

Universidad de Murcia

GRADO EN INGENIERÍA INFORMÁTICA

COMPRESIÓN MULTIMEDIA

CURSO 22/23

Proyecto Final

COMPRESIÓN DE IMÁGENES BASADA EN DCT:
ESTUDIO EXPERIMENTAL

Javier Polo Gambín

javier.polog@um.es

Grupo: PCEO

Profesor: FELIX GOMEZ MARMOL

23 de junio de 2023

Índice

Índice de figuras	3
Índice de tablas	4
1 Resumen	5
2 Introducción	6
3 Metodología	7
3.1 Los compresores desarrollados	7
3.2 Obtención de datos experimentales	8
3.3 Comparación cualitativa de los resultados	8
3.4 Comparación cuantitativa de los resultados	9
3.5 Medios empleados para las pruebas experimentales	9
3.5.1 Hardware empleado	9
3.5.2 Software empleado	10
3.6 Imágenes seleccionadas	10
3.6.1 candados	10
3.6.2 lennon	11
3.6.3 lena	13
3.6.4 mandrill	14
3.6.5 noise	15
3.6.6 pattern	16
3.6.7 color_bars	17
3.6.8 explorer	18
3.6.9 gradient	19
3.6.10 color_shapes	20
3.6.11 x-ray	21
4 Resultados experimentales	22
4.1 Estudio cualitativo	22
4.1.1 candados	22
4.1.2 lennon	24
4.1.3 lena	25
4.1.4 mandrill	25
4.1.5 noise	26
4.1.6 pattern	27
4.1.7 color_bars	28
4.1.8 explorer	29
4.1.9 gradient	30
4.1.10 color_shapes	31
4.1.11 x-ray	32
4.1.12 Conclusiones del estudio cualitativo	33
4.2 Estudio cuantitativo	35

4.2.1	candados	35
4.2.2	lennon	37
4.2.3	lena	38
4.2.4	mandrill	39
4.2.5	noise	40
4.2.6	pattern	41
4.2.7	color_bars	42
4.2.8	explorer	43
4.2.9	gradient	44
4.2.10	color_shapes	45
4.2.11	x-ray	46
4.2.12	Conclusiones del estudio cuantitativo	46
4.2.13	Extra: Análisis de SSIM (Structural Similarity Index)	47
5	Conclusión	49

Índice de figuras

1	Valores tomados por el parámetro CaliQ	8
2	candados.bmp	11
3	lennon.bmp	12
4	lena.bmp. Fuente:[7]	13
5	mandrill.bmp	14
6	noise.bmp	15
7	pattern.bmp	16
8	color_bars.bmp	17
9	explorer.bmp	18
10	gradient.bmp	19
11	cshapes.bmp	20
12	x-ray.bmp	21
13	Referencia - candados	23
14	Resultados experimentales - candado	23
15	Referencia - lennon	24
16	Resultados experimentales - lennon	24
17	Referencia - lena	25
18	Resultados experimentales - lena	25
19	Referencia - mandrill	26
20	Resultados experimentales - mandrill	26
21	Referencia - noise	27
22	Resultados experimentales - noise	27
23	Referencia - pattern	27
24	Resultados experimentales - pattern	28
25	Referencia - color_bars	28
26	Resultados experimentales - color_bars	29
27	Referencia - explorer	29
28	Resultados experimentales - explorer	30
29	Referencia - gradient	30
30	Resultados experimentales - gradient	31
31	Referencia - color_shapes	31
32	Resultados experimentales - color_shapes	32
33	Referencia - x-ray	32
34	Resultados experimentales - x-ray	33
35	log(MSE) vs RC - candados	35
36	SNR vs RC - candados	35
37	log(MSE) vs RC - lennon	37
38	SNR vs RC - lennon	37
39	log(MSE) vs RC - lena	38
40	SNR vs RC - lena	38
41	log(MSE) vs RC - mandrill	39
42	SNR vs RC - mandrill	39
43	log(MSE) vs RC - noise	40
44	SNR vs RC - noise	40

