



# PRÁCTICA 6

Compresión de Imágenes con JPEG

Universidad Pontifical de  
Comillas ICAI

Andrés Sánchez de Ágreda – 202108563  
Servicios Telemáticos Multimedia



## ÍNDICE

1. Introducción.....	1
2. YCbCr .....	2
3. DCT e IDCT.....	4
4. Cuantización .....	6
5. Codificación en zigzag.....	8
6. Codificación de entropía. DPCM, Huffman.....	10
7. Decodificación.....	11
8. Relación de compresión .....	12
9. Relación señal ruido .....	13
10. Conclusión.....	13
11. Anexo Código.....	14



## 1. Introducción

En esta práctica se estudia en detalle el proceso de compresión JPEG aplicado a una imagen real utilizando MATLAB. El objetivo principal es comprender, analizar y reproducir todas las etapas fundamentales del estándar JPEG: conversión de espacio de color, transformada DCT por bloques, cuantización, organización en zig-zag, codificación de entropía y, finalmente, reconstrucción y evaluación de la calidad resultante.

El flujo completo permite observar cómo se reduce la redundancia espacial y estadística de la imagen, a la vez que se controla la pérdida de información mediante matrices de cuantización. Asimismo, se evalúan las métricas más relevantes en este contexto, como la relación de compresión ( $R_c$ ), el error de reconstrucción y los valores de relación señal-ruido (PSNR y SNR).

La práctica se ha desarrollado en un Live Script de MATLAB, siguiendo una estructura modular donde cada apartado corresponde a una fase concreta del proceso JPEG. Además, se incluyen representaciones visuales e histogramas que permiten analizar de forma más profunda el comportamiento de la imagen en cada etapa.



Ilustración 1: Imagen elegida para la práctica

Para esta práctica se ha seleccionado una imagen real que muestra un caza F-35 realizando una maniobra de despegue vertical sobre la cubierta de un portaaviones. La escena combina elementos de alta complejidad visual: zonas oscuras de la cubierta, gradientes suaves del cielo al atardecer, el mar en movimiento y un halo generado por la potencia de la tobera del avión. Esta variedad de texturas y contrastes convierte a la imagen en un ejemplo adecuado para estudiar la eficacia de los procesos de compresión JPEG. La imagen tiene una resolución de 1728x1080 píxeles, con sus canales RGB teniendo una profundidad de 8 bits por canal.

A continuación, se representa la imagen original y se muestran los histogramas de intensidad de cada canal (R, G y B). Estos histogramas permiten observar la distribución de niveles de color presentes en la escena, mostrando cómo la mayor parte de los valores se concentran en intensidades medias y altas, especialmente en los canales G y B debido al tono dominante del cielo.

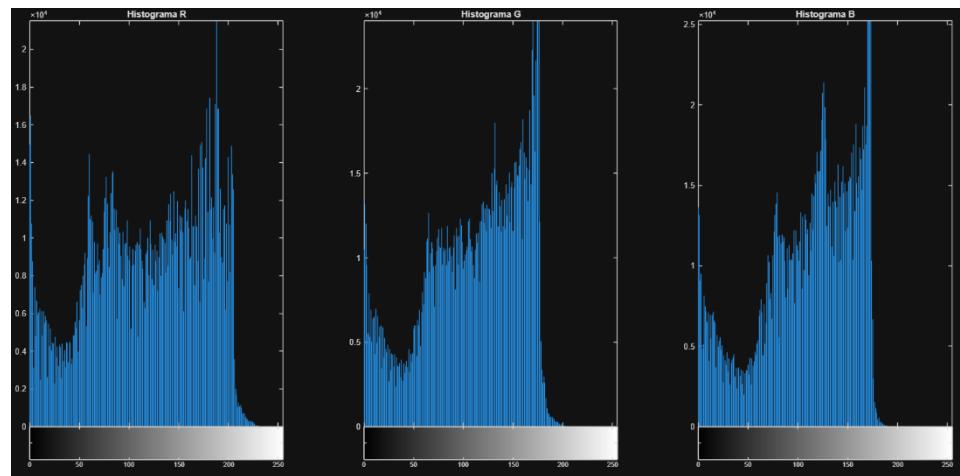


Ilustración 2: Histograma de colores RGB de la imagen

El análisis preliminar indica que:

- El canal **R** presenta mayor presencia en zonas cálidas del cielo (atardecer).
- El canal **G** concentra más información en la región marina.
- El canal **B** presenta gran densidad en valores medios-altos, típico de imágenes con fuerte componente celeste.

