



# PRÁCTICA 5

## TRATAMIENTO DE IMÁGENES

Universidad Pontifical de  
Comillas ICAI

Andrés Sánchez de Ágreda – 202108563  
Servicios Telemáticos Multimedia



## ÍNDICE

<b>1. Introducción.....</b>	<b>2</b>
<b>2. Análisis de imagen.....</b>	<b>2</b>
2.1. Elegir una imagen en formato no comprimido (por ejemplo, BMP), representarla y si es necesario pasarl a escala de grises. Mostrar resultados.....	2
2.2. Analizar el histograma de la imagen en escala de grises y ajustar la escala si fuese necesario. Mostrar resultados.....	3
2.3. Analizar el deterioro de la imagen original usando diferente número de niveles de gris. Comentar lo observado.....	5
<b>3. Análisis Huffman.....</b>	<b>6</b>
3.1. Usar a imagen original con una escala de grises de sólo 8 valores y obtener la tabla de codificación de Huffman. .....	6
3.2. Codificar la imagen en binario según la tabla obtenida. ....	7
3.3. Decodificar la imagen según la tabla y árbol de Huffman antes obtenidos. ....	8
3.4. Analizar la ratio de compresión respecto a la imagen original con escala de grises de 8 valores. ....	9
3.5. Analizar la redundancia de datos para el caso de 8 niveles de gris. ....	9
<b>4. Análisis de Transformada Discreta del Coseno .....</b>	<b>10</b>
4.1. Realizar la transformada discreta del coseno para la imagen codificada antes con Huffman de 8 niveles de gris. La ventana de la transformada será de tamaño 8x8 .....	10
4.2. Analizar la ratio de compresión y la redundancia de datos.....	12
4.3. Recuperar la imagen de la transformada discreta del coseno y comparar el resultado con la imagen original codificada. Probar con diferentes máscaras de valores significativos de la DCT y cuantificar las diferencias. ....	12
<b>5. Análisis con otras transformadas.....</b>	<b>14</b>
5.1. Analizar y comparar lo que ocurre al hacer una transformada de Fourier y comparar con los resultados obtenidos con la transformada del coseno .....	14
5.2. Analizar y comparar lo que ocurre al hacer una transformada tipo wavelet y comparar con los resultados obtenidos con la transformada del coseno .....	15
5.3. Analizar y comparar lo que ocurre al hacer una transformada tipo KLT y comparar con los resultados obtenidos con la transformada del coseno .....	16
<b>6. Conclusión.....</b>	<b>17</b>
<b>7. Anexo Código.....</b>	<b>18</b>



## 1. Introducción

El procesamiento digital de imágenes es fundamental en numerosos ámbitos tecnológicos, desde la fotografía digital hasta la compresión multimedia y la transmisión de información. Esta práctica tiene como objetivo comprender cómo se representa una imagen, cómo se transforma su información y de qué manera puede reducirse su tamaño manteniendo la mayor calidad posible.

En primer lugar, se analiza la imagen original, su histograma y los efectos de limitar el número de niveles de gris, lo que permite estudiar la degradación asociada a la cuantización. Posteriormente, se aplica la codificación de Huffman para explotar la redundancia estadística de la imagen y reducir el número medio de bits por píxel sin pérdida de información.

A continuación, se examinan distintas transformadas aplicadas a la compresión con pérdida: la DCT, la FFT, la transformada Wavelet y la KLT. Mediante la conservación selectiva de coeficientes y la reconstrucción de la imagen, se evalúa la capacidad de cada método para concentrar la energía y su rendimiento en términos de error y calidad visual.

En conjunto, la práctica permite observar cómo las distintas técnicas —cuantización, codificación entropía y transformadas— explotan diferentes tipos de redundancia presentes en las imágenes, proporcionando una visión global de los principios que sustentan los sistemas modernos de compresión.

## 2. Análisis de imagen

### 2.1. Elegir una imagen en formato no comprimido (por ejemplo, BMP), representarla y si es necesario pasarl a escala de grises. Mostrar resultados.

El primer paso de la práctica consiste en seleccionar una imagen en formato sin compresión, con el objetivo de analizar sus características básicas y preparar los datos para etapas posteriores de procesamiento. Se ha optado por una imagen en formato BMP, ya que este formato almacena la información completa de color sin aplicar compresión con pérdida, conservando así toda la información de intensidad y detalle. Este tipo de archivo es especialmente adecuado para análisis de tipo radiométrico o morfológico, en los que se requiere trabajar con valores de píxel precisos.



Ilustración 1: Imagen original a color

La imagen elegida corresponde a un monoplaza de competición de la Universidad Pontificia Comillas (Equipo Formula Student). En su formato original presenta una resolución de 1600 ×



1143 píxeles, distribuidos en tres canales (RGB). Su tamaño en disco es de 5357.95 KB, lo que refleja la gran cantidad de información que contiene cada canal de color.

En la Ilustración 1 se muestra la imagen original en color. Se aprecia un alto nivel de detalle y contraste, con una amplia gama cromática que permite diferenciar con claridad los elementos del vehículo, la pista y el fondo. Esta representación es adecuada para análisis visuales, pero no resulta óptima para cálculos de intensidad, ya que cada píxel está compuesto por tres valores.

Posteriormente, la imagen se ha convertido a escala de grises mediante la función `rgb2gray` de MATLAB. Esta función realiza una combinación ponderada de los tres canales de color según la sensibilidad del ojo humano a las distintas longitudes de onda, generando una única matriz de intensidades. El resultado es una imagen de idénticas dimensiones ( $1600 \times 1143$  píxeles) pero con un solo canal, lo que reduce el tamaño a 1786.99 KB, aproximadamente una tercera parte del original.



Ilustración 2: Imagen original en escala de grises

En la Ilustración 2 se observa que la información estructural y de luminancia se mantiene prácticamente intacta. Los contrastes entre las zonas claras y oscuras del coche, así como los reflejos y las sombras, continúan siendo visibles. Sin embargo, la eliminación de la información cromática provoca una reducción perceptual en la riqueza visual, aunque sin afectar la información espacial ni los detalles esenciales.

