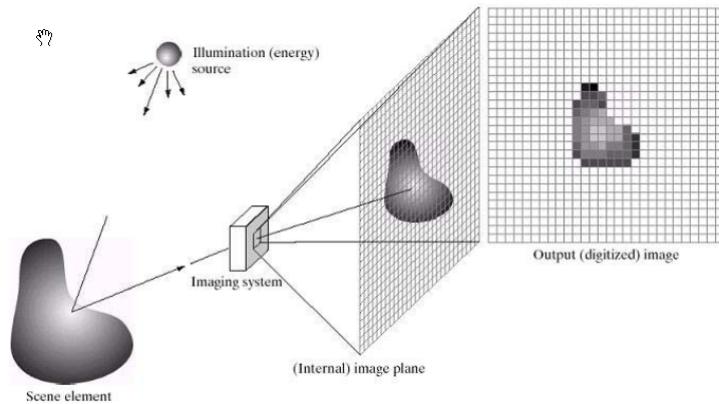
Los tipos de imágenes que nos interesan se generan por una combinación de una fuente de *iluminación* y la reflexión o absorción de energía de esta fuente por parte de los elementos de la escena. Esta *iluminación* puede ser también radar, infrarrojo, o alguna fuente que no es parte del EEM, como ultrasonido.

Dependiendo de la naturaleza de la fuente, la energía "luminosa" puede ser reflejada o transmitida por los objetos.

En algunas aplicaciones, la energía transmitida o reflejada se enfoca a un fotoconvertidor (p. Ej. Una pantalla de fósforo) que convierte la energía a luz visible.

La energía de la fuente de iluminación es reflejada o transmitida por un elemento de la escena, la energía obtenida se colecta y se enfoca al plano focal (en el caso de luz, por medio de una lente).

Dichos sensores pueden ser puntuales, lineales o arreglos. En el caso del ultimo proveen la imagen 2D a través del enfoque del patrón de energía directamente sobre los sensores de la matriz.



Elección de la luz: Interesa elegir la fuente de luz apropiada al ambiente de trabajo que permita obtener la mejor imagen (máximo número de intensidades y mayor contraste con el fondo), que minimice la cantidad de procesamiento requerido para restauración y que provea una iluminación homogénea y temporalmente constante sobre el área de interés. Entre las fuentes de luz se encuentran:

- Luz de día (depende del clima)
- Lámparas halógenas (baratas, iluminación no uniforme y genera calor)
- Lámparas fluorescentes (iluminación homogéneo, sin calor excesivo y con limitación espectral)
- Lámparas de descarga (flash)
- LEDs (Monocromáticos, no liberan calor, no poseen modulaciones de intensidad, bajo voltaje, gran vida útil)
- Láser (radiación coherente, usados para alimentar a fibras ópticas)

7. Física de la imagen

Recordemos que una imagen puede ser representada por una función bidimensional de la forma $f(x, y)$, en donde el valor o amplitud de f en las coordenadas espaciales (x, y) da la intensidad de la imagen en ese punto. Cuando la imagen se genera a partir de un proceso físico, sus valores son proporcionales a la energía radiada por la fuente física (p. Ej. Ondas EM). Como consecuencia, $f(x, y)$ debe ser diferente de cero y finita.

La función $f(x, y)$ se caracteriza por dos componentes:

Iluminación: la cantidad de luz incidente procedente de la fuente sobre la escena.

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

Reflectancia: la cantidad de luz reflejada por los objetos de la escena.

$$0 < r(x, y) < 1 \quad \begin{cases} 0 & \text{Absorción total} \\ 1 & \text{Reflexión total} \end{cases}$$

El producto de estas dos componentes forman la función imagen:

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

teniendo en cuenta que: $0 < i(x, y) < \infty$ y $0 < r(x, y) < 1$. Estas restricciones dejan ver que la naturaleza de $i(x, y)$ está determinada por la fuente de iluminación, y la de $r(x, y)$, por las características de los objetos

8. Muestreo y cuantización de imágenes

Para crear una imagen digital, necesitamos convertir los datos detectados (sensados) a forma digital, esto involucra los procesos de muestreo y cuantificación.

Una imagen puede ser continua tanto respecto a sus coordenadas x e y , como a su amplitud. Para convertirla a forma digital, hay que digitalizarla en los dos aspectos (espacialmente y en amplitud). La digitalización de las coordenadas espaciales (x, y) se denomina **muestreo** de la imagen y la digitalización de su amplitud se conoce como **cuantificación**.

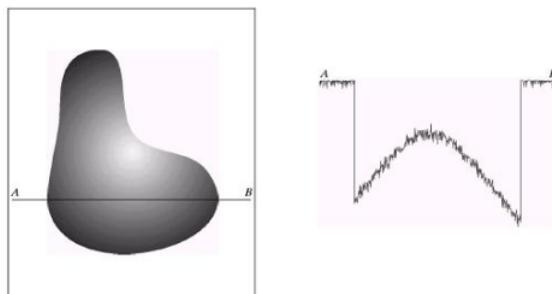


Figura 1: Original y Amplitud en AB

La función bidimensional mostrada en la segunda figura es una gráfica de los valores de amplitud (el nivel de gris) de la imagen en la primera figura a lo largo del segmento de línea AB. Las variaciones aleatorias se deben a ruido de la imagen.

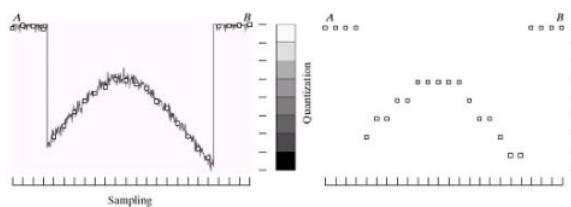


Figura 2: Muestreo y Cuantización

Para muestrear esta función, tomamos muestras a espacios iguales a lo largo de AB, indicadas por los cuadritos blancos. El conjunto de estos cuadritos nos da la función muestreada.

Sin embargo los valores de las muestras aún se encuentran en un rango continuo de valores de niveles de gris. Para obtener una función digital, debemos convertir (cuantificar) los valores de gris a cantidades discretas. Esto se hace simplemente asignando uno de los ocho valores de la figura a cada muestra. La última figura muestra las muestras digitales que resultan del muestreo y cuantificación.

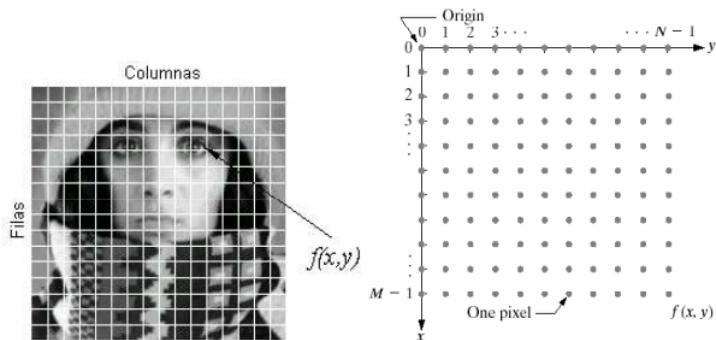
Claramente, la calidad de una imagen digital se determina en gran medida por el número de muestras y niveles de gris utilizados en el muestreo y cuantificación.

Digitalización con un arreglo de sensores: Una imagen continua $f(x,y)$ se describe de forma aproximada por una serie de muestras igualmente espaciadas organizadas en forma de una matriz $N \times M$, donde cada elemento de la matriz es una cantidad discreta. No se requiere un valor especial de M y N , salvo que sean enteros positivos.

En el caso de escalas de grises, el brillo se representa con L niveles de grises, siendo estos una potencia entera de 2 ($L = 2^k$). El número b de bits necesarios para almacenar una imagen digitalizada es $N \times M \times k$ y N^2k cuando $M = N$.

En este ámbito se define *Rango Dinámico* como la dispersión de valores de grises en el intervalo $[0; (L - 1)]$ sobre el rango completo.

Notación: Un elemento de la matriz se identifica mediante una coordenada discreta (x, y) con $0 < x < M - 1$ y $0 < y < N - 1$.



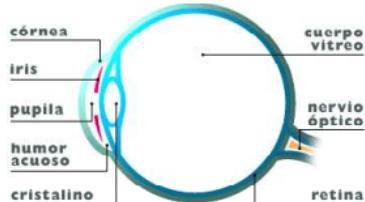
9. Formatos de las imágenes digitales

- **Mapa de bits:** Bitmap (BMP), Tagged Image File Format (TIFF), Portable Network Graphics (PNG), Joint Photographic Experts Group (JPG).
- **Vectorial:** No almacenan el valor de cada punto de la imagen, sino que realizan una representación mediante vectores y formas. Fácil escalado de imágenes. No hay pérdida de información al aumentar o reducir el tamaño. Es útil para dibujos, y no para fotografías (archivos de gran tamaño). Ejemplos: Portable Document Format (PDF), Postscript (PS).
- **Metaformato:** Formatos híbridos de bitmaps y vectoriales que tratan de aprovechar las ventajas de ambos formatos. En una misma escena se usa bitmap para las imágenes y vectorial para las líneas, texto o dibujos. No son utilizados en PDI porque son dependientes del programa que los soporta. Ejemplos: Windows Meta File (WMF), Word Perfect Graphics (WPG), entre otros.

10. El mecanismo de la visión humana

El campo del procesamiento digital de imágenes está construido sobre bases matemáticas y probabilísticas, pero la intuición y análisis humanos juegan un importante papel al momento de escoger una técnica u otra. Esta elección se basa usualmente en juicios visuales subjetivos. Dentro de los componentes principales del ojo tenemos:

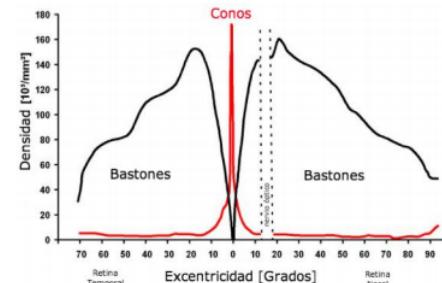
- **Córnea:** membrana transparente en contacto con el exterior.
- **Pupila:** diafragma en el centro del iris que regula la cantidad de luz que ingresa.
- **Cristalino:** lente óptico que efectúa la convergencia de rayos luminosos sobre la retina.
- **Retina:** cuando el ojo está enfocado, la imagen de un objeto exterior es proyectada en la retina. Es quien realiza la transducción de una señal electromagnética (imagen de luz visible) a una señal electroquímica. Sobre su superficie se encuentran los receptores que permiten la percepción visual: los conos y los bastones. Los primeros (6.5 millones) se encargan de la detección de color y detalles de la imagen. Requieren niveles elevados de iluminación (visión fotópica) y poseen alta resolución. Los segundos (130 millones) se encargan de la detección de formas y objetos en movimiento. Funcionan con niveles bajos de iluminación (visión escotópica) y poseen baja resolución.
- **Nervio óptico:** vía neurotransmisora. (neurotransmisión)



(a) Estructura del Ojo



(b) Conos y Bastones

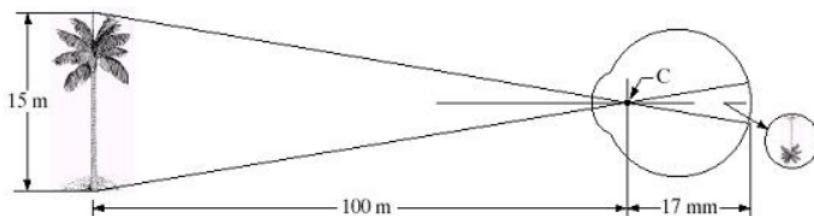


(c) Distribución de Conos y Bastones

11. Formación de imágenes en el ojo

La principal diferencia entre una lente óptica ordinaria y el cristalino es que este último es flexible, y su forma es controlada por la tensión de las fibras del cuerpo ciliar. Para enfocar objetos lejanos, se aplana, para enfocar objetos cercanos, se ensancha. La distancia entre el centro del cristalino y la retina (que llamaremos distancia focal), varía de aproximadamente 17mm a 14mm.

Con esta información podemos calcular el tamaño del objeto reflejado en la retina. Por ejemplo, el observador mira una palmera de 15m de altura desde una distancia de 100m (al punto focal C). Si h es la altura en mm del objeto en la imagen retinal, por geometría obtenemos que: $15/100 = h/17$, entonces $h = 2,55\text{mm}$ (fóvea).

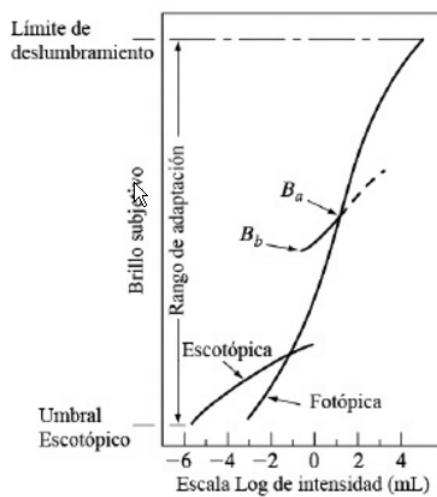


Otro dato a tener en cuenta es que el cambio de curvatura permite enfocar objetos distantes y cercanos. A mayor distancia, menor curvatura y menor refracción (objeto claro). En objetos cercanos, la curvatura aumenta para llevar la imagen que se forma detrás de la retina hacia adelante.

12. Adaptación al brillo y discriminación

Recordemos que las imágenes digitales son un conjunto discreto de intensidades. Por lo tanto tiene sentido hablar de la capacidad de discriminación del ojo entre diferentes niveles de intensidad, es decir, que debemos conocer y considerar la habilidad del ojo para discriminar intensidades de grises diferentes. El rango de intensidades de luz al que el ojo humano puede adaptarse del rango de 10^{10} .

El sistema visual no puede operar sobre este rango simultáneamente, sino que trabaja modificando su sensibilidad: fenómeno de **adaptación al brillo**. Así, el rango simultáneo de discriminación es mucho menor que el total. Evidencia experimental ha demostrado que la iluminación subjetiva o brillo subjetivo (la intensidad luminosa como la percibe el sistema visual humano) es una función logarítmica de la intensidad de luz que incide en el ojo.



En la figura se ilustra lo último mencionado. La línea larga sólida representa el rango de intensidades al que se puede adaptar el sistema visual. Sin embargo se debe tener en cuenta que el rango total de niveles de intensidad distintos que puede discriminar simultáneamente es relativamente pequeño cuando se le compara con el rango total de adaptación que hemos visto.

El nivel de adaptación para cierto conjunto de condiciones es llamado Nivel de adaptación a la iluminación (por ejemplo, Ba). Este nivel está bastante restringido, pues debajo del nivel Bb todo estímulo es percibido como negros indistinguibles unos de otros. El nivel superior, mostrado con guiones, no está restringido pero, si se extiende mucho pierde el significado, ya que intensidades mayores simplemente elevarían el nivel de adaptación Ba.

Otra habilidad del ojo que interesa conocer es la **capacidad para discriminar variaciones en luminosidad**. Un experimento clásico consiste en poner a un observador a mirar una gran área iluminada uniformemente que ocupa su campo visual entero. Esta área es típicamente iluminada por la parte de atrás por una fuente cuya intensidad I , puede variar. Se incrementa la iluminación, ΔI , con destellos de corta duración con forma de círculo en el centro del campo iluminado. Si ΔI no es suficientemente brillante, el observador dirá "no". Mientras ΔI crece, si el observador lo percibe dirá "sí", cuando lo perciba. Finalmente, cuando ΔI sea lo suficientemente fuerte, el observador dirá "sí" todo el tiempo.