Estas características influyen directamente en cómo se comportan los coeficientes DCT y en cómo se distribuyen los errores tras la cuantización, afectando de forma distinta a cada componente del color.

## 2. YCbCr

El estándar JPEG trabaja internamente en el espacio de color YCbCr, ya que este modelo separa la información de luminancia (Y) de la crominancia (Cb y Cr). Esta separación permite explotar características del sistema visual humano, que es mucho más sensible a variaciones de luminosidad que a cambios sutiles en el color. Por ello, JPEG aplica mayor compresión sobre Cb y Cr que sobre Y sin producir una degradación perceptible significativa.

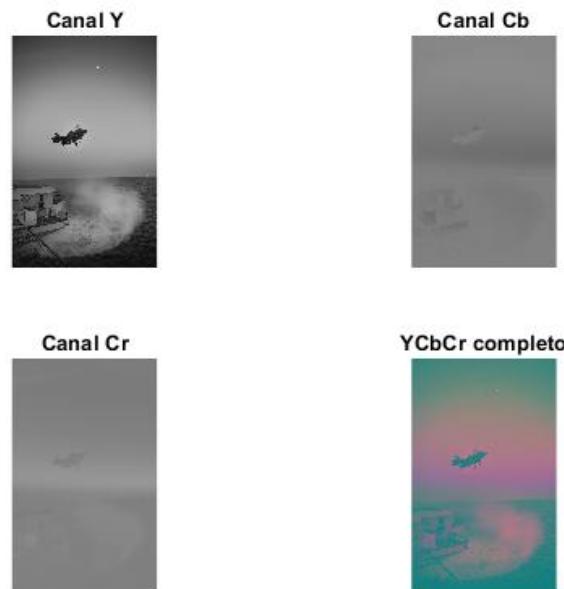


Ilustración 3: Representación de canales YCbCr



Tras convertir la imagen original del espacio RGB a YCbCr, se analizan por separado los tres canales:

El canal Y representa la intensidad o brillo de cada píxel. Este componente conserva prácticamente toda la estructura visual de la imagen:

- forma del F-35,
- borde del mar,
- la cubierta del portaaviones,
- gradientes suaves del cielo al atardecer.

Al observar el canal Y, se aprecia que la imagen conserva sus detalles esenciales incluso sin color, lo que confirma que la mayor parte de la información perceptual se encuentra en esta componente.

El canal Cb describe cuánto tiende cada píxel hacia tonalidades azuladas o amarillentas. En este caso:

- El cielo y el mar presentan niveles intermedios de Cb.
- La cubierta y el avión muestran valores más bajos, lo que significa menor presencia de componente azul.
- La distribución general es suave, sin altos contrastes, lo cual es ideal para la compresión.

Se observa también que muchas zonas de la imagen apenas contienen variación cromática relevante, algo que JPEG podrá aprovechar reduciendo la cantidad de coeficientes necesarios para representarlo.

El canal Cr indica la contribución rojiza de cada píxel. En este canal:

- El cielo al atardecer muestra niveles algo más altos debido a los tonos cálidos.
- El mar presenta valores más bajos por su tonalidad azulada.
- La estructura del barco y el avión apenas contienen componente rojo, por lo que se muestran más apagados.

Al igual que en Cb, la información de Cr es mucho más uniforme y con menos detalle, lo que facilita una compresión efectiva.

La visualización del conjunto YCbCr evidencia:

- la separación clara entre luminancia (muy detallada)
- y crominancia (mucho más uniforme y de baja frecuencia espacial)

Esta representación no debe confundirse con una imagen en RGB: aunque se vea con colores extraños, dicha visualización es útil para comprobar que la conversión se ha realizado correctamente.

La conversión al espacio YCbCr permite que la información visual se reparta de forma más eficiente para la compresión:

- El canal Y conserva la estructura y los detalles finos.



- Los canales Cb y Cr contienen variaciones suaves y de baja frecuencia.
- JPEG podrá aplicar más compresión sobre Cb y Cr sin que el resultado sea perceptible para el usuario.

Este paso es fundamental, ya que prepara la imagen para el tratamiento por bloques y para la cuantización posterior, reduciendo considerablemente la redundancia y el tamaño final tras la compresión.