La conversión de la imagen a escala de grises ha permitido reducir significativamente el tamaño del archivo sin pérdida apreciable de información estructural. Esta reducción se debe a la eliminación de los canales de color, que contenían información redundante para el análisis de intensidad. El resultado obtenido constituye una base adecuada para los apartados siguientes del análisis, en los que se estudiará la distribución de niveles de gris, el histograma y la cuantización de la imagen.

## 2.2. Analizar el histograma de la imagen en escala de grises y ajustar la escala si fuese necesario. Mostrar resultados.

Tras la conversión a escala de grises realizada en el apartado anterior, el siguiente paso consiste en analizar la distribución de intensidades de la imagen, representada mediante su histograma.



Este análisis permite evaluar el rango dinámico de los niveles de gris presentes, detectar posibles zonas sobreexpuestas o subexpuestas y decidir si es necesario realizar un ajuste de contraste.

El histograma muestra el número de píxeles que presentan cada nivel de intensidad entre 0 (negro absoluto) y 255 (blanco puro), proporcionando una visión cuantitativa del brillo y del contraste global de la imagen.

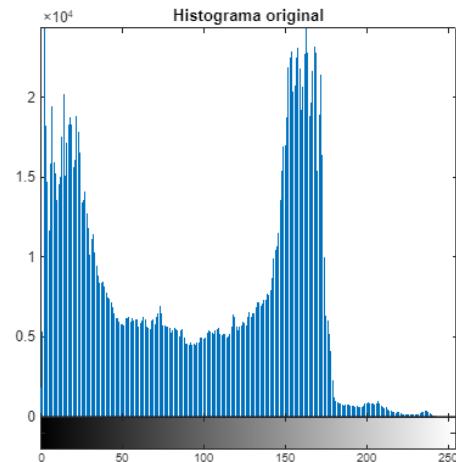


Ilustración 3: Histograma de la distribución de grises

La Ilustración 3 muestra la imagen en escala de grises junto a su histograma correspondiente. Se observa una distribución bimodal, con dos picos principales: uno concentrado en los valores bajos (zonas oscuras del vehículo y las sombras) y otro en los valores altos (zonas claras del fondo y los reflejos del chasis). Entre ambos grupos existe un valle que representa las regiones intermedias de luminancia.

El análisis numérico arroja un valor mínimo de gris de 0 y un valor máximo de 254, lo que indica que la imagen ya ocupa prácticamente todo el rango dinámico disponible, aunque no de forma uniforme. Esto sugiere que existen áreas con alto contraste local, pero también regiones de luminosidad concentrada, lo que puede limitar la percepción de detalle en determinadas zonas.



Ilustración 4: Comparación entre imágenes de grises

Para mejorar la distribución tonal y resaltar los detalles, se ha aplicado un ajuste automático de contraste utilizando la función imadjust. Esta función redistribuye linealmente los niveles de intensidad, extendiendo los valores más frecuentes hacia los extremos del rango [0, 255]. El resultado puede apreciarse en la Ilustración 4, donde se comparan la imagen original en escala de grises y la versión con contraste ajustado. En la segunda, las zonas oscuras del monoplaza y



los reflejos metálicos del chasis presentan una mayor diferenciación, lo que mejora la percepción de profundidad y la definición de contornos.

El histograma de la imagen original revela un uso amplio pero irregular del rango tonal, con predominio de intensidades bajas y altas. Mediante el ajuste de contraste se logra una redistribución más uniforme de los niveles de gris, lo que incrementa la visibilidad de detalles sin introducir artefactos ni pérdida de información.

Este proceso permite obtener una representación más equilibrada en términos de luminosidad y contraste, mejorando la calidad perceptual y preparando la imagen para los análisis de cuantización y compresión que se desarrollarán en los siguientes apartados.

### 2.3. Analizar el deterioro de la imagen original usando diferente número de niveles de gris. Comentar lo observado.

Una de las formas más directas de evaluar la calidad radiométrica de una imagen digital es analizar cómo se ve afectada cuando se reduce el número de niveles de gris disponibles para representar la intensidad de cada píxel. Este proceso, denominado cuantización de niveles de gris, consiste en limitar el rango de valores posibles, sustituyendo los 256 niveles originales (de 0 a 255) por un número menor de niveles equiespaciados.

La reducción de niveles implica una pérdida de información tonal, que puede manifestarse visualmente como *banding* o aparición de zonas planas con saltos bruscos de intensidad entre regiones.

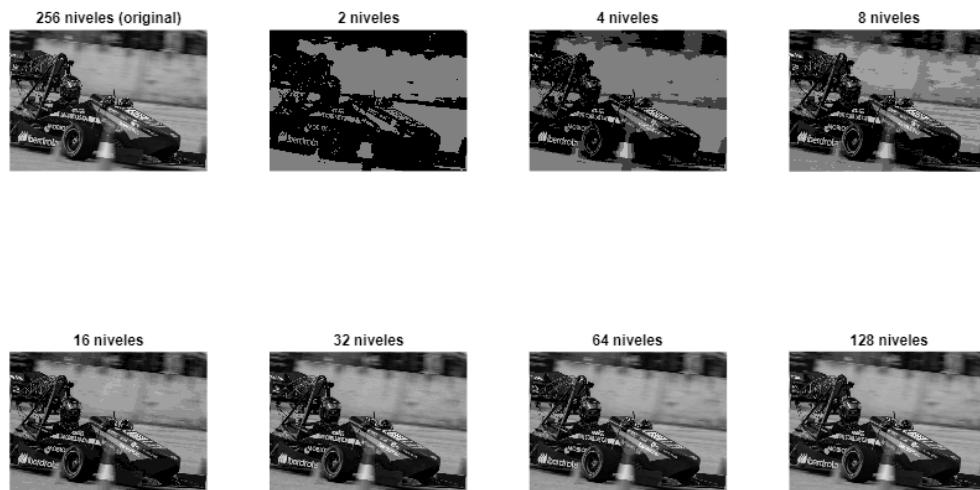


Ilustración 5: Comparación de niveles de gris

En la Ilustración 5 se muestra la comparación entre la imagen original (256 niveles) y sus versiones cuantizadas a 128, 64, 32, 16, 8, 4 y 2 niveles. Todas las versiones se han generado aplicando una cuantización uniforme, de manera que el rango total de intensidades se divide en intervalos equidistantes y cada píxel se asigna al nivel más próximo.