La cantidad $\Delta I_c/I$, donde ΔI_c es el incremento de iluminación que puede discernirse el 50 % de las veces con iluminación I , se llama cociente de Weber. Un valor pequeño de este cociente indica que un cambio de porcentaje pequeño en intensidad puede distinguirse, es decir, buena discriminación de iluminación. Un valor grande quiere decir que se necesita un gran cambio en intensidad para poder distinguirlo. Es decir, mala discriminación.

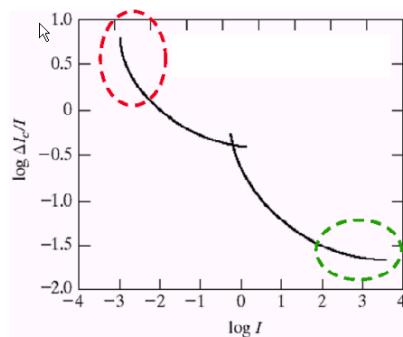


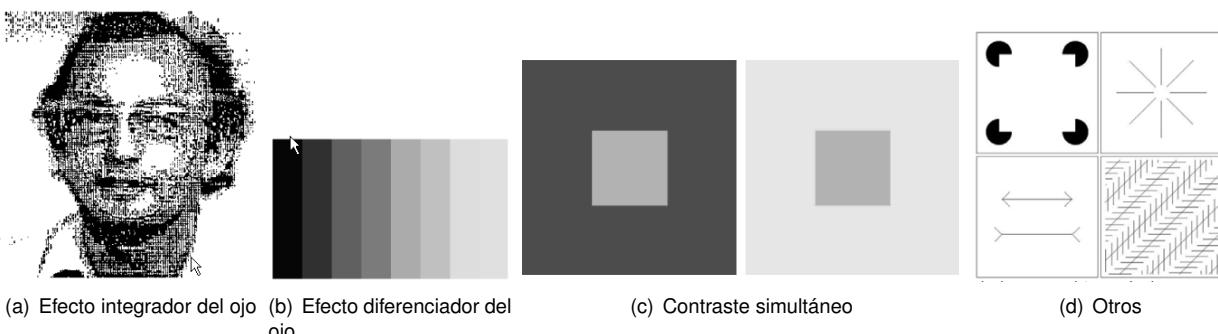
Figura 3: **Baja capacidad de discriminación a baja luminosidad - Gran capacidad de discriminación a alta luminosidad**

La figura muestra una gráfica de $\log(\Delta I_c)/I$ en función de $\log I$. La curva muestra que la discriminación de iluminación es mala (el cociente de Weber es grande) con bajos niveles de iluminación, pero crece significativamente cuando se aumenta la iluminación de fondo. Las dos ramas de la curva reflejan el hecho de que los niveles bajos de iluminación son manejados por bastones, y los altos – con mejor discriminación – por los conos.

13. Fenómenos de percepción humana

Puede demostrarse que la iluminación percibida no es simplemente una función de la intensidad. Algunos ejemplos que nos demuestran esto son:

- **Efecto integrador del ojo:** Cuando se dibuja con "puntos negros" se puede hacer que algo parezca que tenga brillo dejándolo en blanco.
- **Efecto diferenciador del ojo:** Tendencia a sobre-estimar (o subestimar) los cambios de intensidad en regiones de cambio brusco. Se perciben intensidades variantes en regiones de intensidad constante, cerca de las fronteras de cambio.
- **Contraste simultáneo:** Un objeto se percibe más oscuro a medida que se aclara el fondo.
- **Otros:** Agrupamiento de objetos, ilusiones ópticas, etc.



14. Resolución y cuantización en la visión humana

El muestreo es el factor principal para determinar la **resolución espacial** de una imagen. Básicamente, la resolución espacial es el grado de detalle discernible en una imagen.

La potencia de 2 que determina el número de niveles de gris es usualmente 8 bits, es decir, 256 diferentes niveles de gris. Algunas aplicaciones especializadas utilizan 16 bits.

Usualmente decimos que una imagen digital de tamaño $M \times N$ con L niveles de gris tiene una resolución espacial de $M \times N$ pixeles y una resolución de nivel de gris de L niveles.

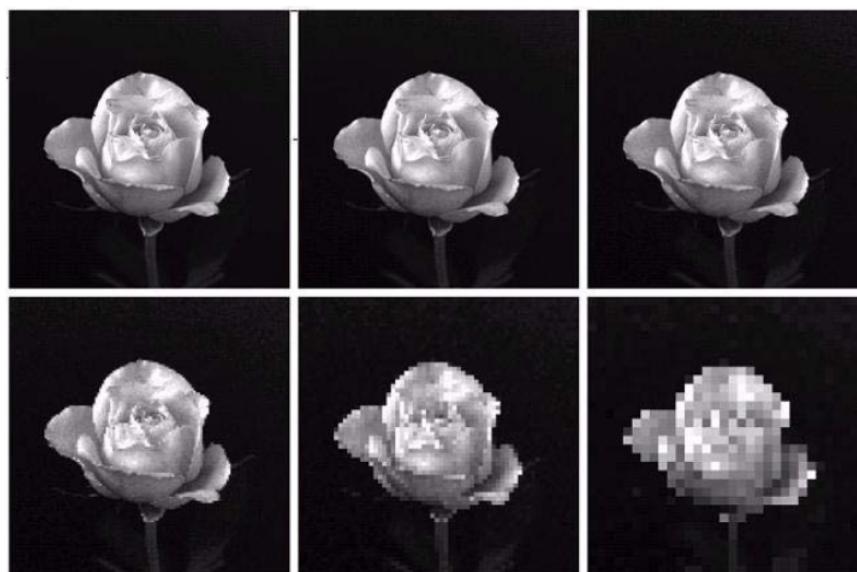


Figura 4: Imagen de 1024x1024 original y sus submuestreos (ampliados al tamaño de la primera) de 512x512, 256x256, 128x128, 64x64 y 32x32

En el primer ejemplo vemos una imagen con resolución espacial de 1024 x 1024 y 8 bits para representar los niveles de gris. Las imágenes siguientes han sido submuestreadas a partir de esta primera imagen, y se han agrandado para comparar los resultados. El submuestreo consistió en eliminar un número apropiado de columnas y renglones de la imagen original. Por ejemplo, en la segunda imagen se borraron una columna y un renglón sí y uno no, obteniéndose una imagen de la mitad del tamaño, 512 x 512. Nótese que a partir de la tercera imagen, de 256 x 256, aparece un fino patrón de "tablero de ajedrez", también llamado "pixelado" en los bordes de la flor. Este patrón se va haciendo más notorio hasta llegar a la última imagen, de 32 x 32.

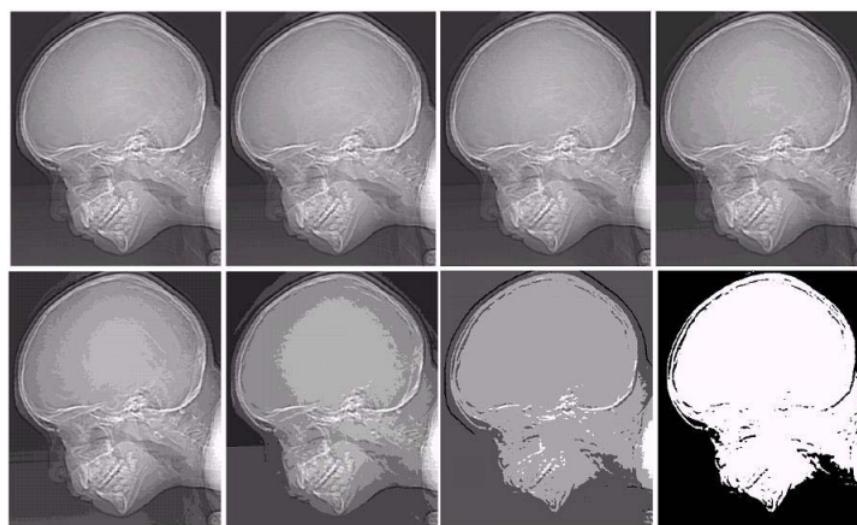


Figura 5: Imagen 472 x 374 en 256, 128, 64, 32, 16, 8, 4 y 2 niveles de gris

En el segundo ejemplo tenemos una imagen CAT con el mismo número de muestras (resolución espacial constante), pero

reducimos el número de niveles de gris desde 256 a 2, en potencias enteras de 2. Nótese a partir de la cuarta imagen, con 32 niveles, un casi imperceptible conjunto de “manchas” delimitadas en áreas suaves de niveles de gris. Este efecto, ya bastante notable en la figura con 8 niveles, es llamado falso contorno.

Un tamaño aproximado en el que una imagen estará libre de falsos contornos y “pixelados” será 256 x 256 con 128 ó 256 niveles de gris. Cuando la imagen tiene alrededor de 30 niveles de gris o menos, aparecen artefactos y falsos contornos que son detectados con facilidad por el sistema visual.

¿Relación perfecta entre muestreo (N) y niveles de grises (k)?: No existen reglas determinadas para fijar la relación entre las variables, sino que la calidad de una imagen es una cualidad subjetiva, y depende del contenido de detalle de la imagen, y de la respuesta del observador. (Experimento de Huang)

2. UNIDAD 2: Operaciones en el Dominio Espacial

1. Operaciones Puntuales: Independientes de las características globales - Operaciones en una sola imagen

La mejora de la imagen se divide en 2 categorías: **métodos del dominio espacial** y **métodos del dominio de la frecuencia**. Los métodos del dominio espacial trabajan sobre el plano de la imagen, y en éste se manipulan directamente los píxeles de una imagen. En los métodos del dominio de la frecuencia se modifica la transformada de Fourier de una imagen. Existen técnicas que se basan en combinaciones de métodos de ambas categorías.

No hay una teoría general de mejora de la imagen. Cuando la imagen se procesa para interpretación visual, el observador es el que juzga qué tan bueno es un método: la evaluación visual de una imagen es un proceso altamente subjetivo. Cuando la imagen se procesa para ser percibida por una máquina, la evaluación es más fácil: el mejor procesamiento de la imagen es aquél que provoca un mejor reconocimiento por parte de la máquina.

1.1. Definiciones Previas

Clasificación de Transformaciones Espaciales:

- **Operaciones puntuales**: el resultado depende únicamente del valor de intensidad en el pixel procesado.
- **Operaciones locales**: el resultado es función del nivel de gris del pixel analizado y de los de su entorno (vecindad).
- **Operaciones globales**: la imagen se transforma sin considerar los píxeles individualmente.
- **Operaciones geométricas**: cambio de posición de cada pixel dependiente de la posición en la imagen original.

En todas las operaciones que se le realizan a las imágenes se aplica un método denominado **Clipping**, el cual consiste en delimitar los valores de salida para que se conserven en el rango $[0 - s_{max}]$.

1.2. Operaciones sobre imágenes

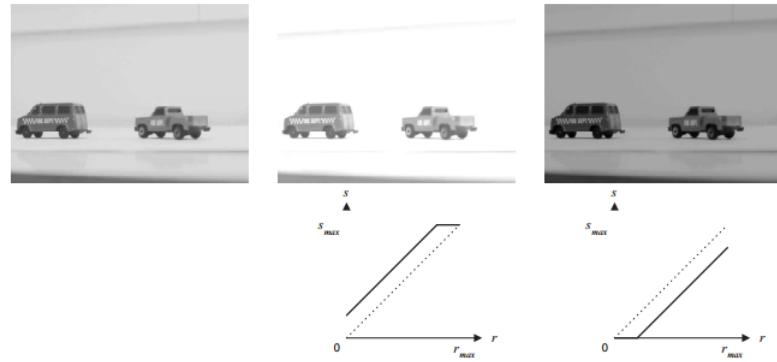
1. Transformaciones Rectilíneas

$$s = ar + c$$

Siendo: r el valor de entrada, a el factor de ganancia y c el offset.

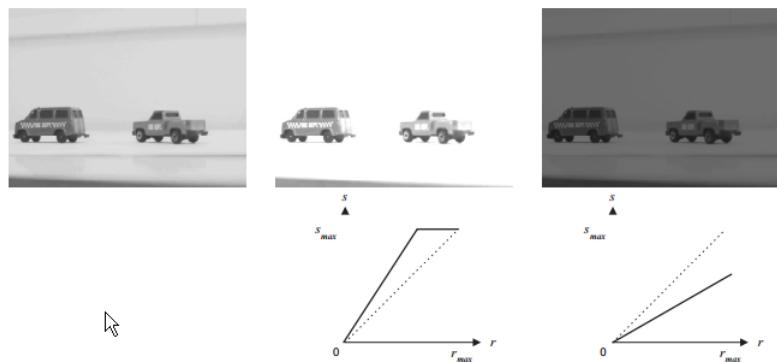
2. Manejo del offset ($a = 1$)

$$s = r + c \begin{cases} c > 0 & \text{imagen mas brillante} \\ c < 0 & \text{imagen mas oscura} \end{cases}$$



3. Amplificación ($c = 0$)

$$s = ar \begin{cases} a > 1 & \text{amplificacion} \\ a < 1 & \text{disminucion} \end{cases}$$



4. Negativos de la imagen

$$s = -r$$

Renormalización a rango positivo: $s = r_{max} - r$ con $r = [0; r_{max}]$.

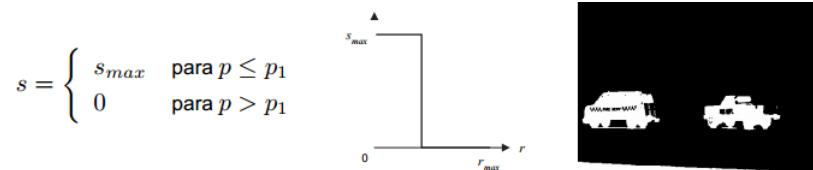
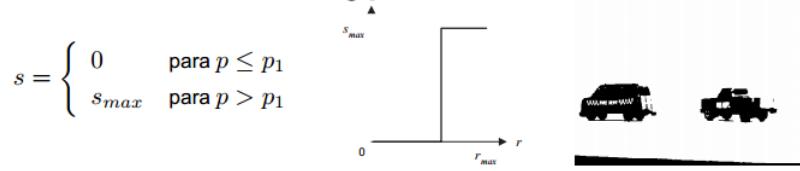
Caso particular de la ecuación general: $a = -1$ y $c = r_{max}$.



Este tipo de procesamiento es útil para mejorar niveles de blanco o gris en regiones oscuras de la imagen, especialmente cuando las áreas negras dominan en tamaño.