45	log(MSE) vs RC - pattern	41
46	SNR vs RC - pattern	41
47	log(MSE) vs RC - color_bars	42
48	SNR vs RC - color_bars	42
49	log(MSE) vs RC - explorer	43
50	SNR vs RC - explorer	43
51	log(MSE) vs RC - gradient	44
52	SNR vs RC - gradient	44
53	log(MSE) vs RC - color_shapes	45
54	SNR vs RC - color_shapes	45
55	log(MSE) vs RC - x-ray	46
56	SNR vs RC - x-ray	46

Índice de tablas

1	Tabla de resultados experimentales	36
2	Tabla de resultados experimentales	37
3	Tabla de resultados experimentales	39
4	Tabla de resultados experimentales	40
5	Tabla de resultados experimentales	41
6	Tabla de resultados experimentales	42
7	Tabla de resultados experimentales	43
8	Tabla de resultados experimentales	44
9	Tabla de resultados experimentales	45
10	Tabla de resultados experimentales	46

1. Resumen

En el siguiente informe se expondrá el trabajo realizado como parte de las prácticas de la asignatura Compresión Multimedia, en la convocatoria de Junio del curso 2022/2023.

Se estudia el desarrollo y pruebas de ejecución de dos compresores de imágenes distintos basados en DCT y codificación Huffman. Se expondrán los principios teóricos necesarios, seguidos de una descripción meticulosa de la metodología seguida en la experimentación. Tras la obtención de los datos experimentales, se procede al análisis pormenorizado de los mismos, realizando tanto un estudio cualitativo como uno cuantitativo. Finalmente, se expondrán las conclusiones obtenidas.

El objetivo de este proyecto será comparar ambos compresores, obtener información detallada sobre sus resultados y factores que puedan influir en estos. Se buscará describir las condiciones en las que se consigue un mejor funcionamiento de los compresores y determinar una configuración de estos (o un conjunto de ellas) en las que se obtienen los mejores resultados de forma generalizada.

2. Introducción

La compresión de imágenes desempeña un papel fundamental en diversos ámbitos en la actualidad, desde el campo de la tecnología y la ciencia hasta acciones cotidianas de nuestra vida diaria.

Con la era digital, la generación y el consumo de imágenes ha crecido exponencialmente. Se envían y reciben más imágenes que en ningún otro momento en la historia, y es aquí donde la compresión de imágenes se vuelve esencial para superar los nuevos desafíos como el almacenamiento y la transmisión eficiente de datos asociados a este consumo masivo. Imaginemos una situación en la que queremos enviar una fotografía de alta resolución a través de una conexión de internet limitada. Sin una técnica de compresión adecuada, la transferencia de esa imagen podría llevar mucho tiempo o incluso ser imposible. Supongamos ahora que un centro médico debe comunicar a otro los resultados de una prueba de resonancia magnética. Estas imágenes suelen tener un tamaño considerable debido a su alta resolución y detalles precisos, que resulta fundamental conservar. Es en casos como estos donde la compresión de imágenes juega un papel crucial al reducir significativamente el tamaño del archivo sin perder información esencial.

Con el desarrollo de la Teoría de la información, las mejoras tecnológicas y la investigación en este campo se han desarrollado diversas técnicas de compresión que buscan cumplir los objetivos anteriores. En este estudio, utilizamos codificadores basados en la técnica de codificación Huffman, empleando dos variaciones de esta técnica

Además de la codificación Huffman, la transformada discreta del coseno (DCT) se ha convertido en un pilar fundamental en la compresión de imágenes. Combinada con la cuantización, permite alcanzar una tasa de compresión significativa sin comprometer demasiado la calidad visual.

En este estudio emplearemos ambas técnicas en conjunto para realizar una aproximación al método JPEG, uno de los más utilizados en la actualidad.