En la imagen original (256 niveles) la transición entre las zonas iluminadas y sombreadas es suave y continua.

Al reducir a 128 o 64 niveles, la degradación es prácticamente imperceptible a simple vista; los matices y las texturas se mantienen con fidelidad. A partir de 32 niveles, comienzan a observarse ligeras pérdidas de suavidad en los degradados, especialmente en las zonas del fondo y en los reflejos metálicos. Con 16 niveles, la imagen presenta ya una apariencia más artificial, con transiciones más abruptas entre áreas de diferente intensidad.



Cuando la cuantización se reduce a 8, 4 o 2 niveles, el efecto de posterización se vuelve evidente: se pierden los detalles finos, las sombras se agrupan en grandes manchas y las regiones de igual nivel se distinguen con bordes marcados.

La reducción progresiva de los niveles de gris provoca un deterioro visible en la calidad radiométrica de la imagen. Aunque los resultados con 64 o 128 niveles son visualmente aceptables, por debajo de 16 niveles el error de cuantización se hace evidente, generando una pérdida notable de detalle y contraste local.

Este experimento demuestra la importancia de mantener un número suficiente de niveles de gris para representar con precisión la información de intensidad, especialmente en imágenes con gradientes suaves o zonas con alta variabilidad tonal.

### 3. Análisis Huffman

#### 3.1. Usar a imagen original con una escala de grises de sólo 8 valores y obtener la tabla de codificación de Huffman.

Una vez cuantizada la imagen a un número reducido de niveles de gris, es posible aplicar técnicas de compresión sin pérdida que aprovechan la redundancia estadística de los datos. Entre ellas, la codificación de Huffman es una de las más utilizadas, ya que asigna códigos binarios más cortos a los símbolos más frecuentes y códigos más largos a los menos comunes, optimizando así el número total de bits empleados para representar la imagen.

En este apartado se ha utilizado la imagen con 8 niveles de gris, lo que equivale a 3 bits por píxel en la representación original. El objetivo es generar una tabla de codificación óptima de Huffman a partir de la distribución de probabilidades de los niveles de gris presentes en la imagen. Se usarán en (c)–(f) para comparar tamaños/ratios, calidad objetiva, y evaluación auditiva.



Ilustración 6: Imagen con 8 niveles de gris

La Ilustración 8 muestra la imagen cuantizada a 8 niveles, utilizada como punto de partida para la generación de la tabla de codificación. Visualmente, la imagen conserva gran parte del detalle estructural, aunque presenta una ligera pérdida de graduación tonal debido a la reducción de niveles.

A partir de la frecuencia de aparición de cada nivel de gris, se ha construido la tabla de codificación de Huffman que se muestra a continuación. En ella, cada símbolo representa un nivel de gris (de 0 a 7) y se le asigna un código binario de longitud variable según su probabilidad de aparición:



Símbolo	Código Huffman
0	01
1	001
2	111
3	1100
4	10
5	000
6	11010
7	11011

Tabla 1: Código Huffman por símbolo

La longitud media de los códigos obtenidos es de 2.6338 bits por símbolo, lo que representa una reducción significativa respecto a los 3 bits por símbolo necesarios en la codificación uniforme de 8 niveles. Esto indica que la imagen presenta una distribución no uniforme de intensidades, con ciertos niveles más frecuentes que otros, permitiendo una compresión efectiva mediante Huffman.

El proceso de codificación Huffman ha permitido obtener una representación más compacta de la imagen, reduciendo el número medio de bits por píxel de 3 a aproximadamente 2.63 sin pérdida de información.

Esta mejora se debe al aprovechamiento de la redundancia estadística inherente a la imagen, donde los tonos de gris más comunes son codificados con menos bits. El resultado obtenido constituye la base para los siguientes apartados del análisis, en los que se evaluará la eficiencia de la compresión lograda y la posterior decodificación de la imagen.

### 3.2.Codificar la imagen en binario según la tabla obtenida.

Una vez obtenida la tabla de códigos de Huffman, el siguiente paso consiste en codificar la imagen completa utilizando dicha tabla. Este proceso transforma cada nivel de gris en su secuencia binaria correspondiente, de longitud variable, generando así una representación comprimida del contenido sin pérdida de información. La principal ventaja de este método es que los píxeles más frecuentes se representan con menos bits, mientras que los menos frecuentes utilizan códigos más largos, logrando una compresión proporcional a la redundancia estadística de la imagen.

Parámetro	Valor
Número total de píxeles	1,828,800



La imagen utilizada contiene 1,828,800 píxeles, una resolución de píxeles. Aplicando la tabla de codificación Huffman obtenida en el apartado anterior, se ha calculado el número total de bits necesarios para representar la imagen completa.

Bits totales con Huffman	4,816,632	cuantizada un total de correspondiente a 1600 × 1143
Bits medios por píxel (Huffman)	2.6338	

Tabla 2: Valores de codificación Huffman

La longitud media de 2.6338 bits por píxel coincide con la obtenida en la tabla de codificación teórica, confirmando la coherencia entre la distribución estadística de la imagen y la eficiencia del algoritmo aplicado. Si se compara con la representación uniforme original de 8 niveles de gris (que requiere 3 bits por píxel), se observa una reducción aproximada del 12.2 % en el tamaño de la codificación. Este ahorro demuestra la efectividad del método Huffman al adaptar la longitud de los códigos a la frecuencia real de los símbolos.

La codificación Huffman ha permitido generar una representación binaria comprimida de la imagen, disminuyendo el número medio de bits por píxel de 3 a 2.63 bits, lo que se traduce en una reducción de tamaño sin pérdida de información. Este resultado confirma que la imagen posee redundancia estadística, ya que algunos niveles de gris aparecen con mayor probabilidad que otros.