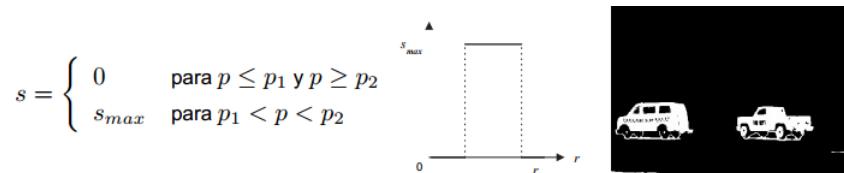
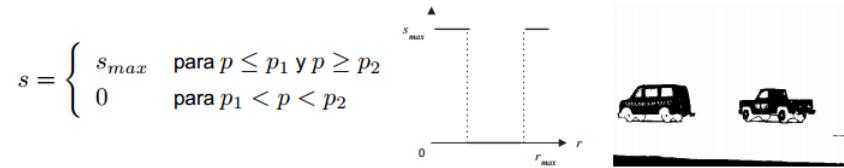
5. Umbral: binariza la imagen

Tomando un valor p_1 como transición para la binarización de la imagen tenemos que:



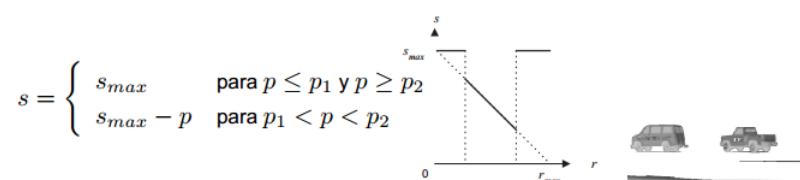
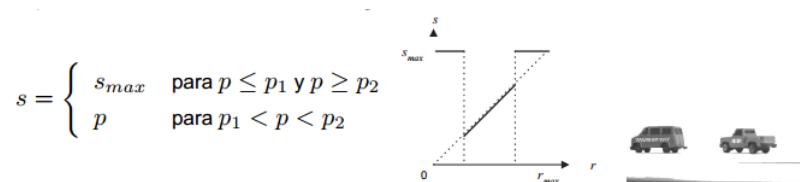
6. Intervalo de Umbral

Tomando dos valores p_1 y p_2 para definir el ancho del umbral tenemos que:



7. Umbral de Escala de Grises

Tomando dos valores p_1 y p_2 para definir el ancho del umbral tenemos que:



8. Transformación Logarítmica

$$s = c \cdot \log(1 + r)$$

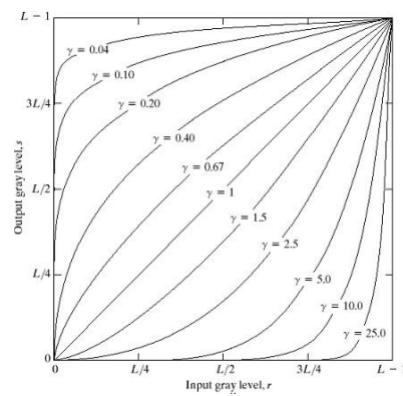


Utilizada en imágenes que tiene un rango dinámico grande ya que expande las intensidades oscuras y comprime las intensidades claras. Para realizar el efecto contrario se utiliza la inversa del logaritmo.

9. Transformación de Potencia

$$s = c \cdot r^\gamma \text{ con } c \text{ y } \gamma \text{ constantes positivas.}$$

Al variar γ obtenemos las diversas transformaciones que se grafican a continuación:



Como en el caso de la transformación logarítmica, las curvas de función de potencia con valores fraccionarios de γ mapean un rango pequeño de valores oscuros en un rango amplio de valores de salida y lo contrario con los valores claros de entrada. Sin embargo, en el caso de esta transformación, aquí obtenemos una familia de curvas de transformación simplemente variando γ . Las curvas con valores de $\gamma > 1$ tienen un comportamiento contrario a las curvas con valores de $\gamma < 1$. Cuando $\gamma = 1$ la transformación es la transformación identidad (si $c = 1$). Esta transformación también se utiliza para manipular el contraste de una imagen.



10. Modificación del rango dinámico

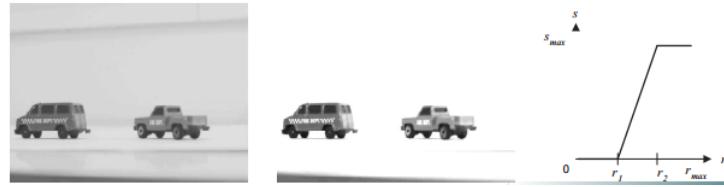
La idea de modificar el rango dinámico es aumentar el contraste de los niveles de grises que se están procesando.

Se definen los puntos de control (r_1, s_1) y (r_2, s_2) que determinan la forma de la función de transformación. Si $r_1 = s_1$ y $r_2 = s_2$ la transformación es una función lineal que no produce cambios en el nivel de gris. Si $r_1 = r_2$, $s_1 = 0$ y $s_2 = L - 1$, la transformación se vuelve una función de umbral. Valores intermedios producen distintos grados de distribución del nivel de gris, afectando el contraste.



11. Ventaneo

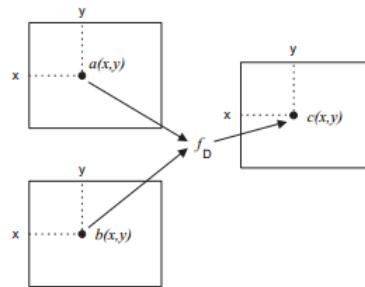
Se utiliza en imágenes con rango mayor al dispositivo de visualización (ej: tomografía computada)



2. Operaciones Puntuales: Independientes de las características globales - Operaciones entre varias imágenes

Transformación punto a punto entre dos o más imágenes, para crear una imagen de salida, de igual dimensión.

$c_{(x,y)} = k \cdot f_D(a_{(x,y)}, b_{(x,y)})$ con f_D siendo una función lineal o no lineal, con factor de escala k apropiado.



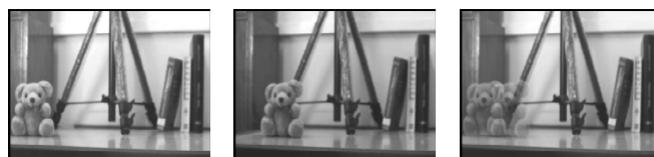
2.1. Operaciones sobre varias imágenes

1. Suma

$$c_{(x,y)} = \frac{a_{(x,y)} + \dots + k_{(x,y)}}{N}$$

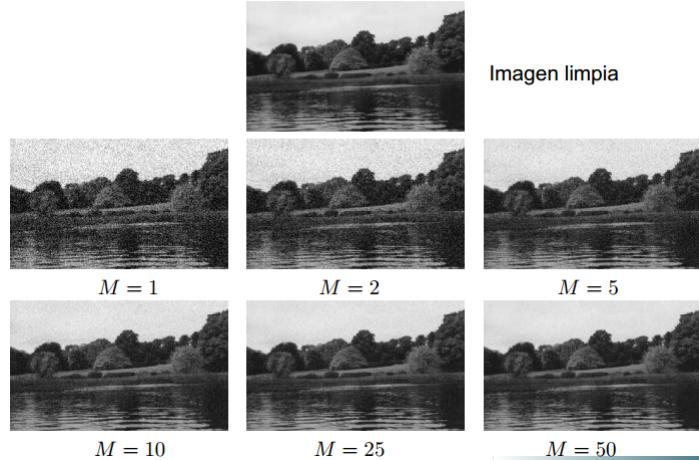
siendo N el número de imágenes a tratar. Esta operación se utiliza para la **fusión de imágenes** y la **reducción del ruido**.

Fusión de imágenes: Una utilidad de esto es para lograr una sensación de movimiento.



Reducción de ruido mediante promediado: A una imagen con ruido se la puede expresar como la suma entre la imagen original más el ruido. Para limpiar estas imágenes con ruido lo que se realiza es un promedio de M observaciones $g_i(x, y)$ con $i = 1, 2, \dots, M$.

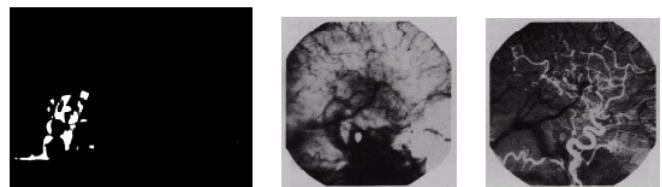
PROMEDIO DE IMÁGENES: $\bar{g}_{(x,y)} = \frac{1}{M} \sum_M^{i=1} g_{(x,y)}$



2. Resta

$c_{(x,y)} = k(a_{(x,y)} - b_{(x,y)})$ siendo k una función de escala.

Esta operación es utilizada en el estudio del movimiento para detectar el cambio producido entre imágenes captadas en instantes de tiempo sucesivos. También se la utiliza para el resaltado de estructuras (segmentación).



3. Multiplicación

Utilizada para aislar regiones de interés en la imagen, multiplicando por una máscara binaria.



4. Operaciones Lógicas: NOT, OR, AND, XOR

Estas operaciones se realizan sobre imágenes binarias.



5. Operaciones Relacionales: $<$, $>$, \leq , \geq

Estas operaciones se realizan sobre imágenes en escala de grises, y dan como resultado imágenes binarias.

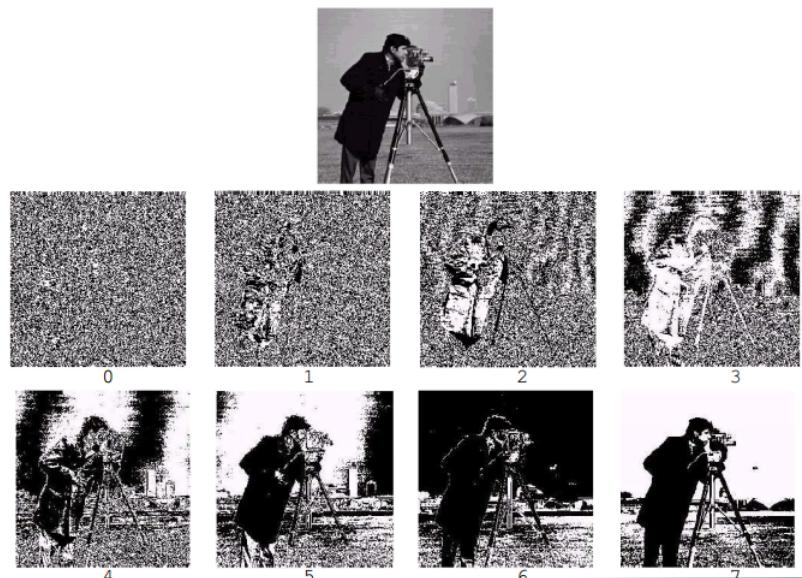


6. Fraccionamiento de planos de bits

Para algunas aplicaciones, es deseable conocer la contribución de cada bit específico en la apariencia total de la imagen. Supóngase que cada pixel de una imagen es representado por 8 bits. Imagínese que la imagen se compone de 8 planos de 1 bit, desde el plano de bits 0 (el bit menos significativo, LSB), hasta el plano de bits 7 (más significativo MSB). El plano de bits cero contiene todos los LSB de cada byte de cada pixel en la imagen y el plano 7 contiene los MSB de cada byte de cada pixel de la imagen.

Corresponde a expresar los niveles de gris de una imagen de p bits como un polinomio de la forma:

$$a_{p-1}2^{p-1} + a_{p-2}2^{p-2} + \dots + a_12^1 + a_02^0 \text{ siendo } a_i \text{ igual a } 0 \text{ o } 1.$$



Separar una imagen en planos de bits es útil para analizar la importancia relativa de cada bit en una imagen. Este proceso ayuda a determinar la cantidad de bits necesaria para cuantificar una imagen. También es útil para la compresión de imágenes, como se verá después.

3. Operaciones Puntuales: Dependiente de las características globales - Manejo de Histogramas

El histograma es una función discreta que describe de manera global la apariencia de una imagen: número de píxeles en función de las intensidades de grises. Ver fig. (6). Probabilidad $P(g)$ de ocurrencia de un determinado nivel g :

$$P(g) = \frac{n(g)}{n}, \text{ donde } \begin{cases} n(g) : \text{número de píxeles de intensidad } g \\ n : \text{número de píxeles de la imagen} \end{cases}$$

$$P(g) \leq 1 \forall g, \text{ y } \sum_g P(g) = 1$$

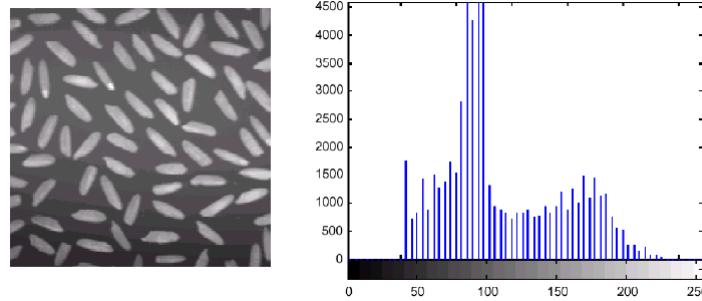


Figura 6: Histograma de una imagen.

3.1. Propiedades estadísticas del Histograma

- **Media:** valor promedio de los niveles de gris (brillo general de la imagen). Ver fig. (7).

$$\bar{g} = \sum_{g=0}^{L-1} g P(g) = \frac{1}{n} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)$$

- **Varianza:** dispersión de valores alrededor de la media (contraste de la imagen). Ver fig. (8).

$$\sigma^2 = \sum_{g=0}^{L-1} (g - \bar{g})^2 P(g) = \frac{1}{n} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x, y) - \bar{g}]^2$$

- **Asimetría sobre la media en la distribución de los niveles de gris:**

$$a = \sum_{g=0}^{L-1} (g - \bar{g})^3 P(g)$$

- **Energía:** informa sobre la distribución de los niveles de gris. Valor máximo 1 para imagen con un único nivel de gris, disminuye con el aumento del número de grises.

$$E = \sum_{g=0}^{L-1} P(g)^2$$

- **Entropía:** incertezza en la distribución, aumenta con el número de grises.

$$e = - \sum_{g=0}^{L-1} P(g) \log_2 [P(g)]$$

En las figuras (9) y (10) se ilustran los histogramas de imágenes con variación de brillo y contraste, respectivamente.



Figura 7: Media de una imagen con distinto brillo.



Figura 8: Varianza de una imagen con distinto brillo.

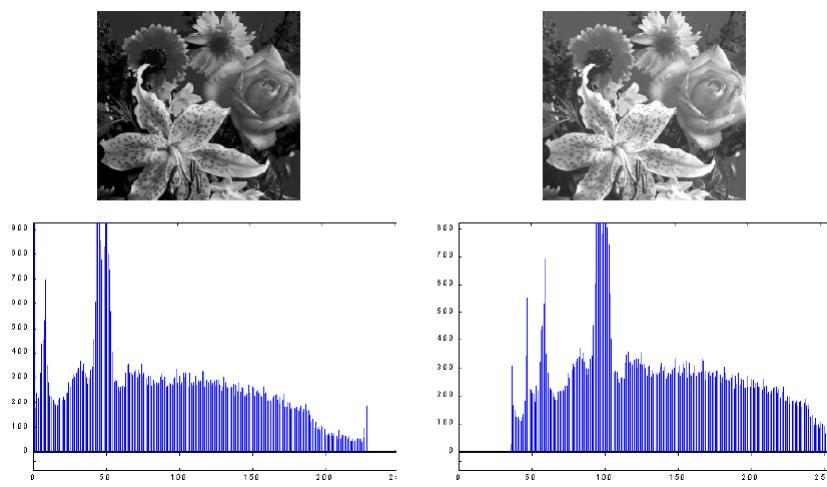


Figura 9: Variación de brillo.

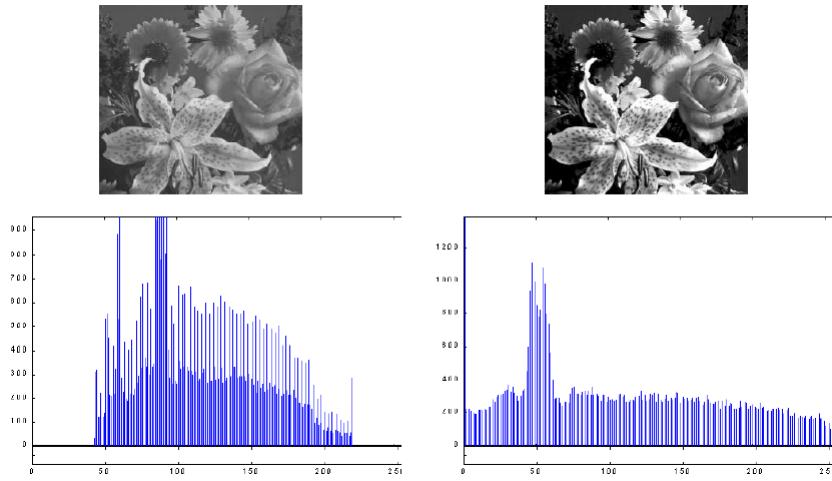


Figura 10: Varición de contraste.