Llevaremos a cabo pruebas experimentales sobre una serie de imágenes cuidadosamente seleccionadas. Estas experiencias nos permitirán valorar las capacidades y los límites de estas técnicas y determinar en qué casos debe usarse y de qué manera para que la compresión sea lo más efectiva posible.

3. Metodología

3.1. Los compresores desarrollados

En el proceso de compresión se hará uso de la transformada discreta del coseno (DCT) [6]. Esto implica la conversión de la imagen a un espacio discreto para reducir su dimensionalidad. Posteriormente, el resultado de esta transformación se somete a cuantización mediante la aplicación de un filtro que elimina ciertos coeficientes, cuya selección depende del valor del parámetro que se denominará **factor de calidad**, en adelante denotado como **CaliQ**. Cuanto mayor sea este factor, mayor será la magnitud de los coeficientes necesaria para preservarse. En este proceso de eliminación, se busca maximizar la cantidad de coeficientes descartados (para aumentar la compresión), preservando en la medida de lo posible la similitud entre la imagen reconstruida y la original.

Debido a la aplicación de la DCT, los compresores utilizados en esta práctica implementan un algoritmo de **compresión con pérdidas**, ya que implica la pérdida de información debido al proceso de cuantización/descuantización. En este tipo de codificaciones, el resultado obtenido difiere del original, por lo tanto, se deberán evaluar las imágenes obtenidas para determinar si los resultados obtenidos son adecuados o no. No tiene sentido lograr una alta compresión de un archivo si no podemos recuperar el archivo original (o algo que percibamos como similar similar).

Para la codificación se empleará codificación Huffman. Esta se basa en la asignación de códigos de longitud variable a los elementos en función de su frecuencia, de forma que emplea un alfabeto de longitud variable. Al utilizar esta codificación, los elementos que aparecen con mayor frecuencia se representan mediante códigos más cortos, mientras que a aquellos menos frecuentes se le asignan códigos más largos. Esta asignación se realiza de manera que no se produzcan ambigüedades en la decodificación de los datos comprimidos. De esta forma, se logra una compresión efectiva al aprovechar las repeticiones y patrones presentes en los datos originales.

Para la construcción del alfabeto de longitud variable pueden emplearse diversas tablas. En nuestro caso emplearemos 2 diferentes:

- Tablas Huffman JPEG estándar de propósito general, determinadas en base a experimentos de percepción visual realizados con observadores humanos, y que se aplican sin variaciones, independientemente de la imagen particular [2].
- Tablas Huffman customizadas, que varían en función de cada imagen, atendiendo a la frecuencia de los valores presentes en la imagen original.

Estas dos tablas en la codificación darán lugar a dos pares de compresor-descompresor. A partir de las tablas Huffman estándar desarrollaremos el compresor-descompresor **default** y a partir de las customizadas el compresor-descompresor **custom**.

Utilizando estos dos compresores se realizarán diferentes pruebas experimentales sobre imágenes seleccionadas, se extraerán los datos pertinentes y se analizarán para describir sus

diferencias, comparar su desempeño y tratar de determinar cuál es la configuración que ofrece mejores resultados de forma general.

3.2. Obtención de datos experimentales

Para obtener los datos experimentales se procede de la siguiente manera:

Las pruebas experimentales de los compresores se realizan sobre imágenes en formato BMP sin compresión. Para cada una se lleva a cabo el proceso de compresión y posterior descompresión utilizando ambos compresores desarrollados, es decir, el compresor por defecto y el compresor customizado (*default* y *custom* respectivamente). Este proceso se repite para cada uno de los siguientes valores del factor de calidad.

$$\text{CaliQ} = [5, 25, 50, 100, 200, 400, 750, 1000]$$

Figura 1: Valores tomados por el parámetro de calidad en la experimentación, descrito en (3.1).

Se ha decidido tomar un número amplio de valores para poder obtener numerosos datos experimentales que permitirán un análisis posterior más exhaustivo.

