

Universidad  
Rey Juan Carlos

ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA INFORMÁTICA

MÁSTER UNIVERSITARIO DE VISIÓN ARTIFICIAL

## Práctica evaluable

TRATAMIENTO DIGITAL DE IMÁGENES

CURSO 2023-2024

Adrián Riaño Martínez

# Índice

<b>Práctica 1: Análisis en el dominio de la frecuencia</b>	<b>3</b>
Apartado A. Filtro alto de Butterworth . . . . .	3
Apartado B. Compresión DCT 2D . . . . .	4
<b>Práctica 2: Segmentación basada en regiones</b>	<b>6</b>
Apartado A. Mapa de disparidad a través de la técnica de ajustes . . . .	6

# Práctica 1: Análisis en el dominio de la frecuencia

## Apartado A. Filtro alto de Butterworth

Un filtro de paso alto es una operación de transformación (dominio de las frecuencias) que excluye frecuencias espaciales bajas. La aplicación de este filtro genera una nueva imagen que conserva los bordes de la imagen original mientras se deshace de las regiones de intensidad altas y bajas. El resultado de este proceso, por orden de aparición, se encuentra en la figura 1: imagen original, imagen con el filtro del paso alto en el espacio de Fourier e imagen filtrada pero en el espacio visible normal.

La literatura afirma que el resultado de aplicar un filtro de paso alto genera una imagen donde los componentes de baja frecuencia se sitúan sobre el punto central  $G(0,0)$ . Esto se aprecia sobre la imagen *Highpass filtered* en 1. Aquí, se crea un radio donde dentro de éste, los valores de la transformada se hacen nulos, y fuera, se mantienen con los valores originales. En cuanto a la imagen *Filtered ifft2*, que es el resultado en el espacio visible, es una figura donde se filtra todo menos los bordes. Para este proceso, se han utilizados los siguientes parámetros: una frecuencia de corte de 0.5 y una orden de magnitud igual a 4.

Por otra parte, se ha calculado los diferentes histogramas 2 de las imágenes 1. De manera general, se puede ver que existe una variación en la distribución de intensidades según el filtrado y el espacio donde se trabaje. La imagen original, convertida a una escala de grises, tiene un rango de 0 a 255 donde los blancos y los negros destacan por encima del resto. Asimismo, se puede apreciar que hay variación general de los valores de los píxeles. En cambio, en la imagen *Highpass filtered*, se puede ver una uniformidad en todo el rango de valores salvo para intensidades nulas o máximas. Por último, en el histograma *Filtered ifft2*, se aprecia que la imagen es binaria, ya que los valores de los píxeles recaen o sobre 0, o sobre 255. Asimismo, destaca cómo la cantidad de píxeles en blanco es ínfima con respecto a la cantidad de píxeles en negro. Esto se debe a que en el filtro de paso alto mantiene los bordes originales representados en color blanco y todo lo demás lo anula.