3.2. Manipulación del Histogramas

Operaciones que se realizan sobre la distribución de grises, a fin de mejorar el aspecto de la imagen. El proceso consiste en el manejo de brillo y contraste visto anteriormente en operaciones puntuales, desplazamiento, corrección gamma, etc. La **Ecualización** (o igualación) es la manipulación del histograma para mejorar el contraste. Redistribuye los grises de la imagen original sobre todos los grises disponibles.

La operatoria se basa en controlar la función de densidad de probabilidad mediante una función de transformación. Suponemos imágenes digitales (valores discretos) con intensidades entre 0 y 1. La relación que define las probabilidades es:

$$p_r(r_k) = \frac{n_k}{n}, \text{ con } \begin{cases} 0 \leq r_k \leq 1 \\ k = 0, 1, \dots, L - 1 \end{cases}$$

La función de transformación utilizada es la Función de Distribución Acumulada (CDF), y está dada por:

$$s_k = T(r_k) = \sum_{j=0}^k \frac{n_j}{n} = \sum_{j=0}^k p_r(r_j), \text{ con } \begin{cases} 0 \leq r_k \leq 1 \\ k = 0, 1, \dots, L - 1 \end{cases}$$

El resultado de esta operación, es decir, la ecualización del histograma trae buenos resultados, como puede verse en la fig. (11). Pero no siempre es así, particularmente cuando el histograma origen está muy localizado (*pico*), pudiendo producir falsos bordes y regiones de diferente intensidad. En estos casos también aumenta la "granulosidad". Ver fig. (12)

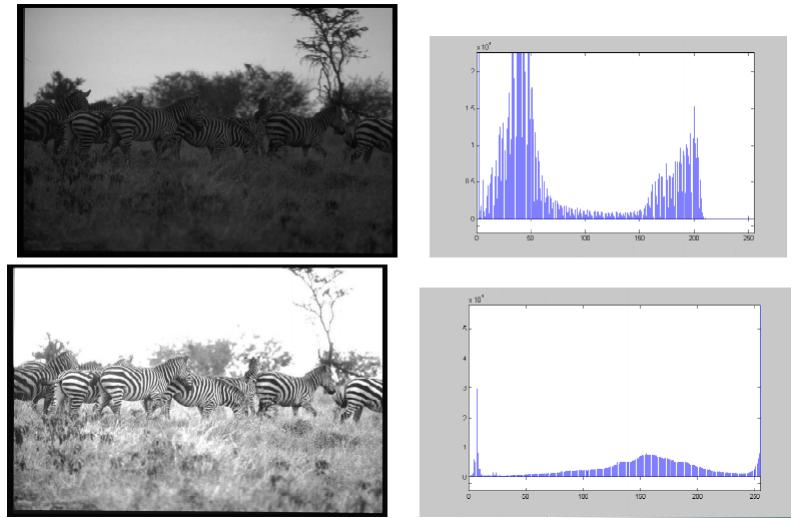


Figura 11: Ejemplo de Ecualización de un Histograma.

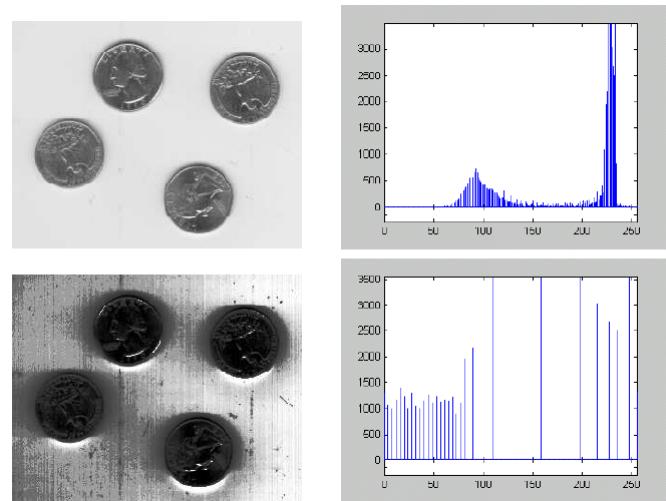


Figura 12: Ejemplo de Ecualización de un Histograma con Picos.

4. Operaciones Locales - Conceptos

Variaciones de los niveles de gris entre píxeles de acuerdo a su distancia:

- **Variaciones rápidas:** componentes de altas frecuencias. Corresponde a bordes de los componentes de la escena (aparición de transiciones claro/oscuro).
- **Transiciones suaves:** componentes de bajas frecuencias. Corresponde al nivel general de iluminación de la escena.

Las vecindades de un pixel particular puede aportar información útil acerca de los niveles de iluminación o detalles de la escena en la zona. Muchas operaciones de realce se efectúan sobre regiones de interés alrededor de cada pixel de la imagen. El resultado es función del nivel de gris del pixel analizado y de los de su entorno.

5. Filtrado Espacial - Conceptos

Modelo de la operación de realce en el dominio espacial:

$$g(x, y) = T(f)(x, y), \text{ con } \begin{cases} f(x, y) & \text{entrada al sistema (imagen bajo análisis)} \\ T & \text{transformación aplicada (lineal o no lineal)} \\ g(x, y) & \text{salida del sistema (imagen realizada)} \end{cases}$$

Transformaciones lineales (extensión del concepto 1-D):

- **Operador T:** lineal e invariante al desplazamiento (LSI).
- **$h(x, y)$:** respuesta al impulso del sistema.
- **Salida del sistema:** convolución.

La convolución 2D continua y discreta viene dada por las siguientes expresiones:

$$\begin{aligned} f(x, y) * h(x, y) &= \int_{-\infty}^{\infty} f(\tau_x, \tau_y) h(x - \tau_x, y - \tau_y) d\tau_x d\tau_y \\ f(x, y) * h(x, y) &= \sum_{s=-\infty}^{\infty} \sum_{t=-\infty}^{\infty} f(s, t) h(x - s, y - t) \\ &= \sum_{s=-\infty}^{\infty} \sum_{t=-\infty}^{\infty} h(s, t) f(x - s, y - t) \end{aligned}$$

Para reducir la complejidad computacional $h(s, t) = 0$, para $(s, t) \notin \Delta$, con Δ conjunto pequeño de $m \times n$ (vecindad9). Entonces, se tiene:

$$g(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b h(s, t) f(x - s, y - t)$$

con $a = \frac{m-1}{2}$, $b = \frac{n-1}{2}$, $x = 0, 1, \dots, M-1$; $y = 0, 1, \dots, N-1$.

El Kernel de convolución es una ventana de coeficientes que definen h . Es de tamaño arbitrario, comúnmente de 3×3 ó 5×5 . Idea del espejado para convolución con talla de $h = 3 \times 3$. Para cada pixel (x, y) de la imagen, la sumatoria se puede ver gráficamente con la siguiente superposición de matrices (ver fig. (13)):

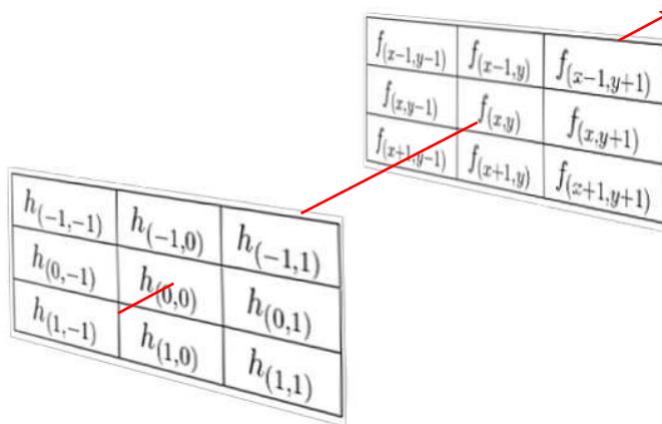


Figura 13: Superposición de matrices.

La sumatoria de convolución opera aplicando sobre f la máscara dada por la siguiente tabla (1), que es la versión rotada de la máscara anterior:

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

Cuadro 1: Máscara.

En vez de operar con el kernel de convolución h , podemos definir los coeficientes de la máscara de filtrado w (versión rotada de h), como se muestra en la fig. (14), y operar con la sumatoria de correlación:

$$\begin{array}{|c|c|c|} \hline w_1 & w_2 & w_3 \\ \hline w_4 & w_5 & w_6 \\ \hline w_7 & w_8 & w_9 \\ \hline \end{array} =
 \begin{array}{|c|c|c|} \hline w_{(-1,-1)} & w_{(-1,0)} & w_{(-1,1)} \\ \hline w_{(0,-1)} & w_{(0,0)} & w_{(0,1)} \\ \hline w_{(1,-1)} & w_{(1,0)} & w_{(1,1)} \\ \hline \end{array} =
 \begin{array}{|c|c|c|} \hline h_{(1,1)} & h_{(1,0)} & h_{(1,-1)} \\ \hline h_{(0,1)} & h_{(0,0)} & h_{(0,-1)} \\ \hline h_{(-1,1)} & h_{(-1,0)} & h_{(-1,-1)} \\ \hline \end{array}$$

Figura 14: Kernel w .

$$g(x, y) = h(x, y) * f(x, y) = w(x, y) \circledast f(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x + s, y + t)$$

$$g(x, y) = w_1 f_{(x-1, y-1)} + w_2 f_{(x, y-1)} + w_3 f_{(x+1, y-1)}$$

$$g(x, y) = w_4 f_{(x-1, y)} + w_5 f_{(x, y)} + w_6 f_{(x+1, y)}$$

$$g(x, y) = w_7 f_{(x-1, y+1)} + w_8 f_{(x, y+1)} + w_9 f_{(x+1, y+1)}$$

Filtrado es el nombre dado al desplazamiento sucesivo de la máscara por la imagen, ya que según la elección de coeficientes de w se limitan o amplifican determinadas frecuencias espaciales. La clasificación de los filtros se puede apreciar en la fig. (15). Para los píxeles del borde se tiene la condición de frontera:

- Borde libre: máscara truncada.
- Borde fijo: la imagen es extendida mediante repetición de la fila/columna, o poniendo valores de intensidad fijos.
- Borde periódico: se convoluciona agregando la fila/columna opuesta (toroide).

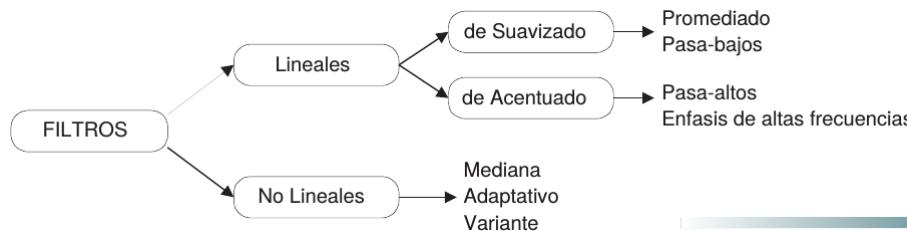


Figura 15: Clasificación de Filtros.

5.1. Filtros Lineales de Suavizado

El suavizado de la imagen es la reducción de variaciones rápidas (saltos bruscos) de intensidad entre píxeles vecinos. Los filtros de suavizado se utilizan para:

- **Desenfoque**: preprocesamiento para eliminar detalles no deseados (pequeños) antes de la extracción de objetos grandes. Corrección de fragmentos de líneas perdidos.
- **Reducción de ruido**: mediante filtros lineales o no lineales.

Dentro de los filtros de Suavizado se encuentran los filtros Promediadores y los filtros Pasa-Bajos. En las imágenes siguientes se pueden ver ejemplos de aplicación de este tipo de filtros. Ver fig. (16) a (19).

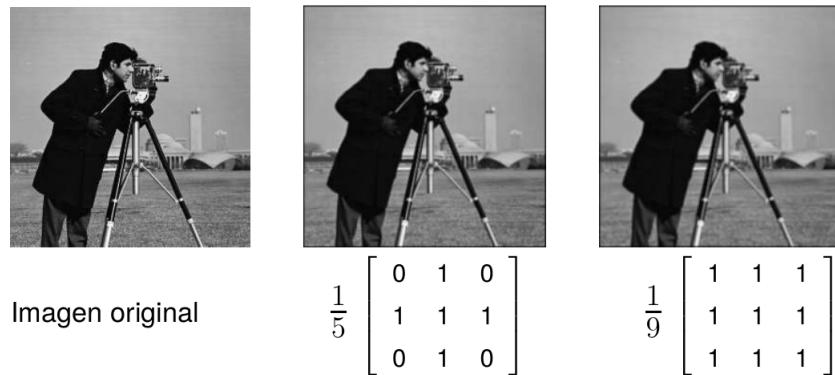


Figura 16: Filtros promediadores: pesos iguales.

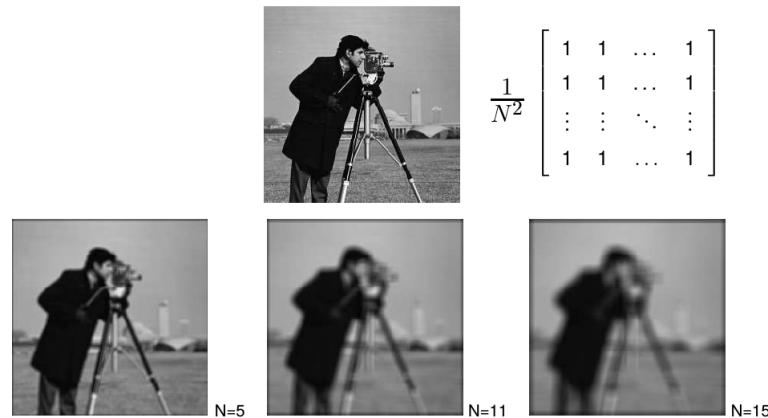
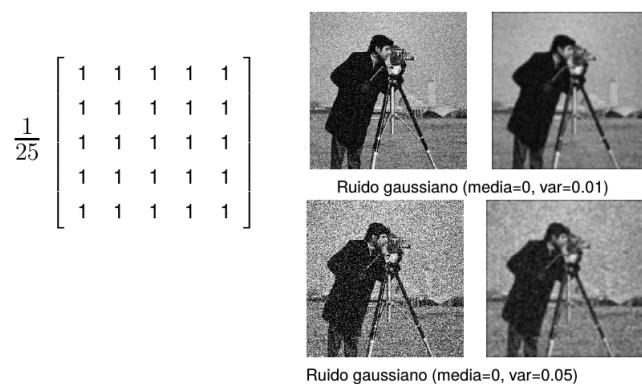
Figura 17: Filtros promediadores: efecto de la talla de w en el desenfoque.

Figura 18: Filtros promediadores: reducción de ruido.

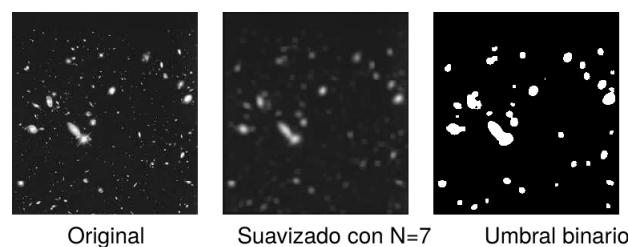


Figura 19: Filtros promediadores: localización de objetos grandes mediante desenfoque.

5.2. Filtros Lineales de Acentuado

El objetivo de estos filtros es resaltar los detalles finos y variaciones rápidas (saltos bruscos) de intensidad entre pixeles vecinos. La máscara es un operador LSI con coeficientes positivos en el centro y valores negativos alrededor. Dentro de esta clasificación se encuentran los siguientes tipos:

- **Pasa-altos:** aplicación directa de una máscara.
- **Enfasis de altas frecuencias:** operación aritmética entre imágenes. Conocido también como "máscara difusa".
- **Alta potencia (high-boost).**

Para los Filtros Pasa-Altos: la suma de coeficientes de la máscara determina la imagen resultante.