Esta metodología nos permite examinar y comparar las diferencias entre los dos compresores/descompresores en un amplio rango de resultados, abarcando desde aquellos con una pérdida mínima de información hasta aquellos con una pérdida significativa de información. En general, se espera que los ejemplos con una mayor pérdida de información alcancen las tasas de compresión más altas.

3.3. Comparación cualitativa de los resultados

Uno de los objetivos de este estudio es determinar las circunstancias en las que cada compresor de imagen resulta adecuado. Intuitivamente, consideraremos que el resultado es adecuado cuando presenta un alto grado de parecido visual en comparación con la imagen original. Aunque este criterio no es el único factor a tener en cuenta, como discutiremos más adelante, resulta fundamental para determinar el desempeño de un compresor como “adecuado”. En última instancia, el objetivo es lograr la compresión de una imagen reduciendo su tamaño perdiendo la menor cantidad posible de información.

En este estudio realizaremos un análisis observando los distintos resultados obtenidos para algunos factores de calidad seleccionados, y señalaremos el grado de parecido visual alcanzado y cualquier diferencia notable encontrada. Como resulta evidente, no existe una manera exacta de comprobar el parecido visual entre dos imágenes percibido por un sujeto (es subjetivo), pero sí que podemos analizar detalles como la legibilidad de los textos, la definición de los bordes, el enfoque, etc., lo que nos proporciona una buena aproximación sobre la percepción que se tiene de la imagen.

3.4. Comparación cuantitativa de los resultados

En esta sección calcularemos diferentes métricas para realizar un análisis cuantitativo de los resultados obtenidos. Las métricas que utilizaremos (formuladas para dos imágenes arbitrarias X, Y de elementos x_i, y_j respectivamente) son las siguientes:

Error cuadrático medio (*MSE*, Mean Squared Error): Establece una métrica entre imágenes basada en las diferencias entre el valor de los píxeles individuales de las imágenes. Como mide la diferencia entre imágenes, en principio cuanto menor sea este valor, más similares serán las dos imágenes y viceversa, aunque no siempre es así. La fórmula utilizada es:

$$MSE = \frac{\sum_i^n (x_i - y_i)^2}{n}$$

Ratio de Compresión (*RC*): Compara el tamaño de la imagen comprimida respecto al tamaño de la imagen original. Se calcula utilizando la siguiente fórmula (T =tamaño):

$$RC = \frac{T_{Original} - T_{Final}}{T_{Comprimido}}$$

De manera **opcional**, se calculará también la métrica Signal-to-Noise-Ratio (métrica estudiada en la asignatura), que creemos aportará información complementaria útil para el análisis cuantitativo.

Signal-to-Noise-Ratio (*SNR*): Calcula el ratio entre la cantidad de información codificada en la señal y la cantidad de error (ruido) obtenido. Se utiliza la siguiente fórmula:

$$SNR(\text{db}) = 10 \cdot \log_{10} \left(\frac{\sum x_i^2}{\sum (x_j - y_j)^2} \right)$$

Estos valores se calcularán para cada imagen, utilizando los diferentes valores del factor de calidad, y se aplicarán tanto al compresor estándar como al compresor customizado. De esta manera podremos establecer una métrica de cuánto difiere cuantitativamente el resultado de cada método respecto a la imagen original, lo que nos proporcionará un método matemático formal para comparar los resultados obtenidos mediante los dos métodos de compresión.

Recordamos que estos no son los únicos factores que han de tenerse en cuenta cuando se evalúa el desempeño de un compresor de imágenes. Es bien sabido que dos imágenes pueden diferir cuantitativamente pero tener un gran parecido visual. En principio, buscamos satisfacer ambas condiciones.

3.5. Medios empleados para las pruebas experimentales

3.5.1. Hardware empleado

Es importante indicar las características del hardware empleado para la ejecución de las pruebas experimentales, para permitir la reproducibilidad de los experimentos y porque los resultados pueden variar considerablemente en función de las prestaciones del hardware sobre el que se trabaja (como se indica en el guión de la práctica [4]).