### 3. DCT e IDCT

La Transformada Discreta del Coseno (DCT) es uno de los pilares fundamentales del estándar JPEG. Su función es convertir la información espacial de cada bloque de la imagen en información frecuencial, permitiendo distinguir entre componentes de baja frecuencia (variaciones suaves) y alta frecuencia (bordes, texturas y detalles finos).

En esta práctica se ha aplicado la DCT de forma independiente sobre cada canal Y, Cb y Cr de la imagen en formato YCbCr, utilizando bloques de  $8 \times 8$  píxeles, tal y como establece el estándar. Posteriormente, se ha reconstruido la imagen mediante la transformada inversa (IDCT), sin aplicar aún cuantización, por lo que el proceso es sin pérdidas salvo por pequeñas imprecisiones numéricas.

El proceso consiste en:

1. Dividir cada canal en bloques de  $8 \times 8$  píxeles.
2. Aplicar la DCT a cada bloque de forma independiente:
  - La DCT concentra la mayor parte de la energía del bloque en las esquinas superiores izquierdas (baja frecuencia).
  - Los coeficientes de alta frecuencia tienden a ser más pequeños y dispersos.
3. Guardar los coeficientes y posteriormente invertirlos con IDCT.

Este procedimiento prepara la imagen para los pasos siguientes, especialmente para la cuantización, donde los coeficientes de alta frecuencia podrán reducirse o eliminarse para lograr compresión.

Tras aplicar la DCT e IDCT a todos los bloques, se reconstruye la imagen completa. A simple vista, la imagen reconstruida en RGB es prácticamente indistinguible de la original. Ambas mantienen el mismo nivel de detalle y gradientes suaves, como se aprecia en la figura:



Ilustración 4: Imagen Reconstruida por IDCT

A simple inspección visual se concluye que:

- El cielo conserva sus gradientes suaves.
- El avión mantiene sus bordes, sombras y contornos.
- El mar y la espuma producida por la fuerza del reactor no muestran alteraciones perceptibles.
- No se observan artefactos, bloques ni pérdidas de nitidez.

Esto confirma que la combinación DCT→IDCT, sin cuantización, es un proceso casi perfectamente reversible.

Para enfatizar más este proceso, se ha incluido una ampliación de la imagen. El zoom proporcionado (ampliación del F-35) demuestra con mayor claridad la fidelidad de la reconstrucción:



Ilustración 5: Ampliación de la imagen reconstruida



Al compararse esta región con la misma zona en la imagen original se observa:

- No hay pérdida de definición ni suavizado de bordes.
- Las líneas del fuselaje, la cabina y las compuertas siguen siendo nítidas.
- La iluminación y los tonos se preservan adecuadamente.
- Los detalles de las compuertas del tren de aterrizaje, especialmente sensibles a los cambios, permanecen intactos.

Esto concuerda con el comportamiento teórico de la DCT: la compresión aún no ha comenzado, y al no eliminar ni modificar coeficientes, la reconstrucción debe ser precisa.

La aplicación de la DCT por bloques de  $8 \times 8$  y su reconstrucción mediante IDCT ofrece una demostración clara de la capacidad de esta transformada para representar la imagen en el dominio frecuencial sin pérdida visible de información. La imagen reconstruida es prácticamente idéntica a la original, lo que confirma que la DCT por sí misma no altera la calidad visual. Este comportamiento será clave para comprender las diferencias que aparecerán en los apartados posteriores, donde la cuantización y la codificación reducirán la información con pérdidas controladas para lograr compresión.

#### 4. Cuantización

La cuantización es el paso clave donde el algoritmo JPEG introduce pérdida de información para reducir significativamente el tamaño de la imagen. Tras aplicar la Transformada Discreta del Coseno (DCT) en bloques de  $8 \times 8$ , los coeficientes obtenidos —especialmente aquellos que representan alta frecuencia— son, en general, poco relevantes para la percepción humana. JPEG aprovecha esta propiedad eliminando o reduciendo la precisión de dichos coeficientes mediante una matriz de cuantización.

Este proceso convierte una representación con valores reales en otra mucho más comprimible y con gran cantidad de ceros, lo cual es fundamental para los pasos posteriores de codificación.