- Suma = 1: realce de altas frecuencias sin alterar las bajas frecuencias.

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

- Suma = 0: extracción de altas frecuencias, eliminando las bajas frecuencias.

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

En la fig. (20) y (21) puede apreciarse el resultado de aplicar cada uno de estos filtros.

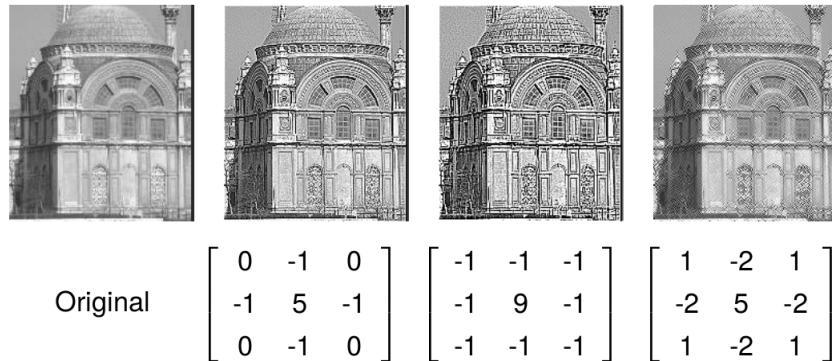


Figura 20: Filtros pasa-altos con suma = 1.

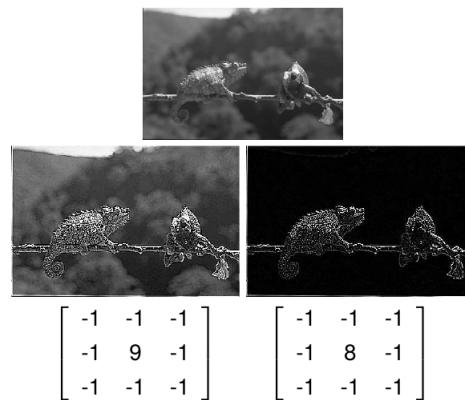


Figura 21: Filtros pasa-altos con suma = 0: elimina zonas homogéneas.

El Filtro de Máscara difusa es un filtro Pasa-Altos obtenido mediante una operación aritmética, calculada como la diferencia entre la imagen original y una versión suavizada. Ver fig. (22).

$$g(x, y) = f(x, y) - PB(f(x, y))$$

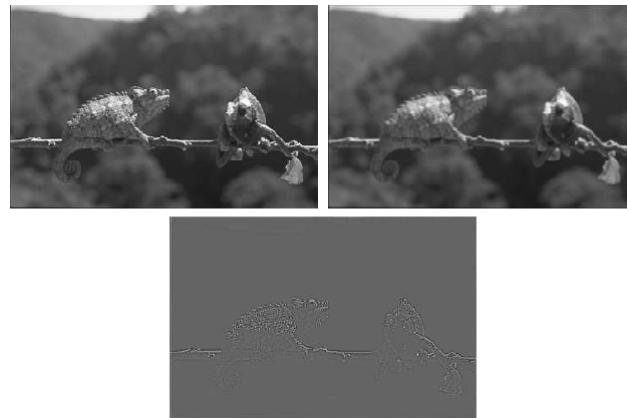
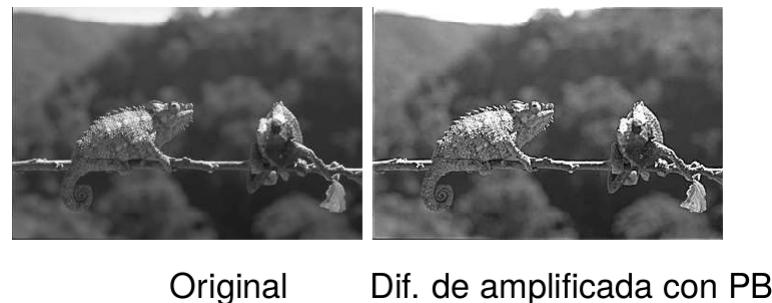


Figura 22: Aplicación del filtro de Máscara Difusa.

El Filtrado de Alta Potencia (high-boost) es una generalización del máscara difusa. Una forma de obtener la salida es mediante la diferencia entre una versión amplificada de la imagen original y una versión suavizada. Ver fig. (23)

$$g(x, y) = Af(x, y) - PB(f(x, y)), \text{ con } A \geq 1$$



Original Dif. de amplificada con PB

Figura 23: Filtro de Alta Potencia (high-boost) - Primera Forma.

Otra forma de obtener la salida de este filtro es la siguiente, ver fig. (24)

$$\begin{aligned} g(x, y) &= Af(x, y) - PB(f(x, y)) \\ &= (A - 1)f(x, y) + f(x, y) - PB(f(x, y)) \\ &= (A - 1)f(x, y) + PA(f(x, y)) \end{aligned}$$

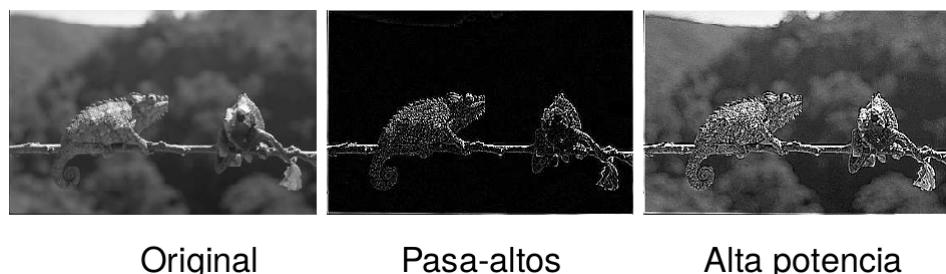


Figura 24: Filtro de Alta Potencia (high-boost) - Segunda Forma.

5.3. Filtros No Lineales

El filtro de promedio reduce el ruido, pero produce imágenes borrosas (pérdida de detalles). Utilizados para reducción de ruido, los filtros no lineales corresponden a técnicas de restauración. Por su amplia utilización, veremos el filtro de mediana:

- Ordenamiento de intensidades en la ventana.
- Asignación del valor intermedio del conjunto.

$$g(x, y) = \text{mediana}\{f(x + s, y + t), \forall(s, t) \in w\}$$

- Excelente capacidad para eliminar ruido "sal y pimienta"(impulsivo).
- No introduce desenfoque como el filtro pasa-bajos lineal.
- Resultados limitados por el tamaño de la ventana.

En la fig. (25) se puede apreciar la diferencia entre aplicar un filtro lineal de suavizado y un filtro no lineal de mediana.

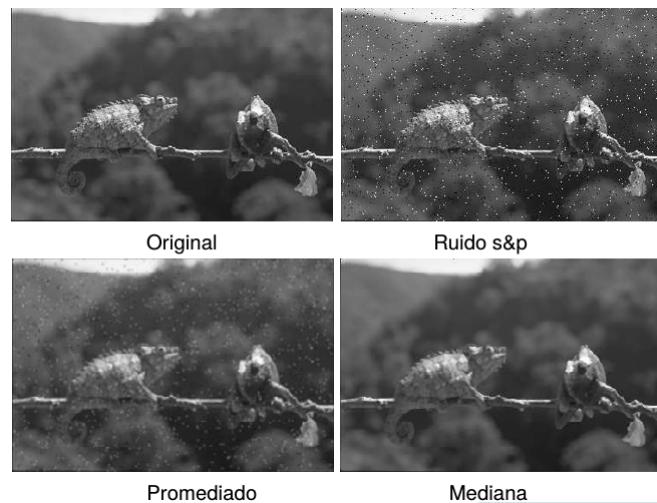


Figura 25: Comparación con filtro de promediado (ambas máscaras de 3×3).

5.4. Combinación de Métodos

Los métodos vistos hasta ahora fueron aproximaciones de realce individuales. Tareas más complejas pueden requerir una secuencia de procesamientos complementarios. Ejemplos:

- En imágenes con ruido gaussiano: filtro de promediado + alta potencia.
- En imágenes con ruido impulsivo: mediana + máscara difusa.
- Operación puntual al final para mejora del rango dinámico (log, ecualización de histograma).

6. Procesamiento de COLOR

6.1. Fundamentos del Color

Importancia del Color en PDI

- El color es un poderoso descriptor que simplifica la identificación de objetos y su extracción de una escena.
- Los seres humanos podemos discernir entre miles de tonalidades de color, comparado a cerca de 20 niveles de gris.

Dos formas de Trabajo

- **Pseudo-color:** imágenes monocromas coloreadas a través de la asignación de un color a una intensidad de gris.
- **Color:** imágenes adquiridas con un sensor de color o multiespectral.

Caracterización de la luz

Básicamente, los colores que los seres humanos y otros animales perciben en un objeto están determinados por la naturaleza de la luz reflejada desde el objeto. La luz visible se compone de una delgada banda de frecuencias del espectro electromagnético. Un cuerpo que refleja luz balanceada en todas las longitudes de onda visibles se verá blanco. Sin embargo, un cuerpo que favorezca la reflectancia en un rango limitado del espectro visible mostrará algunos tonos de color.

- **Luz acromática** (sin color): intensidad como único atributo. Nivel de gris: medida escalar de intensidad.
- **Luz cromática:** posee dispersión de energía desde los 400 a los 700 nm aprox.

Parámetros que describen la calidad de la fuente lumínica cromática

- **Radiancia:** cantidad total de energía que sale de la fuente (Watts [W]).
- **Luminancia:** cantidad de energía procedente de la fuente que percibe un observador (Lumens, [lm]). Ej: fuente infrarroja (alta Radiancia, baja Luminancia).
- **Brillo:** descriptor subjetivo (noción de intensidad). Ej: azul y verde de igual intensidad.

Clasificación de Colores

- **Colores primarios de luz (modelo aditivo):** rojo = 700 nm, verde = 546.1 nm y azul = 435.8 nm.
- **Colores secundarios:** cyan, magenta y amarillo. Formados por la suma de colores primarios.
- **Colores primarios de pigmento:** aquellos que absorben un color primario de luz y reflejan los restantes (modelo sustractivo). Pigmentos primarios: amarillo, magenta y cyan.

Características del Color

- **Brillo:** noción acromática de intensidad.
- **Tono (hue):** λ dominante en la onda percibida ("color" del objeto).
- **Saturación:** pureza relativa del color. Inversamente proporcional a la mezcla de luz blanca. Ejemplos: rojo, verde y azul: completamente saturados. Rosado, lavanda y verde manzana: menos saturados.

Se llama cromaticidad al tono y saturación juntos. Entonces podemos caracterizar un color por su cromaticidad y su brillo.

6.2. Modelos de Color

Los modelos (o espacios) de color son estándares de especificación de colores. Especifican un sistema de coordenadas 3D y un subespacio en el cual queda definido cualquier color como un punto único. Surgen con el hardware: modelos para monitores, impresoras, etc.; o con el software: modelos para PDI, computación gráfica, animación.

Modelo RGB (Red-Green-Blue)

En el modelo RGB, cada color aparece en sus componentes primarios del espectro de rojo, verde, y azul. Este modelo se basa en un sistema de coordenadas cartesianas que forma un cubo unitario (valores entre [0,1]) en el que los valores RGB están en 3 esquinas. Los valores de cian, magenta y amarillo están en las otras 3 esquinas, el negro está en el origen y el blanco está en el vértice más alejado del origen.

La escala de gris (puntos con valores de RGB iguales) se extiende del negro al blanco en la línea que une los puntos

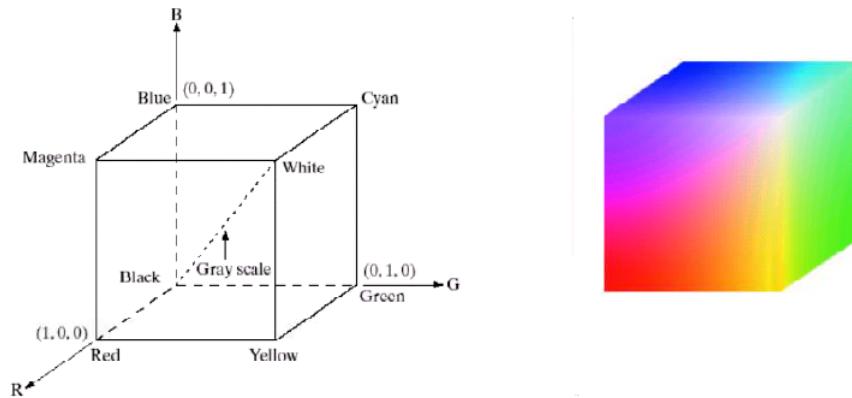


Figura 26: Modelo RGB.

negro y blanco. Los diferentes colores del modelo son puntos sobre o dentro del cubo, y se definen por medio de vectores que se extienden desde el origen.

Las imágenes representadas en el modelo RGB consisten en 3 imágenes componentes, cada una de un color primario. El número de bits utilizado para representar cada pixel en el espacio RGB se llama **profundidad de pixel**. Considérese una imagen en la que cada una de las imágenes roja, azul y verde es una imagen de 8 bits. En estas condiciones se dice que cada pixel de color RGB (una tupla (R, G, B)) tiene una profundidad de 24 bits. El número total de colores en una imagen RGB de 24 bits es de $(2^8)^3 = 2^{24} = 16,777,216$ colores.

La imagen se forma a partir de 3 planos independientes, uno para cada color primario:

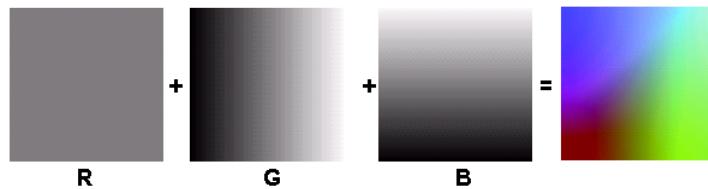


Figura 27: Ejemplo de composición de imagen RGB a partir de los 3 planos R, G y B por separado.

Modelos CMY y CMYK

El cian, magenta y amarillo son los colores secundarios de la luz o, alternativamente, los colores primarios de los pigmentos. Por ejemplo, cuando una superficie cubierta por pigmento cian se ilumina con luz blanca, no se refleja luz roja de la superficie; es decir, el cian sustrae luz roja de la luz blanca reflejada. La mayoría de los dispositivos que depositan pigmentos de color en el papel, como las impresoras y copiadoras, requieren entrada de datos en CMY o realizan una conversión interna de RGB a CMY. Esta conversión se lleva a cabo con la operación:

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Esta ecuación muestra que la luz reflejada de una superficie cubierta de cian puro no contiene rojo ($C = 1 - R$ en la ecuación). De igual manera, el magenta puro no refleja verde, el amarillo puro no refleja azul.

Modelo CMYK: impresión en 4 colores, donde el negro se agrega al modelo para evitar generar el negro mediante suma de pigmentos primarios.

Modelos HSV y HSI

- **HSV:** Hue-Saturation-Value: Tono-Saturación-Valor
- **HSI:** Hue-Saturation-Intensity: Tono-Saturación-Intensidad

Desgraciadamente los modelos RGB y CMY no son convenientes para describir colores en los términos prácticos para la interpretación humana. Cuando un ser humano ve un objeto de color, lo describe por su tono, saturación y brillo.

El modelo de color HSI separa el componente de intensidad de la información de color en una imagen de color. Como resultado, este modelo es una herramienta ideal para desarrollar algoritmos basados en descripciones de color naturales e intuitivas para los seres humanos. En los modelos RGB y CMYK la sensación de color está acoplada al brillo, por lo que resultan útiles en generación de colores.

Ventaja del HSV: desacopla la información de crominancia (H, S) de la luminancia (V). Aplicaciones en algoritmos basados en descripción natural de colores. Ej: verificación automática de madurez de frutas, inspección del acabado de productos, etc.