En este caso, la ejecución de los experimentos la he realizado sobre un ordenador con procesador AMD Ryzen™ 7 2700X @ 3.7Ghz , con 16GB de RAM instalada.

3.5.2. Software empleado

Por las mismas razones de antes se indican las versiones concretas del software empleado.

La implementación de los distintos compresores y descompresores la he realizado en lenguaje *Matlab*. Para la ejecución de las pruebas y obtención de los datos he empleado el software *Matlab R2018b*, sobre sistema operativo Microsoft Windows 10 Pro. Para crear las gráficas de los resultados he empleado Microsoft Excel 2019.

3.6. Imágenes seleccionadas

La selección de imágenes para el experimento se ha realizado siguiendo un criterio intencionado. De esta manera, las imágenes abarcan una amplia gama de características, como tamaño, composición, número de elementos, patrones y colores. También se ha tenido en cuenta el potencial de análisis que presenta cada imagen, como se explica más adelante. El objetivo de esta selección es obtener ejemplos representativos que permitan apreciar las características fundamentales de los compresores utilizados y comprender las condiciones en las que su desempeño es mejor y peor.

Todas las imágenes son en formato *BMP* por tratarse de un formato sin pérdidas (no podríamos concluir nada de un formato comprimido con pérdidas, como *jpeg*, porque se estarían comparando las diferencias relativas de dos compresiones consecutivas pero no la compresión del fichero original). En el caso de las imágenes de fuente propia, algunas las he obtenido realizando una captura de pantalla, lo que me permite almacenar el mapa de bits en el formato deseado (en oposición de imágenes descargadas de internet, de las que no siempre puedo asegurar si se encuentran en su formato original). Otras las he generado directamente como mapa de bits usando un código en python o programas de edición de imagen y una de las imágenes la he escaneado del soporte físico directamente en formato *BMP*. También he utilizado algunas imágenes disponibles en conocidas bases de datos de investigación [5].

Como observación, se ha creado un subdirectorio en la carpeta de la entrega para contener los ficheros de imagen, para que todo quede más organizado y facilitar así su análisis y procesamiento. Los ficheros de las distintas imágenes pueden encontrarse en el directorio *|PoloGamin-Javier|source|Images|original*.

3.6.1. candados

Esta imagen la he obtenido del material proporcionado por los profesores de la asignatura en el aula virtual. La he seleccionado debido a su abundancia de elementos detallados, ya que considero que esto permitirá apreciar las diferencias en el nivel de detalles preservados por cada tipo de compresor y factor de calidad. Además, gran parte de la imagen presenta tonos similares, principalmente tonos dorados, los cuales podrían ser aprovechados por el compresor para reducir el tamaño de la imagen.

La imagen es de 480×640 píxeles y ocupa 900KB.



Figura 2: candados.bmp

3.6.2. lennon

He seleccionado esta imagen debido a su fondo uniforme, lo cual permitirá al compresor aprovecharlo y lograr un buen ratio de compresión. Además, la imagen contiene elementos detallados, como las gafas y el micrófono, que resultará interesante analizar en términos de cómo conservan sus detalles en diferentes configuraciones de los compresores.

La imagen es de 1375×1031 píxeles y ocupa 4.05MB. Se trata de una imagen bastante grande, puesto que creo que resultará interesante extraer datos sobre el desempeño de los compresores en imágenes de distintos tamaños (grandes y pequeños).



Figura 3: lennon.bmp

3.6.3. lena

Esta imagen ha sido ampliamente utilizada en el campo de la compresión de imágenes como referencia para evaluar el desempeño de los algoritmos compresores, y es probablemente la imagen más famosa en este campo de investigación.

Es una elección interesante debido a las áreas con alto nivel de detalle, como el cabello y los ojos, y su fondo no uniforme pero relativamente plano. La hemos incluido en nuestro estudio debido a su relevancia histórica y por ser un estándar ampliamente reconocido en la industria.