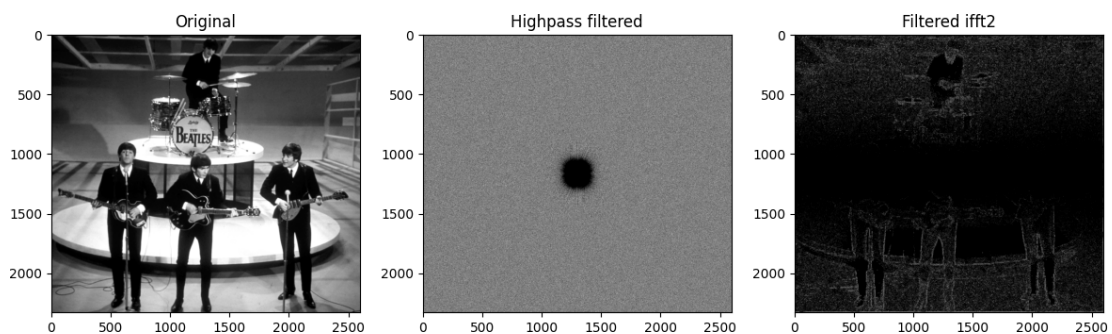


Figura 1: Imágenes del proceso del filtro de paso alto de Butterworth

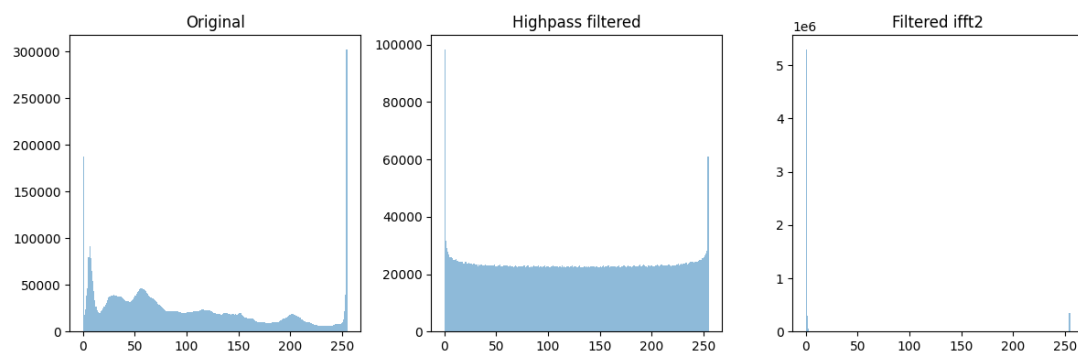
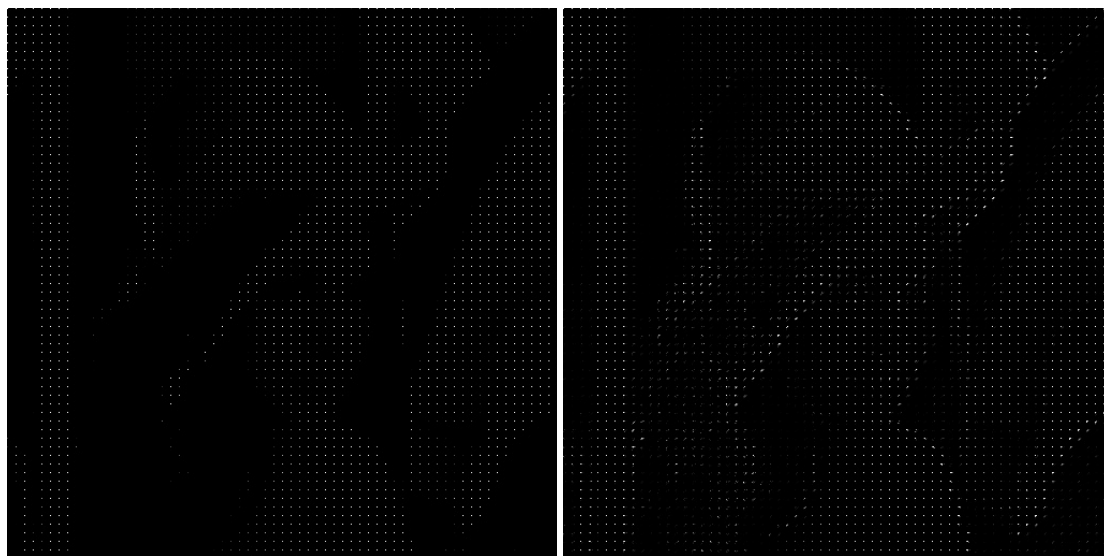


Figura 2: Histogramas del proceso del filtro de paso alto de Butterworth

## Apartado B. Compresión DCT 2D

La Transformada Discreta de Coseno (DCT, por sus siglas en inglés) se trata de una alternativa a la transformada de Fourier que destaca sobre esta última porque el proceso de compresión de energía es más eficiente. Esto se debe, según indica la literatura, porque la mayor parte de la información de la señal tiende a concentrarse en unos pocos componentes de baja frecuencia de la DCT.



(a) Factor de compresión 1/64

(b) Factor de compresión 10/64

Figura 3: Compresión por DCT 2D

El objetivo de este apartado B consiste en la compresión de imágenes en escala de grises aplicando DCT. Para ello, se ha tomado una nueva foto y se ha llevado a cabo el proceso de compresión y descompresión. En la figura 3, se muestra el proceso de compresión por DCT. Si se comparan los resultados, lo primero que salta a la vista es que los bordes de la imagen con un factor de compresión 10/64 se encuentran más marcados y luminosos. Esto tiene sentido porque con un factor de 1/64, se especifica que por cada bloque 8x8 se debe guardar el primer pixel

(en su defecto un coeficiente DCT) de ese bloque. Por tanto, por cada bloque donde se calcula DCT, la cantidad de información perdida es tan alta que en el proceso de descompresión, los valores finales no coincidirán con los iniciales. De manera más técnica, el resultado DCT presenta tantos valores nulos en la imagen que solo se puede recuperar los valores correspondientes a las esquinas definidas en la matriz de compresión.



(a) Factor de compresión 1/64

(b) Factor de compresión 10/64

Figura 4: Descompresión por IDCT 2D

Por otra parte, si comparamos las imágenes fruto de la descompresión 4, se distingue que aquella que se ha aplicado un factor alto de compresión, la imagen está pixelada, mientras que aquella con un factor 10/64, la imagen pierde cierta calidad aunque a simple vista es inapreciable. Con este experimento, también se demuestra que con factores de compresiones altos, se introduce ruido en la imagen. Este permite la formación de perturbaciones sobre ella. No obstante, si la imagen estuviese en una escala RGB, también se apreciaría aberraciones cromáticas.