La codificación binaria obtenida servirá como base para los apartados siguientes, donde se realizará la decodificación y se analizarán la ratio de compresión y la redundancia global de los datos.

### 3.3. Decodificar la imagen según la tabla y árbol de Huffman antes obtenidos.

Tras la codificación de la imagen mediante el esquema de Huffman, el siguiente paso consiste en realizar la decodificación a partir del flujo binario generado y del árbol de códigos obtenido previamente. Este proceso tiene como objetivo comprobar la correcta reconstrucción de la información y verificar que la codificación implementada es completamente reversible, tal como debe ocurrir en los métodos de compresión sin pérdida.

La decodificación se ha llevado a cabo utilizando la misma tabla de correspondencias entre símbolos y códigos binarios obtenida en el apartado A. Cada secuencia de bits se tradujo



Ilustración 7: Imagen reconstruida con C. Huffman



nuevamente a su símbolo de nivel de gris correspondiente, reconstruyendo así la matriz de intensidades original. El resultado se muestra en la Ilustración 7, donde se aprecia la imagen decodificada a 8 niveles de gris, coincidente con la versión cuantizada utilizada antes de la codificación.

La comparación visual entre la imagen cuantizada original y la imagen decodificada no muestra diferencias perceptibles. Para confirmar este resultado, se ha calculado el error absoluto entre ambas matrices, obteniéndose un valor nulo en todos los píxeles. Este resultado demuestra que el proceso de codificación y decodificación es totalmente reversible, manteniendo la integridad completa de los datos.

El experimento confirma que la decodificación Huffman se ha realizado correctamente, obteniendo una imagen reconstruida idéntica a la versión cuantizada original. Esto demuestra la naturaleza sin pérdida del método de Huffman, el cual permite reducir el número medio de bits por píxel sin alterar la información visual ni radiométrica. De este modo, se garantiza que la compresión alcanzada en los apartados anteriores no compromete la calidad de la imagen ni su fidelidad respecto a los datos originales.

### 3.4. Analizar la ratio de compresión respecto a la imagen original con escala de grises de 8 valores.

Una vez completados los procesos de codificación y decodificación de Huffman, es posible evaluar la eficiencia de la compresión obtenida. La medida más común para cuantificar esta eficiencia es la ratio de compresión ( $R_c$ ), definida como la relación entre el número de bits necesarios para representar la imagen original y el número de bits empleados tras la compresión. Este indicador permite conocer cuánto se ha reducido el tamaño de la imagen sin pérdida de información.

En la imagen cuantizada a 8 niveles, cada píxel requiere 3 bits para su representación uniforme, ya que  $2^3 = 8$ . Dado que la imagen contiene un total de 1,828,800 píxeles, el número total de bits en la representación original es:

$$\text{Bits}_{\text{original}} = 1,828,800 \times 3 = 5,486,400 \text{ bits.}$$

Tras aplicar la codificación Huffman, el número total de bits utilizados es de 4,816,632 bits. Por tanto, la ratio de compresión se calcula como:

$$R_c = \frac{\text{Bits}_{\text{original}}}{\text{Bits}_{\text{Huffman}}} = \frac{5,486,400}{4,816,632} = 1.1391$$

La codificación Huffman aplicada a la imagen de 8 niveles ha permitido obtener una ratio de compresión de 1.1391, lo que representa una reducción de tamaño cercana al 14 %. Aunque no se trata de una compresión muy elevada, resulta significativa considerando que el método es sin pérdida y que los niveles de gris de la imagen presentan una distribución relativamente amplia.

Estos resultados evidencian la capacidad de Huffman para mejorar la eficiencia del almacenamiento de imágenes cuando existe redundancia estadística en los valores de intensidad, manteniendo la calidad y fidelidad de los datos originales.

### 3.5. Analizar la redundancia de datos para el caso de 8 niveles de gris.

El análisis de redundancia permite evaluar en qué medida los datos de una imagen contienen información repetida o predecible que puede aprovecharse para mejorar la eficiencia de la codificación.



En el caso de una imagen cuantizada a 8 niveles de gris, el número máximo de bits necesarios para representar cada símbolo mediante un código fijo es de 3 bits por píxel. No obstante, si la distribución de intensidades no es uniforme, la entropía de la imagen será menor, y por tanto existirá redundancia susceptible de ser eliminada mediante métodos como Huffman.

La entropía de la imagen cuantizada se ha calculado a partir de la probabilidad de aparición de cada nivel de gris, obteniéndose un valor de:

$$H = 2.5547 \text{ bits/símbolo}$$

Este valor representa el límite teórico mínimo de bits necesarios para representar cada píxel sin pérdida de información. Dado que la codificación fija utiliza 3 bits por símbolo, la redundancia absoluta presente en la imagen es:

$$R = 3 - H = 0.4453 \text{ bits/símbolo}$$

lo que indica que aproximadamente medio bit por píxel no aporta información nueva. La redundancia relativa, expresada como porcentaje respecto al código fijo, es:

$$R_{\text{rel}} = \frac{3 - H}{3} \times 100 = 14.84\%$$

Este valor confirma que cerca de un 15 % de los bits de la representación uniforme de 8 niveles son redundantes. Finalmente, la longitud media del código de Huffman obtenido es:

$$L_{\text{Huffman}} = 2.6338 \text{ bits/símbolo}$$

un valor cercano a la entropía teórica, aunque ligeramente superior, tal como es esperado en un código práctico que cumple las condiciones de prefijo y unicidad.

El estudio de la redundancia confirma que la imagen cuantizada contiene una proporción significativa de información repetida o predecible, con una entropía inferior a los 3 bits teóricos del código fijo. La codificación Huffman se aproxima al límite marcado por la entropía, reduciendo la longitud media de los códigos hasta 2.6338 bits, lo que evidencia un uso eficiente de la redundancia estadística de la imagen. En conjunto, este análisis demuestra que la imagen presenta un grado moderado de redundancia, suficiente para justificar el uso de técnicas de compresión sin pérdida como la empleada en este apartado.