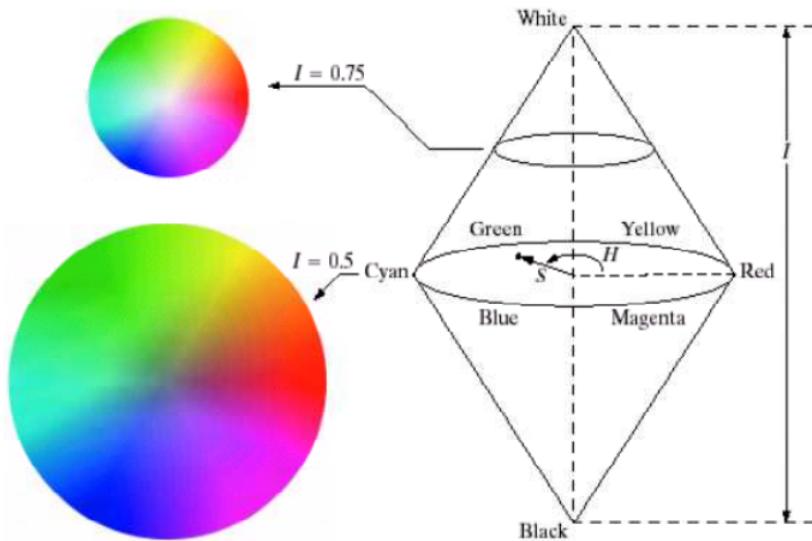


Figura 28: Modelo HSI.

Conversión de RGB a HSI

- **Tonalidad:**

$$H = \begin{cases} \theta & \text{si } B \leq G \\ 360 - \theta & \text{si } B > G \end{cases} \quad \text{donde } \theta = \cos^{-1} \left\{ \frac{\frac{1}{2}[(R-G)+(R-B)]}{[(R-G)^2+(R-B)(G-B)]^{\frac{1}{2}}} \right\}$$

- **Saturación:**

$$S = 1 - \frac{3}{R+G+B} [\min(R, G, B)]$$

- **Intensidad:**

$$I = 1 - \frac{1}{3}(R+G+B)$$

Se asume que los valores RGB han sido normalizados al rango [0, 1] y que el ángulo θ se mide desde el eje rojo del espacio HSI. El tono puede normalizarse al rango [0, 1] dividiendo entre 360. Los otros componentes ya se obtienen normalizados si los valores de RGB estaban normalizados.

Manipulación de los componentes HSI Para cambiar un color individual en cualquier región de una imagen RGB, cambiamos los valores de la región correspondiente en la imagen de tono, después convertimos la nueva imagen H, sin cambiar las imágenes S e I a RGB. Para cambiar la saturación, hacemos lo mismo con la imagen S.

6.3. PDI en Pseudo-Color

El procesamiento del falso color (a veces llamado seudo color) consiste en asignar colores a valores de niveles de gris según un criterio especificado. El uso primordial del pseudo-color es para visualización humana e interpretación de eventos en escala de gris en una imagen o una secuencia de imágenes. Una de las principales motivaciones en el uso de color es el hecho de que los seres humanos podemos discernir miles de tonalidades e intensidades de color, comparado con cerca de 20 tonos de grises.

Rodajas de Intensidad (división del color)

Consiste en tratar a una imagen monocroma como una función 3D: intensidad en función de las coordenadas espaciales. Colocación de un plano paralelo al plano de coordenadas, asignando colores diferentes a cada lado (binarización de color). Si se asigna un color diferente a cada lado del plano en la figura, cada pixel cuyo nivel de gris esté sobre el plano será codificado con un color, y cualquier pixel bajo el plano será codificado con otro. El resultado es una imagen en 2 colores cuya apariencia relativa puede ser controlada moviendo el plano divisor.

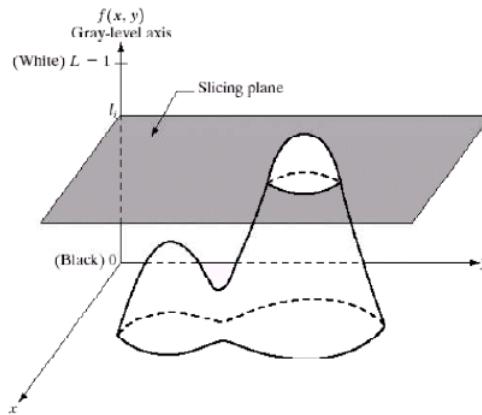


Figura 29: Rodajas de Intensidad.

Si se mira de arriba son 2 óvalos pegados y aplicando este método, la parte mas blanca, que es la 'circunferencia' del ovalo grande, la pinta de un color y el resto de otro. Esto se puede extender a p planos.

Transformación de grises a color Aplicación de tres transformaciones independientes a los grises de la imagen. Las salidas se inyectan a los canales R, G y B. Técnica utilizada por la flexibilidad de la definición de las funciones.

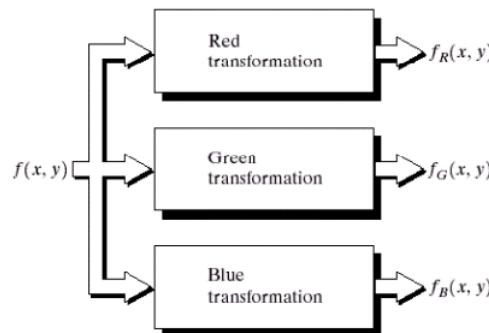


Figura 30: Transformación de grises a color.

Ejemplos de aplicación: Rayos X en control de equipajes. Se logran diversos acabados utilizando mapas de colores diferentes (por ejemplo Rainbow de 1 o 2 ciclos).

6.4. Procesamiento a todo color

Hay 2 categorías. En la primera se procesa cada componente de imagen individualmente y después se forma una sola imagen compuesta a partir de todas. En la segunda categoría se trabaja directamente con píxeles de color. Los píxeles de color se toman como vectores con componentes dados por los 3 (o más, en otros modelos) planos de color.

Para equivalencia entre aproximaciones deben satisfacer las condiciones: El proceso sea aplicable a escalares y vectores. La operación en cada componente del vector sea independiente de las otras componentes.

Transformaciones de Color Concepto: procesamiento de los componentes de color de una imagen en un modelo de color particular. Formulación general:

$$g(x, y) = T[f(x, y)]$$

$$s_i = T_i(r_1, r_2, \dots, r_n), \text{ con } i = 1, 2, \dots, n$$

donde

- r_i y s_i son las componentes de color de $f(x, y)$ y $g(x, y)$, respectivamente
- n es el número de componentes de color
- $T_i(r_1, r_2, \dots, r_n)$ son las funciones de mapeo de color

Cualquier transformación puede ser aplicada en cualquier modelo de color. Sin embargo, algunas transformaciones son más adecuadas para ciertos modelos y debe tenerse en cuenta la conversión entre modelos. Por ejemplo para oscurecimiento de una imagen, en RGB y CMY se tienen que modificar s_i con $i = 1, 2, 3$ mientras que para HSI solo modifíco s_3 que sería la intensidad.

Complemento de Color

Es análogo al negativo en imágenes en tonos de gris. Matices opuestos según el círculo de colores (se le suma 180 grados al canal H...tener en cuenta que se trata de un círculo) + el negativo de I (255 - I, suponiendo un canal de 8 bits).

Rebanado de Color (*color slicing*)

Resaltado de un rango de colores específico. Útil para:

- Mostrar colores de interés separados del fondo.
- Generar una máscara con la región del color de interés.

Mapeo de colores en un cubo de radio W centrado en a:

$$s_i = \begin{cases} 0,5 & \text{si } [|r_j - a_j| > \frac{W}{2}]_{1 \leq j \leq n}, i = 1, 2, \dots, n \\ r_i & \text{resto} \end{cases}$$

Mapeo de colores en una esfera de radio R_0 centrado en a:

$$s_i = \begin{cases} 0,5 & \text{si } \sum_{j=1}^n (r_j - a_j)^2 > R_0^2, i = 1, 2, \dots, n \text{ (distancia Euclídea)} \\ r_i & \text{resto} \end{cases}$$

Corrección de tono

Ajuste del brillo y el contraste. Modelo RGB y CMY(K): mapeo de las tres (cuatro) componentes con la misma función. Modelo HSI: modificación solamente de la componente I. Funciones de transformación: curvas lineales por tramos o polinomios que aumentan el contraste, o varían la intensidad en forma de potencia.

Balance de colores

Ajustes de información de canales (calibración). La percepción de un color es afectada por los colores vecinos. La rueda de colores es útil en la elección de colores a cambiar. Formas de ajuste: para incrementar un color podemos - Decrementar su complementario. - Aumentar la proporción de los dos colores contiguos en la rueda (similar a decrementar la proporción de los dos colores contiguos al complementario).

Manejo del histograma

Operaciones de ecualización (especificación) de valores de intensidad. En imágenes a todo color, el manejo de las componentes por separado conduce a resultados erróneos. Para uniformizar valores de intensidad sin cambiar los matices, se utiliza el modelo HSI.

Suavizado y acentuado

Extensión de las operaciones locales estudiadas para imágenes de intensidad. En el modelo RGB: las tres componentes se transforman mediante la aplicación convencional de una máscara. En el modelo HSI: aplicación de la máscara a la componente I. Resultados diferentes, dado que en RGB los colores se promedian, mientras que en HSI permanecen inalterados.

3. UNIDAD 3 - Operaciones en el Dominio Frecuencial

1. Introducción

La transformada de Fourier de una función discreta de una variable, $f(x)$, está dada por la ecuación:

$$F(u) = \frac{1}{M} \sum_{x=0}^{M-1} f(x) e^{-j\frac{2\pi ux}{M}}, \quad \text{para } u = 0, 1, \dots, M-1$$

$$f(x) = \sum_{u=0}^{M-1} F(u) e^{j\frac{2\pi ux}{M}}, \quad \text{para } x = 0, 1, \dots, M-1$$

Lo mismo ocurre para el caso bidimensional, teniendo $f(x, y)$, la transformada de Fourier esta dado por la ecuación:

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

$$\begin{cases} u = 0, 1, \dots, M-1 \\ v = 0, 1, \dots, N-1 \end{cases}$$

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

$$\begin{cases} x = 0, 1, \dots, M-1 \\ y = 0, 1, \dots, N-1 \end{cases}$$

Existe una relación entre el *paso espacial* y el *paso frecuencial*:

$$\Delta u = \frac{1}{M \Delta x} \quad \Delta v = \frac{1}{N \Delta y}$$

En coordenadas polares, se tiene:

$$F(u, v) = |F(u, v)| e^{j\phi(u, v)}$$

donde

$$\text{Magnitud (espectro): } |F(u, v)| = \sqrt{R^2(u, v) + I^2(u, v)}$$

$$\text{Fase: } \phi(u, v) = \tan^{-1} \left[\frac{I(u, v)}{R(u, v)} \right]$$

La base de Fourier para una imagen de $M \times N$ pixels, es un conjunto de $M \times N$ elementos de $M \times N$ pixels cada uno. Para dar una idea de esto, en la fig. (31) se muestran bases de Fourier para imágenes de 4×4 y 16×16 . Mientras que en la fig. (32) se muestra un ejemplo.

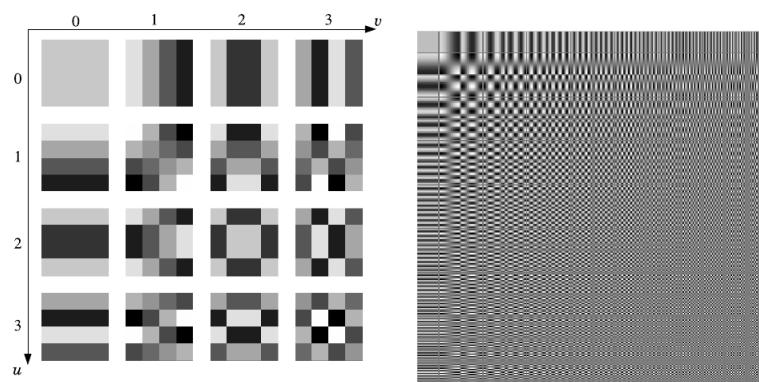


Figura 31: Base para imagen de 4×4 y 16×16 .

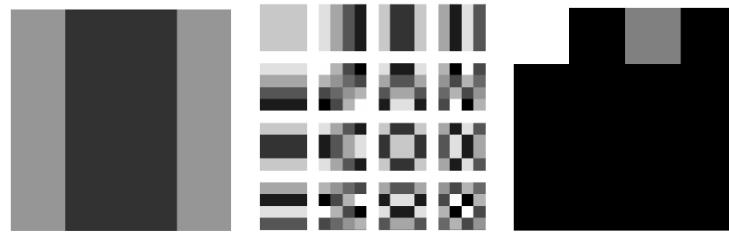


Figura 32: Ejemplo con imagen de 4×4 : imagen, imágenes base de 4×4 y coeficientes positivos de la DCT2.

1.1. Propiedades de la TFD

Traslación

$$\begin{aligned} f(x, y) e^{j2\pi(\frac{u_0x}{M} + \frac{v_0y}{N})} &\Leftrightarrow F(u - u_0, v - v_0) \\ f(x - x_0, y - y_0) &\Leftrightarrow F(u, v) e^{-j2\pi(\frac{u_0x}{M} + \frac{v_0y}{N})} \end{aligned}$$

cuando $u_0 = M/2$ y $v_0 = N/2$:

$$\begin{aligned} e^{j2\pi(\frac{u_0x}{M} + \frac{v_0y}{N})} &= e^{j\pi(x+y)} \\ &= (-1)^{(x+y)} \end{aligned}$$

Por lo tanto:

$$\begin{aligned} f(x, y)(-1)^{(x+y)} &\Leftrightarrow F(u - M/2, v - N/2) \\ f(x - M/2, v - N/2) &\Leftrightarrow F(u, v)(-1)^{(x+y)} \end{aligned}$$

Periodicidad $f(x, y)$ y $F(u, v)$ son funciones periódicas de período M y N :

$$\begin{aligned} F(u, v) &= F(u + M, v) = F(u, v + N) = F(u + M, v + N) \\ f(x, y) &= f(x + M, y) = f(x, y + N) = f(x + M, y + N) \end{aligned}$$

Simetría Conjugada Si $f(x, y)$ es real, entonces:

$$F(u, v) = F^*(-u, -v) \quad \text{y} \quad |F(u, v)| = |F(-u, -v)|$$

lo que dice que el espectro es simétrico respecto al origen.

Rotación En coordenadas polares:

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

Substituyendo:

$$\begin{aligned} f(r, \theta) &\Leftrightarrow F(\omega, \varphi) \\ f(r, \theta + \theta_0) &\Leftrightarrow F(\omega, \varphi + \varphi_0) \end{aligned}$$

Si $f(x, y)$ se gira un ángulo θ_0 , su transformada de Fourier $F(u, v)$ se gira la misma cantidad. Esto puede verse en la fig. (33).

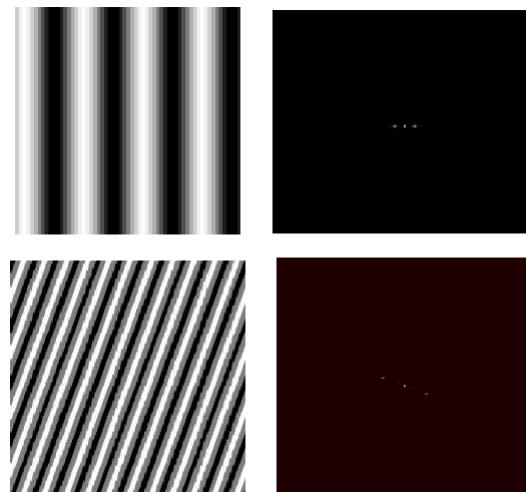


Figura 33: Ejemplo la propiedad de Rotación.