Los pasos realizados en este apartado son:

1. Cuantización

Cada bloque DCT se divide elemento a elemento entre una matriz de cuantización  $Q$ , lo que reduce la precisión de los coeficientes de alta frecuencia:

$$C_q(i,j) = \text{round}\left(\frac{C(i,j)}{Q(i,j)}\right)$$

2. De-cuantización

Para reconstruir la imagen, los coeficientes se multiplican de nuevo por la misma matriz:

$$C'(i,j) = C_q(i,j) \cdot Q(i,j)$$

3. Aplicación de IDCT en los bloques reconstruidos  
Esto produce una imagen aproximada a la original, pero con una pérdida controlada de detalle debido al proceso de cuantización.

La imagen reconstruida tras cuantización y de-cuantización sigue manteniendo una apariencia muy cercana a la original, como se observa en la siguiente figura:



Ilustración 6: Imagen con proceso de cuantización

A primera vista:

- El cielo mantiene sus gradientes suaves sin aparición evidente de artefactos.
- La estructura del barco y el mar se perciben naturales.
- El color general de la escena se mantiene sin alteraciones perceptibles.

Esto indica que la cuantización empleada no ha sido excesivamente agresiva y conserva la mayoría de la información relevante para la percepción visual.

El zoom sobre el F-35 permite observar con más detalle los efectos de la cuantización:



Ilustración 7: Ampliación de la imagen

Aquí se aprecian los primeros indicios visibles de pérdida de información:

#### 1. Reducción de nitidez en bordes

- Los contornos del fuselaje y las superficies metálicas muestran una ligera suavización.



- Las transiciones ya no son tan precisas como en la reconstrucción completa del apartado 3.

### 2. Aparición de bloqueos (blockiness)

- Aunque muy leve, en zonas de baja variación como el cielo es posible observar patrones cuadrados característicos de JPEG.
- Esto se debe a que la DCT trabaja por bloques independientes, y la cuantización amplifica diferencias entre ellos.

### 3. Detalles finos atenuados

- Remaches, sombras pequeñas y elementos del tren de aterrizaje pierden algo de definición.
- Esto confirma que los coeficientes de alta frecuencia (responsables del detalle fino) han sido reducidos o eliminados.

La cuantización constituye el primer paso del proceso donde aparece pérdida visible de información. Aunque el resultado general es muy similar a la imagen original, se evidencian ligeras pérdidas de nitidez en bordes y texturas, así como un incipiente efecto de bloque en ciertas regiones.

Este comportamiento es coherente con el estándar JPEG:

- conserva la luminancia y estructura global de la imagen,
- pero reduce el detalle de alta frecuencia para lograr compresión eficiente,
- manteniendo un aspecto visual satisfactorio.

Estos cambios serán aún más relevantes en los apartados posteriores, donde la eliminación de coeficientes y la codificación de entropía intensificarán el impacto perceptual.

## 5. Codificación en zigzag

El siguiente paso en el proceso JPEG consiste en transformar cada bloque de  $8 \times 8$  coeficientes DCT en un vector unidimensional siguiendo un recorrido en forma de zig-zag. Esta operación no altera la información, pero reorganiza los coeficientes desde las bajas frecuencias hacia las altas, de modo que los valores más relevantes (frecuencias bajas) quedan agrupados al inicio del vector, mientras que los coeficientes menos importantes (altas frecuencias) se sitúan al final.

Este reordenamiento facilita los procesos posteriores de compresión, especialmente la codificación run-length y la codificación de Huffman, ya que tiende a agrupar largas secuencias de ceros tras la cuantización.

En lugar de escoger arbitrariamente un bloque de la imagen, se implementó un método para identificar automáticamente el bloque con mayor energía (suma del valor absoluto de los coeficientes DCT). Esto garantiza que el bloque analizado:

- contiene variación espacial significativa,
- no pertenece a zonas planas (como el cielo),
- es representativo de la acción de la imagen (en este caso, el F-35).