#### 4. Análisis de Transformada Discreta del Coseno

##### 4.1. Realizar la transformada discreta del coseno para la imagen codificada antes con Huffman de 8 niveles de gris. La ventana de la transformada será de tamaño 8x8

La Transformada Discreta del Coseno (DCT) es una herramienta fundamental en la compresión de imágenes, ya que permite representar bloques locales mediante coeficientes que concentran la energía en pocas componentes de baja frecuencia. Este comportamiento es especialmente útil en algoritmos como JPEG, donde la información perceptualmente menos relevante (altas frecuencias) puede eliminarse o cuantizarse con mayor agresividad sin producir deterioro visual significativo.



El objetivo de este apartado es aplicar la DCT en bloques de  $8 \times 8$  píxeles sobre la imagen previamente cuantizada a 8 niveles y analizar la distribución de sus coeficientes. Para ello, se prepara la imagen mediante un proceso de padding para asegurar que sus dimensiones son múltiplo de 8, requisito indispensable para el procesado por bloques.

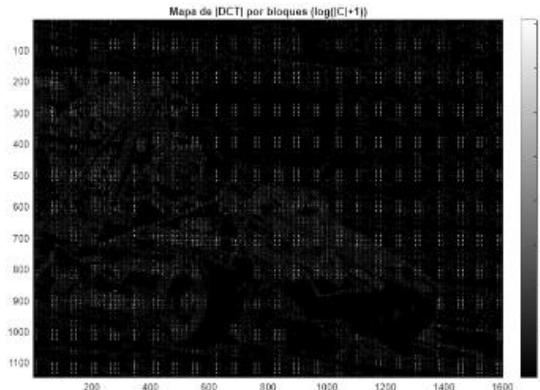


Ilustración 8: Mapa de la DCT

La imagen cuantizada tiene una resolución original de  $1143 \times 1600$  píxeles. Dado que la dimensión vertical no es múltiplo de 8, se añade una fila adicional al final, obteniendo una imagen con dimensiones  $1144 \times 1600$  píxeles, sobre la que se ejecuta la DCT en bloques regulares.

La matriz muestra un mapa de calor que representa la magnitud de los coeficientes DCT para todos los bloques de la imagen, escalados mediante la función logarítmica  $\log(|C|+1)$  para mejorar su visualización. En este mapa se aprecia que la mayor parte de la energía se concentra en las componentes de baja frecuencia de cada bloque (zona superior izquierda de cada cuadrícula), mientras que las altas frecuencias presentan coeficientes cercanos a cero.

Como referencia, se muestra también un bloque DCT de la esquina superior izquierda, correspondiente al primer bloque de  $8 \times 8$  píxeles de la imagen:

$$\begin{bmatrix} 16 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{bmatrix}$$

Este bloque revela un caso particularmente simple: la DCT presenta un único coeficiente distinto de cero, correspondiente a la frecuencia DC, lo cual indica que el bloque tiene una intensidad prácticamente uniforme. Este comportamiento es característico de regiones de la imagen con poca variación local, lo que demuestra la capacidad de la DCT para separar las componentes estructurales de baja frecuencia de los detalles finos.

La aplicación de la DCT por bloques de  $8 \times 8$  ha permitido observar cómo la energía de la imagen se concentra fundamentalmente en el coeficiente DC y en unas pocas componentes de baja frecuencia. Las altas frecuencias, que representan variaciones rápidas de intensidad, presentan valores muy pequeños o nulos en muchos bloques, lo que sugiere que pueden eliminarse o cuantizarse sin afectar de manera significativa a la calidad perceptual.

Este comportamiento confirma la idoneidad de la DCT como herramienta para la compresión de imágenes, preparando el terreno para los análisis posteriores sobre ratio



de compresión, redundancia y reconstrucciones con máscaras de coeficientes significativos.

#### 4.2. Analizar la ratio de compresión y la redundancia de datos

La DCT permite representar bloques de  $8 \times 8$  píxeles mediante 64 coeficientes ordenados por importancia en zig-zag. Al conservar únicamente los coeficientes más significativos es posible reducir de forma notable la cantidad de información almacenada, obteniendo compresión a costa de una pérdida progresiva de detalle.

Se ha estimado la ratio de compresión para diferentes valores de  $K$ , donde  $K$  es el número de coeficientes DCT conservados por bloque. Los resultados se resumen en la tabla:

<b>K coef./bloque</b>	<b>Bits aproximados</b>	<b>Ratio Rc</b>
64	5,491,200	1.00
32	2,745,600	2.00
16	1,372,800	4.00
10	858,000	6.39
6	514,800	10.66

Tabla 3: RC para diferentes valores de  $K$

Estos valores muestran que la compresión aumenta rápidamente al reducir  $K$ . Con 32 coeficientes ya se obtiene una compresión del 50 %, mientras que con 10 o 6 coeficientes la reducción es muy elevada, aunque acompañada de un deterioro visible en la reconstrucción de la imagen.

El análisis confirma que la DCT es altamente eficiente para la compresión: conservar entre 16 y 32 coeficientes ofrece un equilibrio razonable entre compresión y calidad, mientras que reducir por debajo de 10 coeficientes produce compresiones muy altas, pero con pérdidas perceptibles. La tendencia observada es coherente con el funcionamiento de esquemas basados en DCT como JPEG.

#### 4.3. Recuperar la imagen de la transformada discreta del coseno y comparar el resultado con la imagen original codificada. Probar con diferentes máscaras de valores significativos de la DCT y cuantificar las diferencias.

Una de las aplicaciones más relevantes de la DCT es la posibilidad de reconstruir una imagen utilizando únicamente un subconjunto de sus coeficientes. Este proceso permite estudiar de forma directa la relación entre el número de coeficientes conservados, la calidad de la imagen reconstruida y el error introducido. En este apartado se evalúa la reconstrucción de la imagen al conservar únicamente los  $K$  coeficientes más significativos de cada bloque de  $8 \times 8$ , ordenados en zig-zag. La calidad se cuantifica mediante el Error Cuadrático Medio (MSE) y la Relación Señal-Ruido (PSNR).