Convolución

$$\begin{aligned} g(x, y) &= f(x, y) * h(x, y) \\ g(x, y) &\Leftrightarrow G(u, v) \\ G(u, v) &= F(u, v)H(u, v) \\ f(x, y) * h(x, y) &\Leftrightarrow F(u, v)H(u, v) \end{aligned}$$

La TF de la salida de un sistema es igual a la multiplicación de la TF de la entrada al sistema por la TF de la respuesta del sistema al impulso unitario.

1.2. Representación Gráfica de la TF

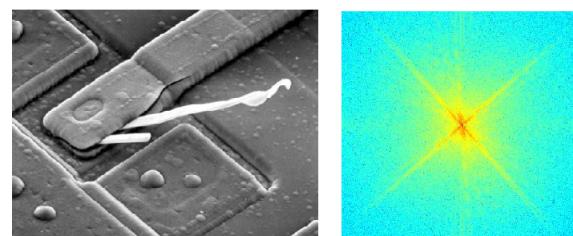


Figura 34: Ejemplo de la representación gráfica de la TF.

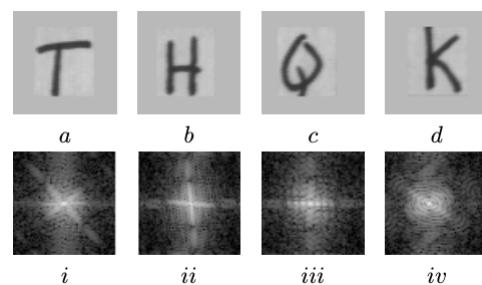


Figura 35: Ejercicio de parcial: a-ii, b-iii, c-iv y d-i.

1.3. Importancia de la Fase

Consecuencias del análisis de Fourier:

- La mayoría de las características más importantes de una señal se preservan sólo si la información de la fase se mantiene.
- Cuando una señal es de longitud finita, la información de la fase es suficiente para reconstruir una señal (podemos ver la imagen a grandes rasgos).

Imagen de sólo módulo:

- la TF tiene como módulo el de la imagen original, y fase nula.

Imagen de sólo fase:

- la TF tiene módulo unidad (o promediado) y fase igual a la de la imagen original.

Las características de la imagen original son identificables en una imagen de sólo fase, ya que la inteligibilidad está asociada a los detalles (puntos, bordes, etc.). En la imagen fig. (36) se ilustra el módulo y la fase de dos imágenes, mientras que en la fig. (37) se encuentra la representación de sólo módulo y sólo fase.

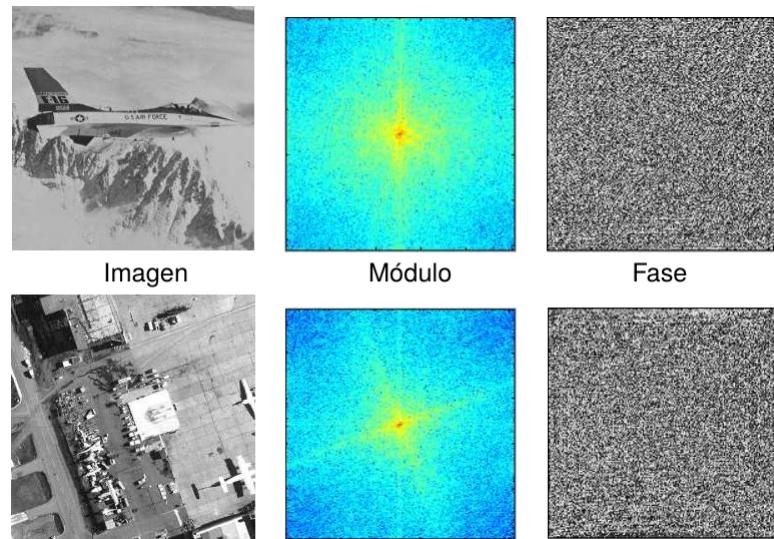


Figura 36: Módulo y Fase.

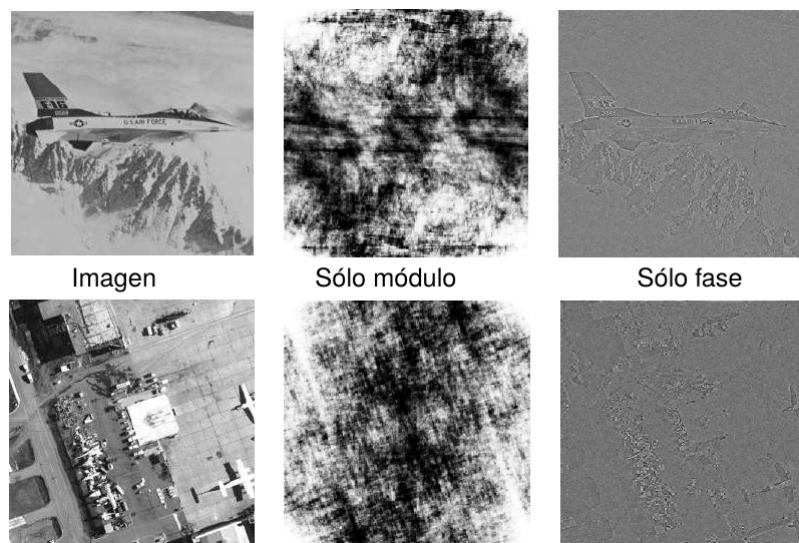


Figura 37: Sólo Módulo y sólo Fase.

Experimento de Oppenheim:

En la representación de Fourier de las señales, el módulo y la fase tienden a representar diferentes papeles y en algunas situaciones la mayoría de las características más importantes de una señal se preservan sólo si la información de la fase se mantiene. Además bajo una variedad de condiciones, cuando una señal es de longitud finita, la información de la fase simplemente basta para reconstruir una señal dentro de un factor de escala. Estas afirmaciones son válidas tanto para señales unidimensionales como multidimensionales.

Una imagen de sólo fase tiene una transformada de Fourier cuya fase es igual a la fase de la señal original y módulo unidad, o tal vez un módulo promediado entre un conjunto de imágenes. Muchas de las características de la imagen original son identificables claramente en la imagen de sólo fase, pero no ocurre lo mismo en la imagen de sólo módulo (aquella cuya transformada de Fourier tiene como módulo el de la imagen original, y fase nula).

Un experimento que ilustra claramente la observación de que las señales de sólo fase capturan mayor inteligibilidad de la señal que las de sólo módulo es el *Experimento de Oppenheim*, que consiste en reproducir una imagen combinando el módulo de una imagen fuente con la fase de otra imagen fuente y observando cuál información de las dos imágenes fuente predomina en la imagen resultante. Obviamente los pasos a seguir son los siguientes:

- Obtener el módulo de la transformada de Fourier de una de las imágenes fuente
- Obtener la fase de la transformada de Fourier de la otra imagen fuente.
- Hallar la transformada de Fourier inversa de la combinación del módulo con la fase obtenidos de los dos pasos anteriores.

El experimento se realizó con OrquideaJAI y se obtuvieron los resultados ilustrados en la fig. (38). En las dos primeras filas se encuentran las imágenes fuentes con sus módulos y fases de sus transformadas de Fourier. En la fila inferior a la izquierda se encuentra la imagen resultante de combinar el módulo de la TF de la flor y la fase de la TF de la mujer; y a su derecha se encuentra la imagen resultante del módulo de la TF de la mujer y la fase de la TF de la flor. Como se observa, en ambas imágenes resultantes predomina la información aportada por la fase. Es decir, la fase de la TF lleva la mayor parte de la información de la imagen.

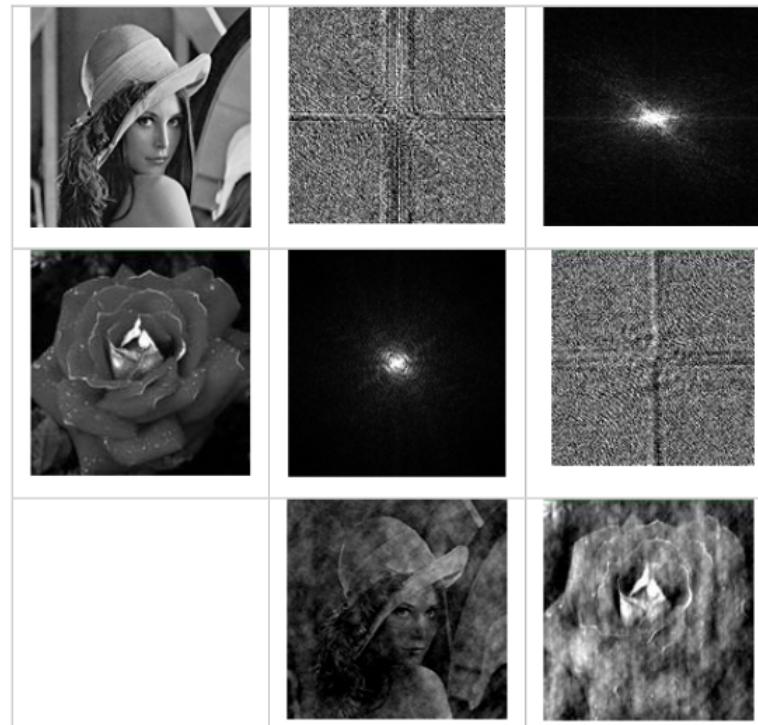


Figura 38: Experimento de Oppenheim.

2. Filtrado Frecuencial

- En el dominio espacial: la convolución puede ser computacionalmente ineficiente cuando se aumenta la talla del filtro.
- En el dominio frecuencial: utilización de la propiedad de correspondencia entre convolución espacial y producto TF.
- Proceso:
 - Generar una función $H(u, v)$ (función de transferencia del filtro).
 - Calcular la TF $F(u, v)$ de la imagen.
 - Multiplicar elemento a elemento las funciones.
 - Calcular la TF inversa.
 - Obtener la parte real.

2.1. Filtros de Suavizado

Filtro Pasa-Bajos Ideal:

$$H(u, v) = \begin{cases} 1 & \text{si } D(u, v) \leq D_0 \\ 0 & \text{si } D(u, v) > D_0 \end{cases}$$

donde: D_0 es la *frecuencia de corte*, y $D(u, v)$ es la distancia euclídea desde el punto (u, v) hasta el origen del plano de frecuencia. Los filtros ideales traen consigo un efecto de *Sobredisparo* por la discontinuidad de la función de transferencia. Esto genera lo que se conoce como **Fenómeno de Gibbs**.

El fenómeno de Gibbs presenta oscilaciones en la imagen filtrada originadas en el cálculo de la TF inversa. Esta aparición de oscilaciones se explica claramente en el dominio espacial. Para solucionar este problema, se multiplica el filtro ideal por una ventana suavizante (Hamming, Kaiser, etc.), o generar una función $H(u, v)$ con caída suave.

En las fig. (39), (40) y (41) se ilustra lo explicado mas arriba.

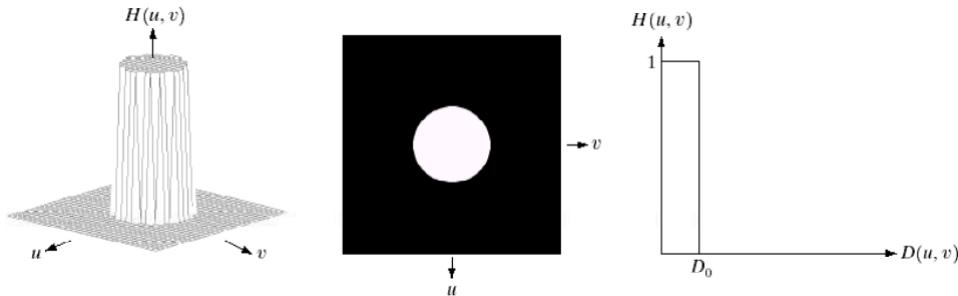


Figura 39: Filtro Pasa-Bajos Ideal.

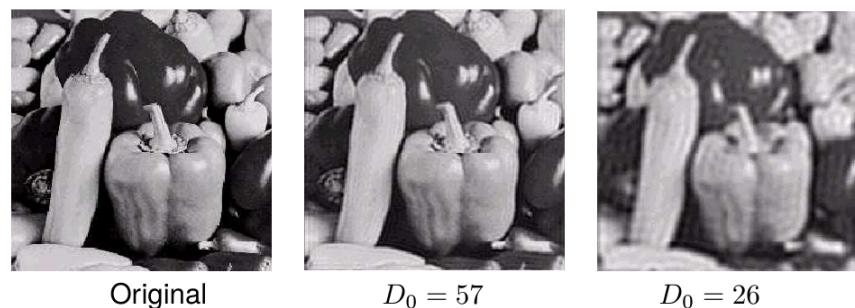


Figura 40: Ejemplo de Aplicación de filtro Pasa-Bajos Ideal.

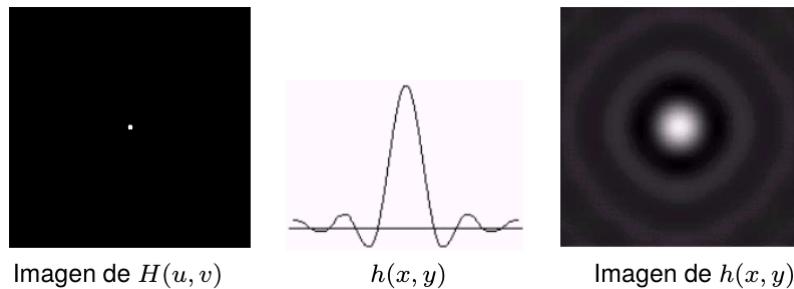


Figura 41: Oscilaciones del Fenómeno de Gibbs.

Filtro Pasa-Bajos Butterworth de orden n

$$H(u, v) = \frac{1}{1 + \left[\frac{D(u, v)}{D_0} \right]^{2n}}$$

Este filtro tiene una transición de corte suave, por lo que no genera el efecto de sobredisparo. Ver fig. (42).

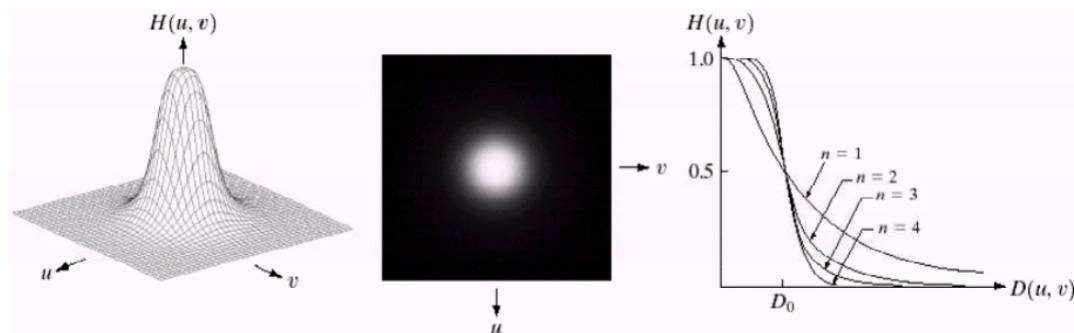


Figura 42: Filtro Pasa-Bajos de Butterworth.

Filtro Pasa-Bajos Gaussiano

$$H(u, v) = e^{-D^2(u, v)/2\sigma^2}$$

σ es la medida de dispersión de la curva gaussiana. Cuando $D(u, v) = \sigma$, la función de transferencia está a 0,607 de su valor máximo. Por otro lado, existe una propiedad importante, la TF inversa de un filtro gaussiano también es gaussiana. Esto tiene la ventaja de que no tiene sobredisparo. La definición del filtro gaussiano se puede ver en la fig. (43).