El recuadro mostrado en la imagen siguiente indica exactamente qué bloque fue seleccionado:



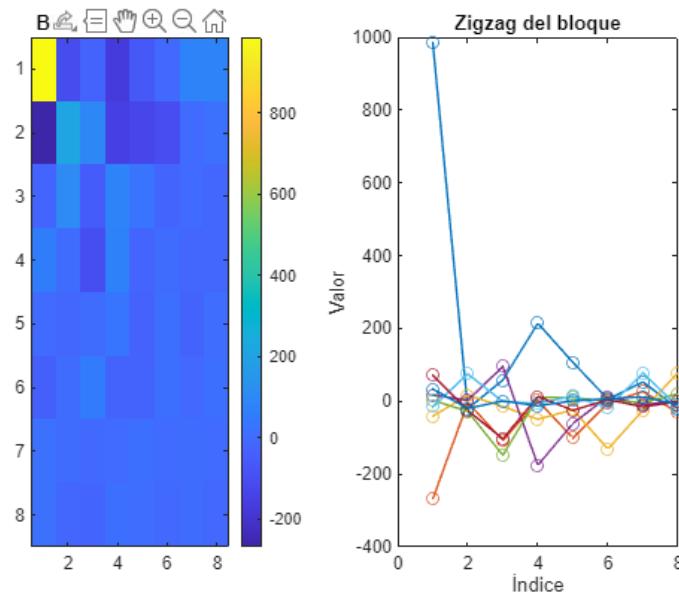
Bloque seleccionado para zigzag



Ilustración 8: Bloque seleccionado para ZigZag

Gracias a este método, el bloque escogido pertenece a una zona con contornos definidos del fuselaje del avión, lo que permite obtener un patrón DCT y un zig-zag mucho más informativo.

Para ilustrar el proceso de reorganización de los coeficientes DCT en el estándar JPEG, se ha seleccionado automáticamente el bloque con mayor energía dentro de la imagen. Esta selección garantiza un bloque con contenido significativo —en este caso, perteneciente a la zona del fuselaje del F-35— y evita regiones uniformes donde la DCT apenas contiene información.



En la figura se muestra simultáneamente:

- a la izquierda, el bloque DCT original (8×8),
- a la derecha, su correspondiente recorrido zig-zag.

Este formato conjunto permite apreciar, de forma directa, la relación entre la distribución de frecuencias espaciales del bloque y la secuencia ordenada que JPEG emplea para su codificación.

El bloque DCT presenta un coeficiente DC dominante en la esquina superior izquierda, seguido de valores moderados en las bajas frecuencias y una rápida atenuación hacia las zonas de alta frecuencia, donde los coeficientes son muy reducidos. Esta estructura es típica de bloques que contienen bordes suaves y variaciones intermedias.

Aplicando el orden de lectura zig-zag, los coeficientes se reorganizan en un vector que prioriza la energía del bloque, situando primero las bajas frecuencias y relegando al final las altas, que suelen ser cercanas a cero. El vector zig-zag obtenido para este bloque comienza con un valor DC muy elevado, seguido de coeficientes de baja frecuencia con valores apreciables y una caída progresiva hacia las frecuencias menos relevantes.

En este ejemplo, la representación mostrada contiene los 8 primeros elementos del zigzag, correspondientes a la zona de mayor energía del bloque.

Esta operación es fundamental para la compresión JPEG, ya que facilita la aparición de largas secuencias de ceros en las posiciones finales del vector tras la cuantización, optimizando de manera significativa la codificación de entropía que se aplicará en los apartados posteriores.

## 6. Codificación de entropía. DPCM, Huffman

Tras la cuantización y el reordenamiento zig-zag, los coeficientes DCT se someten a la fase final de compresión JPEG: la codificación de entropía. Este proceso reduce el tamaño de los datos aprovechando su redundancia estadística.

Los coeficientes DC suelen variar poco entre bloques consecutivos. Por ello, se aplica una codificación diferencial (DPCM), donde se almacena la diferencia entre el DC actual y el anterior. Esto genera valores más pequeños y concentrados alrededor de cero, lo que facilita la codificación posterior.



Tanto las diferencias DC como los coeficientes AC (el resto del zig-zag) se codifican mediante Huffman, que asigna códigos más cortos a los valores más frecuentes. Debido a la gran cantidad de ceros generados tras la cuantización, Huffman permite una reducción significativa del tamaño total.

El tamaño total obtenido tras la codificación de entropía es de 7 678 778 bits. Este valor es coherente con la resolución y complejidad de la imagen utilizada, que contiene numerosas zonas con detalle (avión, mar, espuma). El resultado confirma que la combinación DPCM + Huffman aporta una compresión eficaz sin necesidad de eliminar más información.