La tabla muestra las reconstrucciones obtenidas para diferentes valores de  $K$ . A medida que el número de coeficientes disminuye, aumentan el error y la pérdida de detalle, especialmente en zonas con bordes o texturas.

Los valores obtenidos son los siguientes:

<b>K coef./bloque</b>	<b>MSE</b>	<b>PSNR</b>
64	0.0000	$\infty$ dB
32	0.0027	42.54 dB
16	0.0191	34.08 dB
8	0.0465	30.22 dB
4	0.1125	26.39 dB
2	0.1916	24.08 dB

Tabla 4: valores de MSE y PSNR para distintos  $K$

En los casos de  $K = 64$  y  $K = 32$ , la reconstrucción es prácticamente perfecta, con PSNR superior a 40 dB. Esto indica que los coeficientes de baja frecuencia contienen la mayor parte de la información visual de la imagen.

Con valores intermedios, como  $K = 16$  u  $8$ , la degradación comienza a ser perceptible, aunque la estructura global se mantiene adecuadamente. Para valores extremos, como  $K = 4$  o  $K = 2$ , el error aumenta de forma notable, produciendo una pérdida significativa de textura y suavizado excesivo debido a la eliminación de las componentes de alta frecuencia.

Los resultados confirman que la DCT concentra la mayor parte de la energía de cada bloque en los coeficientes de baja frecuencia. Conservando únicamente entre 16 y 32 coeficientes es posible obtener una reconstrucción de alta calidad, mientras que valores muy bajos de  $K$  generan un deterioro visual perceptible.

Este comportamiento refleja el fundamento de los esquemas de compresión basados en DCT, donde se descartan o cuantizan más agresivamente los coeficientes asociados a altas frecuencias, manteniendo la fidelidad visual con un coste de almacenamiento reducido.



## 5. Análisis con otras transformadas

### 5.1. Analizar y comparar lo que ocurre al hacer una transformada de Fourier y comparar con los resultados obtenidos con la transformada del coseno

La FFT 2D de la imagen revela cómo se distribuye la energía en el dominio frecuencial. La Figura correspondiente muestra un espectro con un claro máximo en el centro, indicando que la mayor parte de la energía se concentra en bajas frecuencias, mientras que las altas frecuencias presentan menor magnitud. Este comportamiento confirma que la información estructural de la

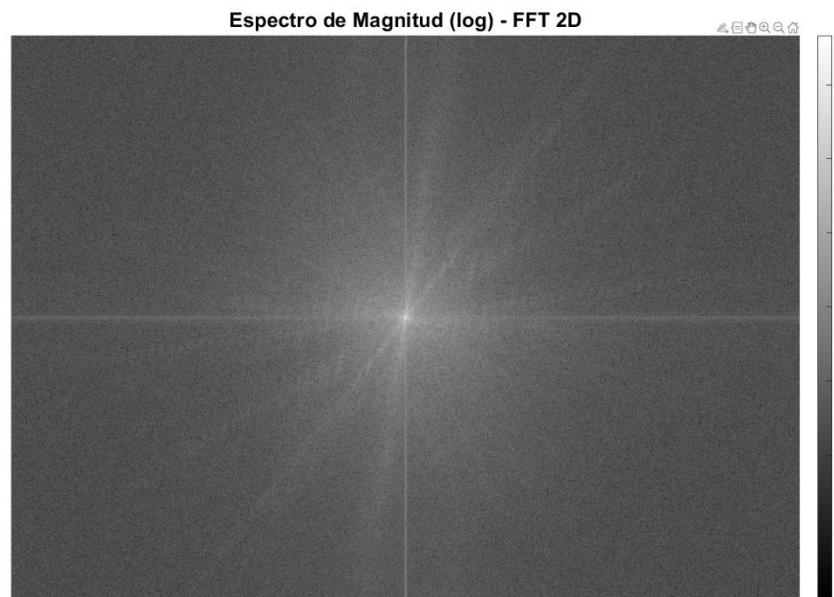


Ilustración 9: Espectro de Magnitud de la FFT

imagen reside principalmente en regiones suavemente cambiantes.

Se reconstruyó la imagen conservando únicamente una región circular de radio normalizado  $r$  alrededor del centro del espectro. Los resultados son:

<b>r</b>	<b>MSE</b>	<b>PSNR</b>
1 (sin pérdida)	0.0000	$\infty$ dB
0.5	0.0177	34.42 dB
0.25	0.0692	28.50 dB
0.125	0.1886	24.15 dB
0.0625	0.3967	20.92 dB

Tabla 5: MSE y PSNR



La calidad de reconstrucción se mantiene razonablemente buena hasta  $r \approx 0.5$ . Por debajo de ese valor, la pérdida de detalle aumenta de forma brusca, apareciendo efectos de suavizado y pérdida de textura. Comparada con DCT, la FFT produce deterioros más notorios para el mismo nivel de reducción debido a que no opera por bloques ni sigue criterios perceptuales.