## Práctica 2: Segmentación basada en regiones

### Apartado A. Mapa de disparidad a través de la técnica de ajustes

Los mapas de disparidad se tratan de imágenes que reflejan la diferencia de posición de los objetos dado dos imágenes estéreas. El objetivo de este apartado consiste en obtener un mapa de disparidad a partir de la técnica de ajustes o *blockmatching*. Dado un tamaño de kernel de 11x11 y un desplazamiento de 50 píxeles máximo, se posiciona el kernel sobre las mismas coordenadas en ambas imágenes. A continuación, se va moviendo el kernel de la imagen derecha posiciones atrás tantas veces como se haya especificado en el desplazamiento. A la vez, por cada desplazamiento, se va calculando la suma de diferencias al cuadrado (SSD por sus siglas en inglés) (1). Una vez alcanzado el máximo de desplazamientos, se obtiene el menor valor SSD y se asigna al trozo del mapa de disparidad el valor del desplazamiento correspondiente al SSD obtenido.

$$\text{SSD}(x, y, d) = 1/1 + \sum_{i=1}^{\text{rows}} \sum_{j=1}^{\text{cols}} (I_L(x + i, y + j) - I_R(x + i - d, y + j))^2 \quad (1)$$



(a) Imagen estérea izquierda

(b) Imagen estérea derecha

Figura 5: imágenes estéreas

Para la puesta en escena, se ha elegido las imágenes estéreas (5). A partir de la teoría explicada anteriormente, se codifica y se obtiene el siguiente mapa de disparidad 6a. Lo primero que hay que tener en cuenta son las escalas de colores, los colores que tienden a blanco significa que la disparidad más baja se encuentra a muchos desplazamientos, mientras que escalas más oscuras indican lo contrario. Se puede apreciar como el mapa de disparidad distingue las formas correspondientes a los objetos. También se puede apreciar la existencia de patrones de diferente color dentro de los objetos o formas. Esto puede ser porque las posiciones de

disparidad varían levemente. En un segundo experimento, se ha intercambiado las imágenes, en vez de tomar la imagen derecha como imagen derecha, se ha asignado la imagen izquierda y al revés con la imagen izquierda. El resultado se puede ver en (6b). Se obtiene un mapa de disparidad formado por manchas. La explicación a este fenómeno podría ser que debido a que los objetos en la toma derecha suelen encontrarse posiciones más a la izquierda, hay una disonancia y se necesita o más desplazamientos o un barrido hacia adelante. Este hecho hace imposible distinguir formas.

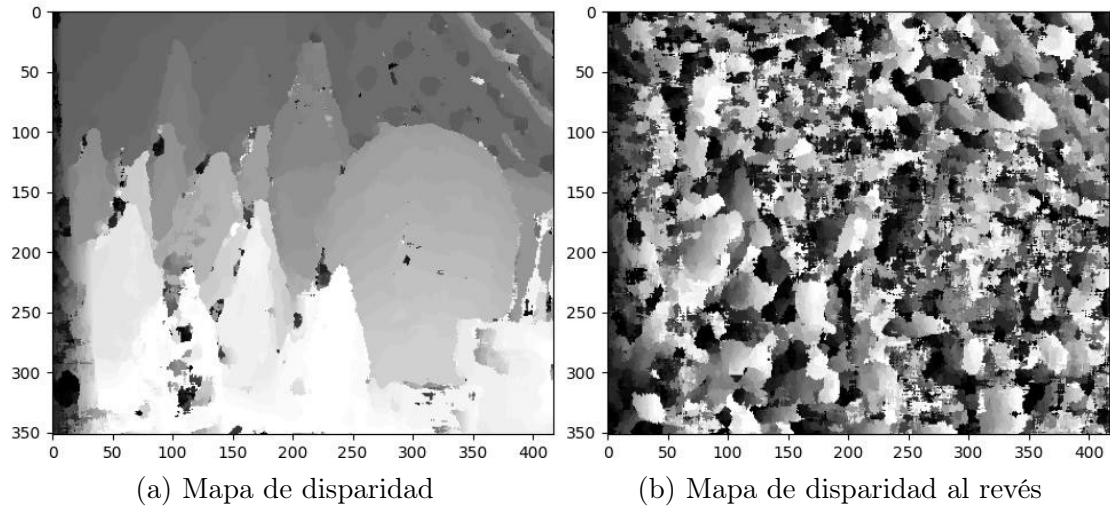


Figura 6: imágenes estéreas