La codificación de entropía organiza y optimiza los coeficientes DCT para reducir el tamaño final del archivo. El uso conjunto de DPCM y Huffman es especialmente eficiente gracias a la estructura del zig-zag y la alta presencia de ceros en los coeficientes cuantizados. El tamaño final obtenido valida el correcto funcionamiento del proceso de compresión.

## 7. Decodificación

Una vez aplicada la codificación de entropía, el proceso se invierte para reconstruir la imagen final. Esto implica realizar la decodificación Huffman, reconstruir los coeficientes DC mediante DPCM inverso, aplicar la descompresión AC, reordenar los coeficientes mediante zig-zag inverso, realizar la de-cuantización y aplicar la transformada inversa IDCT por bloques. Finalmente, la imagen YCbCr reconstruida se convierte nuevamente al espacio RGB.

El resultado obtenido mantiene una apariencia muy similar a la imagen original gracias a que las pérdidas introducidas durante la cuantización han sido moderadas.

Para evaluar la calidad de la reconstrucción, se calcula la diferencia entre la imagen original y la imagen reconstruida:

$$\text{Error} = I_{\text{original}} - I_{\text{reconstruida}}$$

El histograma del error obtenido se muestra en la figura:

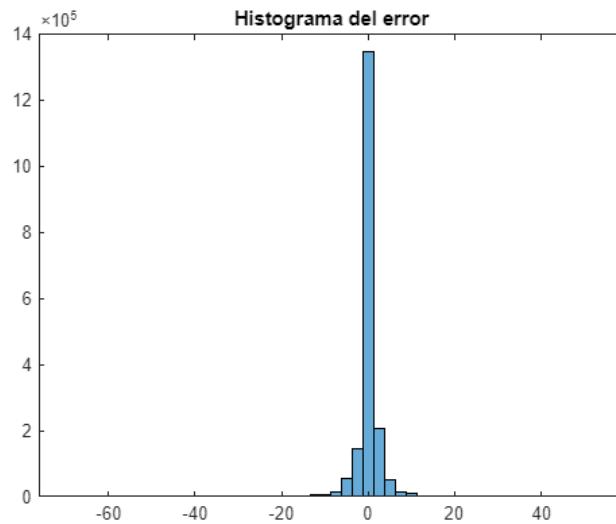


Ilustración 9: Histograma del error correspondiente a la reconstrucción de la imagen



Este gráfico revela que:

- La inmensa mayoría de los errores se encuentran alrededor de **0**,
- Solo aparece una dispersión muy pequeña hacia valores positivos y negativos,
- No existen errores extremos ni desplazamientos significativos,
- El patrón estrecho y centrado confirma que la reconstrucción es de alta calidad.

El hecho de que el histograma esté tan concentrado alrededor del cero indica que las diferencias entre ambas imágenes son mínimas y prácticamente imperceptibles al ojo humano.

La reconstrucción final preserva fielmente la estructura visual de la imagen original. El histograma del error confirma que las pérdidas se limitan a pequeñas variaciones introducidas por la cuantización, sin degradación significativa del contenido. La mayoría de los valores de error se sitúan en torno a cero, demostrando que el proceso JPEG aplicado ha sido eficiente y ha mantenido una calidad visual elevada, acorde al nivel de compresión utilizado.

## 8. Relación de compresión

La relación de compresión ( $R_c$ ) permite cuantificar la eficacia del proceso JPEG aplicado sobre la imagen. Se define como:

$$R_c = \frac{\text{Tamaño de la imagen original}}{\text{Tamaño de la imagen comprimida}}$$

En esta práctica se han considerado:

- El tamaño original de la imagen en RGB (8 bits por canal  $\times$  3 canales)
- El tamaño obtenido tras la codificación de entropía mediante DPCM + Huffman

A partir de estos valores, la relación de compresión calculada ha sido:

$$R_c = 5.833$$

Esto significa que el proceso completo de compresión reduce el volumen de información a aproximadamente:

$$\frac{1}{5.833} \approx 17.1\%$$

del tamaño original, es decir, la imagen comprimida ocupa solo una sexta parte de la información inicial.

Un valor de  $R_c \approx 6$  indica que:

- Se ha logrado una compresión significativa,
- Sin provocar una degradación apreciable en la calidad visual,
- Gracias a que la imagen contiene amplias zonas con gradientes suaves (como el cielo y el mar) que son altamente comprimibles mediante DCT y cuantización,
- Mientras que las zonas de mayor detalle (avión y espuma del agua) han sido preservadas razonablemente bien.