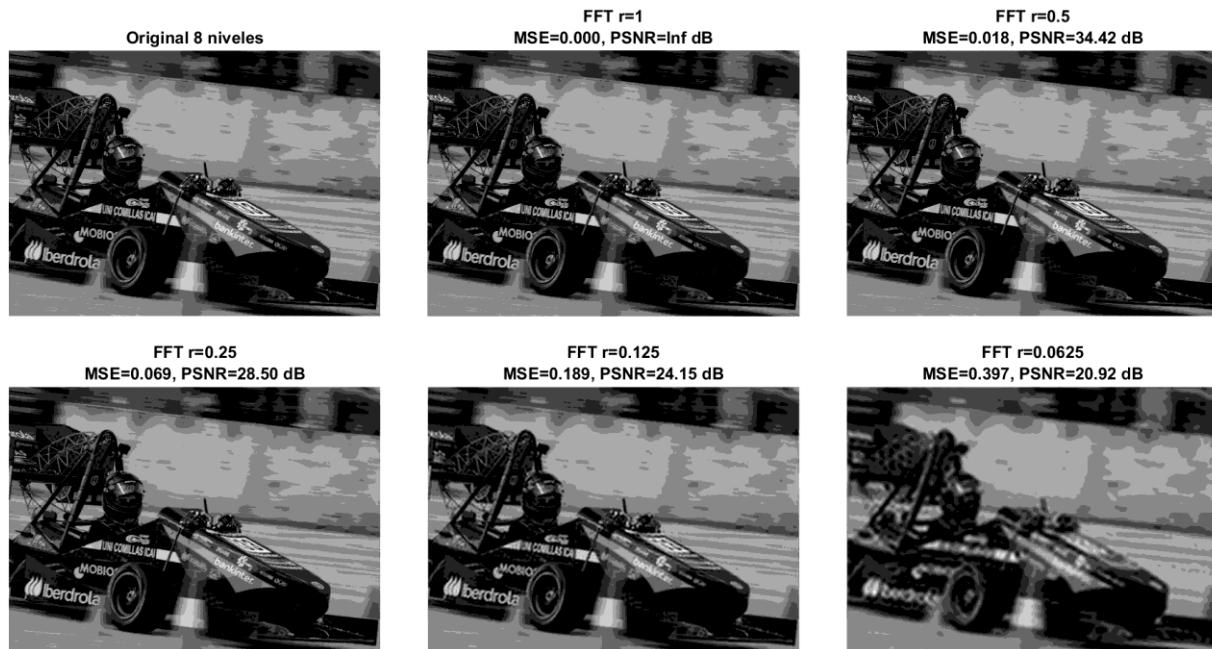


Ilustración 10: Comparación de imágenes por FFT

## 5.2. Analizar y comparar lo que ocurre al hacer una transformada tipo wavelet y comparar con los resultados obtenidos con la transformada del coseno

Se aplicó la Transformada Wavelet discreta (Haar) en niveles, eliminando coeficientes por orden de magnitud. Los resultados obtenidos fueron:

% coef.	Rc aprox.	MSE	PSNR
100%	1.00	0.0000	$\infty$ dB
50%	2.00	0.0000	$\infty$ dB
25%	4.00	0.0000	$\infty$ dB
10%	9.99	0.0165	34.73 dB
5%	19.98	0.1036	26.75 dB

Tabla 6: Comparación de coeficientes Wavelet

La wavelet demuestra una extraordinaria capacidad de compactación:

- Hasta el 25% de coeficientes se pueden eliminar sin ningún error.
- Incluso con solo 10%, la calidad sigue siendo alta ( $PSNR \approx 34.7$  dB).
- Aún con un extremo 5%, la estructura global de la imagen permanece reconocible.



Esto se debe a que la wavelet representa muy eficientemente bordes y discontinuidades, algo que ni la FFT ni la DCT logran tan bien.

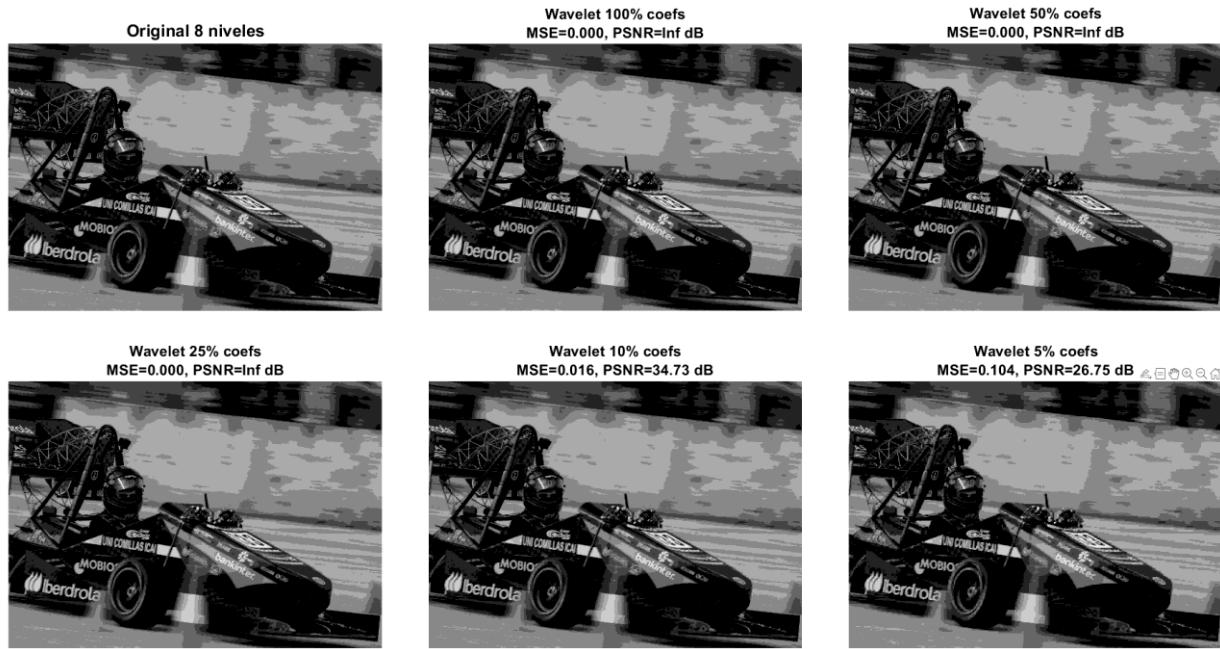


Ilustración 11: Comparación de imágenes con wavelet

### 5.3. Analizar y comparar lo que ocurre al hacer una transformada tipo KLT y comparar con los resultados obtenidos con la transformada del coseno

La KLT maximiza la compactación de energía porque sus vectores propios están adaptados a las estadísticas reales de la imagen. Es, teóricamente, la transformada óptima para compresión.