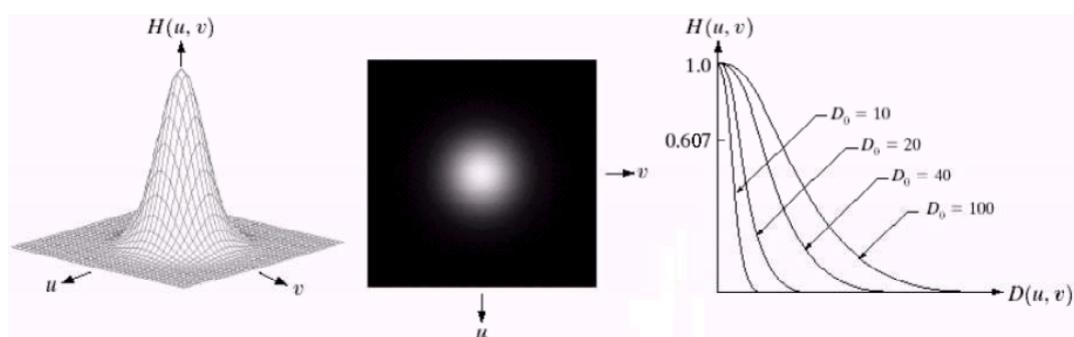


Figura 43: Filtro Pasa-Bajos Gaussiano.

2.2. Filtros de Acentuado

En el filtrado pasa-altos, el brillo medio es anulado. El resultado “visualmente” no es óptimo. Solución: trabajar sobre la imagen original, agregando componentes de alta frecuencia.

Filtro Pasa-Altos Ideal

$$H(u, v) = \begin{cases} 0 & \text{si } D(u, v) \leq D_0 \\ 1 & \text{si } D(u, v) > D_0 \end{cases}$$

donde D_0 es la frecuencia de corte, y $D(u, v)$ es la distancia euclídea desde el punto de vista (u, v) hasta el origen del plano de frecuencia. Su definición puede verse en la fig. (44).

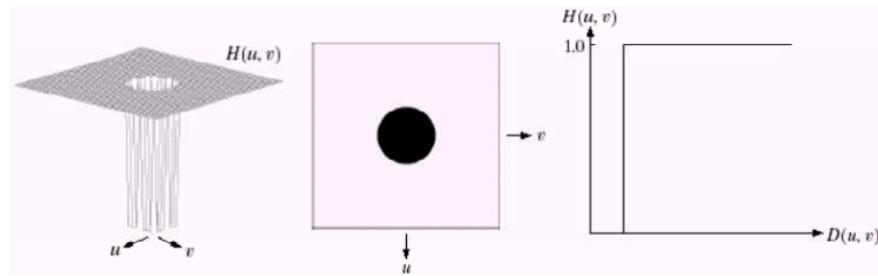


Figura 44: Filtro Pasa-Altos Ideal.

En la fig. (45) se puede ver un ejemplo de la aplicación del filtro pasa-altos ideal. Al igual que en el filtro pasa-bajos, se observa la aparición del fenómeno de Gibbs.

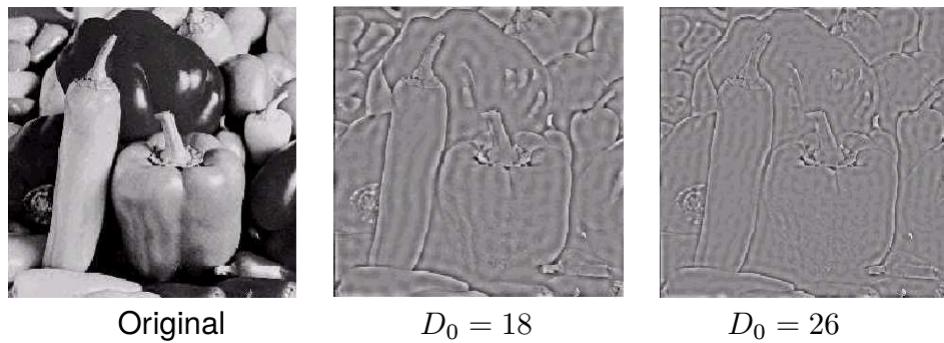


Figura 45: Ejemplo de aplicación de filtro Pasa-Altos Ideal.

Filtro Pasa-Altos de Butterworth de orden n

$$H(u, v) = \frac{1}{1 + \left[\frac{D_0}{D(u, v)} \right]^{2n}}$$

Al igual que el filtro PB, para orden bajo no introduce sobredisparo ($n < 20$). En la fig. (46) se puede ver la definición del filtro mientras que en la fig. (47) se hace una comparación entre el filtro ideal y el Butterworth.

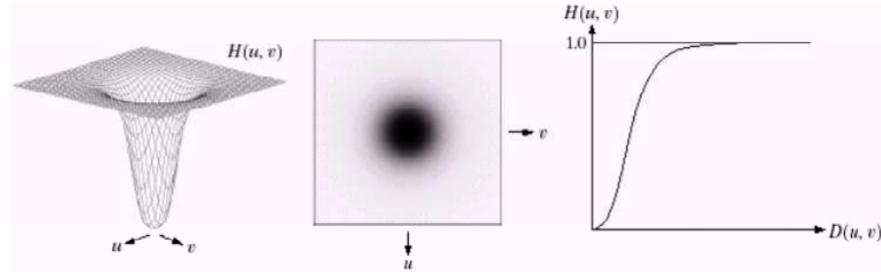


Figura 46: Filtro Pasa-Altos Butterworth.

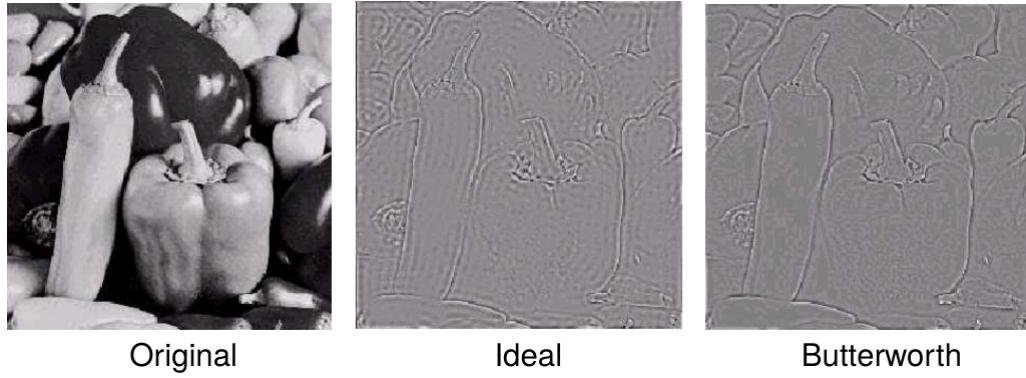


Figura 47: Comparación filtro Ideal y filtro Butterworth.

Filtro Pasa-Altos Gaussiano

$$H(u, v) = 1 - e^{-D^2(u, v)/2\sigma^2}$$

siendo σ^2 la varianza de la curva gaussiana ($\sigma = D_0$). La definición de este filtro se puede ver en la fig. (48).

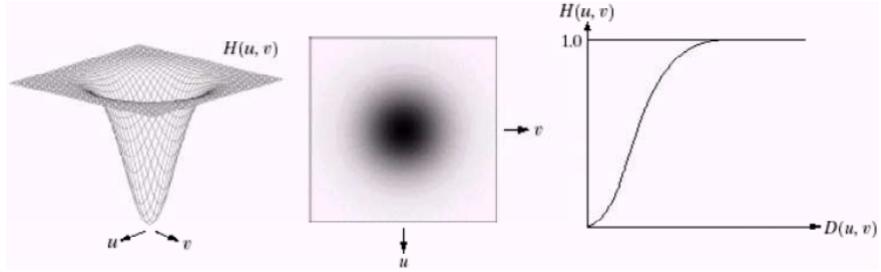


Figura 48: Filtro Pasa-Altos Gaussiano.

2.3. Filtrado de Mascara Difusa

Operaciones:

en el dominio espacial: $f_{PA}(x, y) = f(x, y) - f_{PB}(x, y)$

en el dominio frecuencial: $F_{PA}(u, v) = F(u, v) - F_{PB}(u, v)$

al ser: $F_{PB}(u, v) = H_{PB}(u, v)F(u, v)$

entonces: $F_{PA}(u, v) = F(u, v) - H_{PB}(u, v)F(u, v)$

despejando: $= F(u, v)(1 - H_{PB}(u, v))$

el filtro es: $H_{PA}(u, v) = 1 - H_{PB}(u, v)$

El resultado que se obtiene es visualmente similar al filtro pasa-altos. La solución es amplificar el aporte de la imagen original.

2.4. Filtrado de Alta Potencia (high-boost)

Generalización de la máscara difusa ($A \geq 1$):

$$f_{AP}(x, y) = Af(x, y) - f_{PB}(x, y)$$

vimos que puede ser reescrita como:

$$f_{AP}(x, y) = (A - 1)f(x, y) + f_{PA}(x, y)$$

en el dominio frecuencial:

$$H_{AP}(u, v) = (A - 1) + H_{PA}(u, v)$$

2.5. Filtrado de Énfasis de Alta Frecuencia

Para aumentar el aporte de los componentes de alta frecuencia a la imagen, se multiplica por una constante al filtro PA y se modifica el *offset* para no eliminar el brillo medio:

$$H_{EAF}(u, v) = a + bH_{PA}(u, v)$$

con: $a \geq 0$ y $b > a$

Casos particulares:

- Si $a = (A - 1)$ y $b = 1$: $H_{EAF} = H_{AP}$
- Si $a = 0$ y $b = 1$: $H_{EAF} = H_{PA}$

En la fig. (49) se comparan los filtros *high-boost* y Énfasis de Alta Frecuencia.

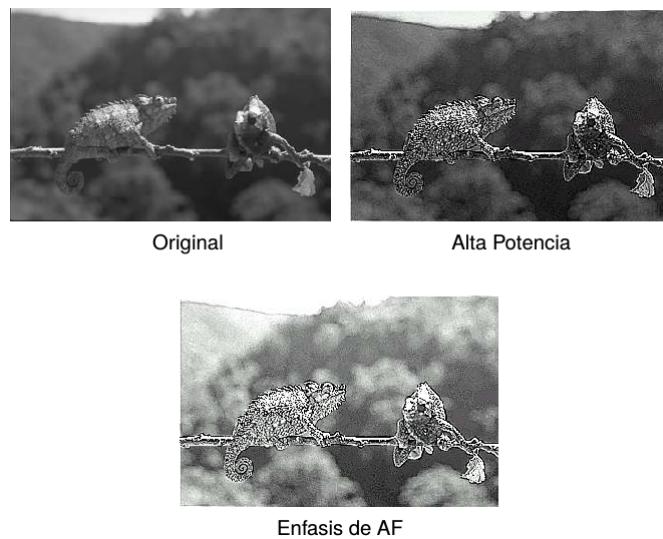


Figura 49: Comparación de filtros high-boost y EAF.

2.6. Filtrado Homomórfico

Las imágenes digitales que procesamos normalmente se crean a partir de imágenes ópticas, las cuales constan de dos componentes:

- **Iluminación $i(x, y)$:** condiciones de luz (cambiante).
- **Reflectancia $r(x, y)$:** propiedad intrínseca de los objetos (fija).

En muchas aplicaciones interesa realizar la componente de reflectancia mientras se reduce la componente de iluminación. El filtrado PA y PB por sí solos no logran ambos cometidos a la vez, por lo que generalmente se utilizan con un procesamiento posterior. El filtrado homomórfico es un proceso que:

- comprime el rango dinámico a partir de las condiciones de iluminación, y
- realza el contraste a partir de las propiedades de reflectancia de los objetos.

Recordar que una imagen puede expresarse como:

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

La ecuación anterior no se puede usar directamente porque la transformada de Fourier de un producto de funciones no es separable:

$$F\{f(x, y)\} \neq F\{i(x, y)\}F\{r(x, y)\}$$

Haciendo:

$$z(x, y) = \ln(f(x, y)) = \ln(i(x, y)) + \ln(r(x, y))$$

entonces

$$F\{z(x, y)\} = F\{\ln(f(x, y))\} = F\{\ln(i(x, y))\} + F\{\ln(r(x, y))\}$$

o bien:

$$Z(u, v) = I(u, v) + R(u, v)$$

Al procesar con un filtro a $Z(u, v)$ se tiene:

$$S(u, v) = H(u, v)Z(u, v) = H(u, v)I(u, v) + H(u, v)R(u, v)$$

En el dominio espacial:

$$\begin{aligned} s(x, y) &= F^{-1}\{S(u, v)\} = F^{-1}\{H(u, v)I(u, v)\} + F^{-1}\{H(u, v)R(u, v)\} \\ &= i'(x, y) + r'(x, y) \end{aligned}$$

La imagen resultante es:

$$\begin{aligned} g(x, y) &= e^{s(x, y)} \\ &= e^{i'(x, y)}e^{r'(x, y)} \\ &= i_0(x, y)r_0(x, y) \end{aligned}$$

Entonces, la secuencia de operaciones es:

$$f(x, y) \rightarrow \langle \ln \rangle \rightarrow \langle F \rangle \rightarrow \langle H(u, v) \rangle \rightarrow \langle F^{-1} \rangle \rightarrow \langle \exp \rangle \rightarrow g(x, y)$$

Clave: definir un filtro homomórfico $H(u, v)$ que actúe sobre las componentes de iluminación y reflectancia de forma separada.

Conceptos:

- La iluminación se caracteriza por variaciones espaciales pequeñas → bajas frecuencias de la TF del logaritmo de la imagen.
- La reflectancia tiende a variar abruptamente en los bordes → altas frecuencias de la TF del logaritmo de la imagen.

El filtro Homomórfico $H(u, v)$, entonces, debe afectar de manera diferente a las bajas y altas frecuencias.

Especificación de $H(u, v)$, ver fig. (50):

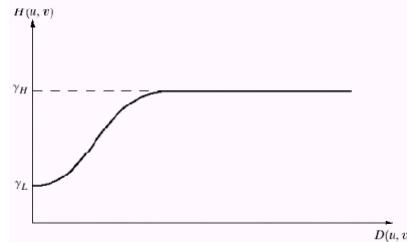


Figura 50: Especificación de $H(u, v)$.

Elección de parámetros:

- $\gamma_L < 1$: decrementa el brillo
- $\gamma_H > 1$: amplifica las altas frecuencias

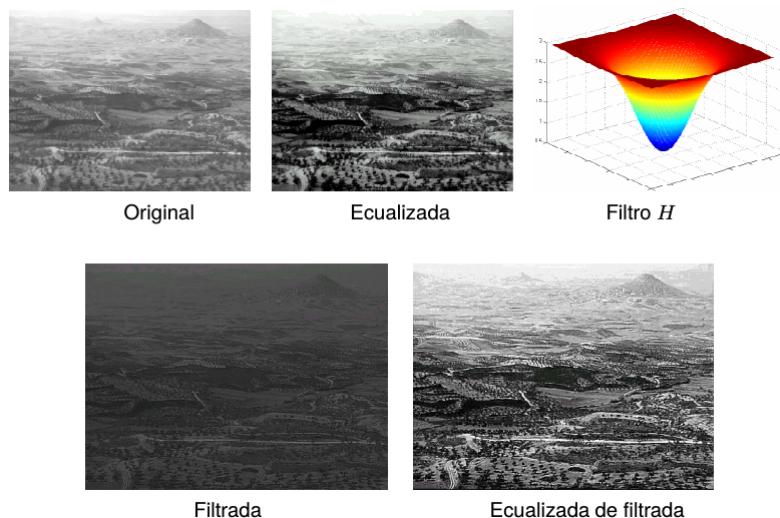


Figura 51: Ejemplo de filtrado Homomórfico.