Este nivel de compresión es típico de imágenes fotográficas de alta resolución cuando se utiliza una cuantización moderada, obteniendo un equilibrio adecuado entre calidad y tamaño.

La relación de compresión obtenida confirma que el proceso JPEG implementado ha sido eficaz, reduciendo el tamaño del archivo en más de un 80% sin pérdida visual notable. Este resultado valida la eficiencia del sistema DCT + Cuantización + Entropía a la hora de comprimir imágenes reales de alta resolución manteniendo una excelente calidad percibida.

## 9. Relación señal ruido

Para cuantificar la calidad de la imagen reconstruida tras todo el proceso JPEG, se emplean métricas basadas en la diferencia entre la imagen original y la reconstruida. En particular, se utilizan:

- MSE (Mean Squared Error)
- PSNR (Peak Signal-to-Noise Ratio)
- SNR (Signal-to-Noise Ratio)

Estas medidas permiten evaluar numéricamente el nivel de degradación introducido por el sistema de compresión.

Los valores calculados en esta práctica han sido:

- MSE = 8.5757
- PSNR = 38.80 dB
- SNR = 32.54 dB

La calidad de la imagen reconstruida se evaluó mediante las métricas MSE, PSNR y SNR. El valor obtenido de MSE = 8.5757 indica que el error medio entre la imagen original y la reconstruida es muy bajo, especialmente considerando que los niveles de intensidad van de 0 a 255; un MSE inferior a 10 suele asociarse a pérdidas mínimas.

La métrica PSNR = 38.80 dB confirma esta apreciación, ya que valores por encima de 35 dB se consideran de muy buena calidad y por encima de 40 dB prácticamente indistinguibles del original. Con casi 39 dB, la imagen reconstruida mantiene un nivel de detalle muy alto, con diferencias prácticamente imperceptibles para el ojo humano.

Por su parte, la SNR = 32.54 dB muestra que la energía de la señal es muy superior al ruido introducido durante el proceso de compresión, lo que evidencia que las alteraciones generadas por la cuantización y transformaciones JPEG son leves y no afectan significativamente a la percepción visual. En conjunto, estas métricas confirman que la compresión aplicada conserva de forma excelente la calidad de la imagen original.

Las métricas de error y calidad obtenidas demuestran que el proceso de compresión JPEG aplicado mantiene un nivel de fidelidad muy elevado. El MSE bajo y los valores altos de SNR y PSNR indican que las pérdidas introducidas son mínimas y que la imagen reconstruida es prácticamente indistinguible de la original desde el punto de vista visual. Esto confirma la eficacia del esquema DCT + cuantización + codificación de entropía implementado en esta práctica.

## 10. Conclusión

A lo largo de esta práctica se ha implementado y analizado en detalle el proceso completo de compresión JPEG, desde el preprocessado inicial hasta la reconstrucción final de la imagen. La



conversión al espacio de color YCbCr permitió separar la información de luminancia y crominancia, optimizando la compresión al explotar las características del sistema visual humano. La aplicación de la DCT en bloques de  $8 \times 8$  demostró cómo la energía de la imagen se concentra mayoritariamente en las bajas frecuencias, mientras que la cuantización redujo de forma controlada la precisión de los coeficientes menos relevantes, introduciendo pérdidas leves pero prácticamente imperceptibles.

El reordenamiento zig-zag y la codificación de entropía mediante DPCM y Huffman mostraron su eficacia al reducir de forma significativa el tamaño final de los datos, alcanzando una relación de compresión  $R_c = 5.833$ , coherente con la resolución y complejidad de la escena analizada. La reconstrucción final confirmó el correcto funcionamiento de todas las etapas, y las métricas obtenidas,  $MSE = 8.5757$ ,  $PSNR = 38.80\text{ dB}$ ,  $SNR = 32.54\text{ dB}$ , evidenciaron una calidad visual muy alta, con diferencias mínimas respecto a la imagen original.

En conjunto, los resultados demuestran la solidez y eficiencia del estándar JPEG, capaz de lograr una compresión notable manteniendo una fidelidad excelente. La práctica ha permitido comprender de forma aplicada cómo cada etapa contribuye al equilibrio entre reducción de datos y calidad visual, ofreciendo una visión completa y profundizada del proceso real de compresión de imágenes digitales.

## 11. 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/P6>