Se evaluó la KLT por bloques, conservando únicamente los primeros  $K$  autovectores de mayor energía. Los resultados obtenidos fueron:

<b>K</b>	<b>Rc aprox.</b>	<b>MSE</b>	<b>PSNR</b>
64	1.00	0.0000	$\infty$ dB
32	2.00	0.0025	42.94 dB
16	4.00	0.0172	34.55 dB
8	7.99	0.0447	30.40 dB
4	15.99	0.0939	27.18 dB

Ilustración 12: Componentes  $K$

La KLT es la transformada óptima en términos de compactación de energía, ya que está adaptada estadísticamente a los datos. Los resultados confirman:

- Rendimiento muy similar a la DCT para niveles bajos de compresión.
- Ligera ventaja sobre la DCT para compresiones extremas.
- Altísima calidad en el caso  $K=32$  ( $PSNR \approx 43$  dB).

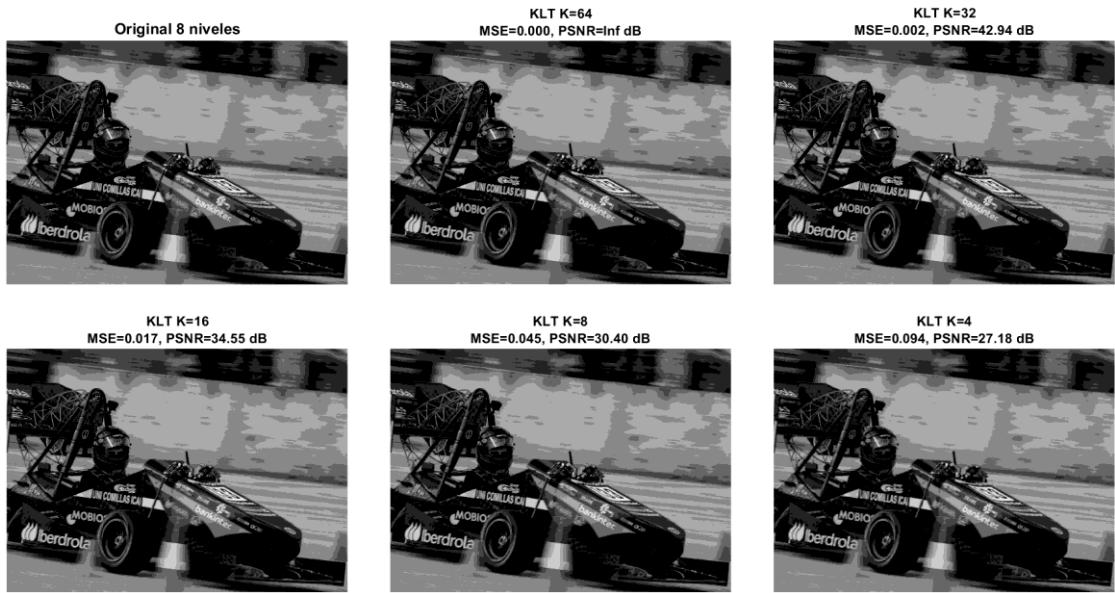


Ilustración 13: Comparación de imágenes por KLT

Su principal inconveniente es su alto coste computacional, lo que limita su uso en sistemas reales.

## 6. Conclusión

La práctica ha permitido realizar un análisis completo del proceso de cuantización y compresión de imágenes mediante técnicas clásicas de procesamiento digital. A partir de la imagen original, primero se estudió su representación radiométrica, su histograma y el impacto de reducir el número de niveles de gris. La cuantización demostró que, aunque la imagen puede conservar un aspecto razonable con un número relativamente bajo de niveles, por debajo de 16 niveles la pérdida de tonalidad se vuelve evidente debido a la aparición de regiones planas y la desaparición de gradientes suaves.

Posteriormente, se evaluó la codificación de Huffman aplicada a la imagen de 8 niveles, comprobándose que la distribución no uniforme de intensidades permite reducir el número medio de bits por píxel desde 3 a 2.6338 sin pérdida de información. La decodificación confirmó la reversibilidad del proceso y el cálculo de entropía reveló una redundancia aproximada del 15%, lo que justifica el uso de codificación estadística para optimizar almacenamiento.

El estudio de la DCT por bloques de  $8 \times 8$  mostró cómo la mayor parte de la energía se concentra en las bajas frecuencias, permitiendo reconstrucciones de alta calidad incluso cuando solo se conservan 16 o 32 coeficientes por bloque. La relación entre compresión y error se manifestó claramente: a medida que disminuye el número de coeficientes, la pérdida de detalle se vuelve más pronunciada, especialmente en texturas y bordes. Este comportamiento valida el fundamento del estándar JPEG, que se basa precisamente en esta propiedad.

A continuación, se compararon otras transformadas: FFT, Wavelet y KLT. La FFT demostró ser menos eficiente para compresión perceptual, ya que pequeñas reducciones en el espectro producen deterioros visibles. Por su parte, la Transformada Wavelet destacó como la más robusta: fue posible eliminar hasta un 25%–50% de los coeficientes sin pérdida alguna, y mantener una calidad elevada incluso con solo el 10% de los coeficientes originales. Finalmente, la KLT, aunque teóricamente óptima en compactación de energía, obtuvo resultados solo ligeramente superiores a los de la DCT, a costa de un mayor coste computacional.



En conjunto, la práctica ha permitido observar de manera integradora cómo distintas transformadas y técnicas de codificación explotan diferentes tipos de redundancia en la imagen (estadística, espacial, frecuencial). Los resultados confirman que:

- la cuantización controla directamente la fidelidad radiométrica,
- la codificación entropía reduce redundancia estadística sin pérdida,
- las transformadas permiten comprimir información conservando solo los componentes más relevantes,
- y la eficacia de cada método depende de cómo representen la energía y la estructura de la imagen.

Se concluye que la DCT y la Wavelet son las transformadas más equilibradas en términos de compresión y calidad visual, mientras que la KLT destaca por su rendimiento teórico y la FFT resulta menos adecuada para compresión perceptual. Esta práctica proporciona una visión global de los principios fundamentales de la compresión de imágenes y sus implicaciones en calidad, eficiencia y representación.

## 7. Anexo Código

Igual que con prácticas anteriores, se ha incluido un enlace al repositorio en Github donde se encuentra todo el contenido de la práctica (códigos, documentos y archivos de audio):  
<https://github.com/MrAndy5/PracticasSTM/tree/main/P5>