Como hipótesis inicial, a pesar de ser una imagen en blanco y negro sin transiciones abruptas de color, considero que los resultados obtenidos no serán óptimos debido a la gran cantidad de detalles que contiene. Es probable que no sea posible alcanzar tasas de compresión elevadas sin perder una considerable cantidad de información visual.

La imagen es de 512×512 píxeles y ocupa 768KB.



Figura 4: lena.bmp. Fuente:[7]

3.6.4. mandrill

La siguiente imagen se ha obtenido de la base de datos de imágenes proporcionada por la USC (University of Southern California) [5].

La imagen muestra un primer plano de la cabeza de un mandril, con muchos detalles. Como en el caso anterior, se utiliza frecuentemente como imagen de referencia para evaluar distintos algoritmos de compresión.

La imagen es interesante debido a que presenta una gran variedad de colores (desde colores vivos a colores más apagados, nos permitirá ver cómo de bien se preserva el color), así como de texturas (distintas texturas del pelo muy complejas, textura más plana y regular de la piel, etc.). En definitiva, se trata de una imagen de la que esperamos obtener información significativa sobre el desempeño de los dos compresores.

Tiene unas dimensiones de 512×512 píxeles y ocupa 768KB.

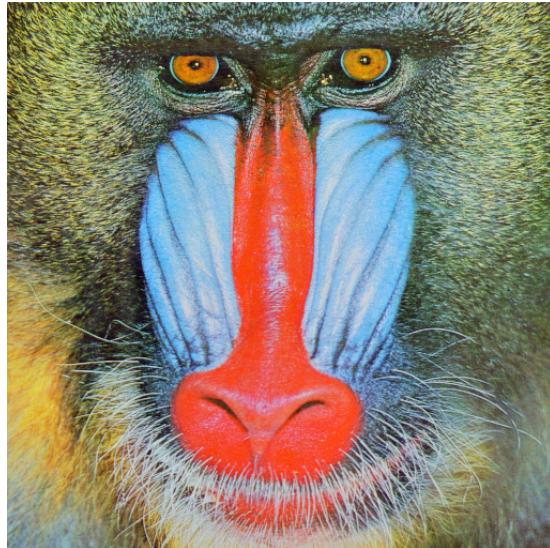


Figura 5: mandrill.bmp

3.6.5. noise

Esta imagen ha sido generada utilizando un sencillo programa en Python, donde los píxeles de colores se han asignado de manera aleatoria. En este caso, se espera que los resultados de la compresión sean muy deficientes, ya que los compresores aprovechan la correlación espacial para lograr una mayor compresión. Sin embargo, en esta imagen, dicha correlación es prácticamente inexistente, lo que implica que para mantener la similitud visual, se requerirán prácticamente todos los coeficientes de la transformada de coseno discreta (DCT), lo que resultará en tasas de compresión extremadamente bajas.

El objetivo de incluir esta imagen es evaluar el rendimiento de cada compresor ante un caso altamente desfavorable. Posteriormente, se compararán estos resultados con los obtenidos en un caso favorable, con el fin de destacar la importancia de la correlación espacial en el rendimiento de estos compresores.

La imagen es de 320×240 píxeles y ocupa 225KB.



Figura 6: noise.bmp

3.6.6. pattern

En el caso de esta imagen, debido a su patrón repetitivo, se espera que la transformada de coseno discreta (DCT) pueda aprovechar la redundancia espacial presente en la imagen. Además, la codificación Huffman podría ser eficiente en la compresión de los elementos más repetitivos dentro del patrón. Con base en esta consideración, mi hipótesis es que esta imagen debería alcanzar altos niveles de compresión sin perder significativamente el parecido visual con la imagen original.

Será interesante realizar una comparación entre los resultados obtenidos en esta imagen y los de la imagen anterior que carecía de correlación espacial. Esto nos permitirá analizar y contrastar los efectos de la correlación espacial en los resultados de compresión (3.6.5).

La imagen es de 320×240 píxeles y ocupa 225KB. La he escogido del mismo tamaño que la anterior para poder comparar ambas mejor.

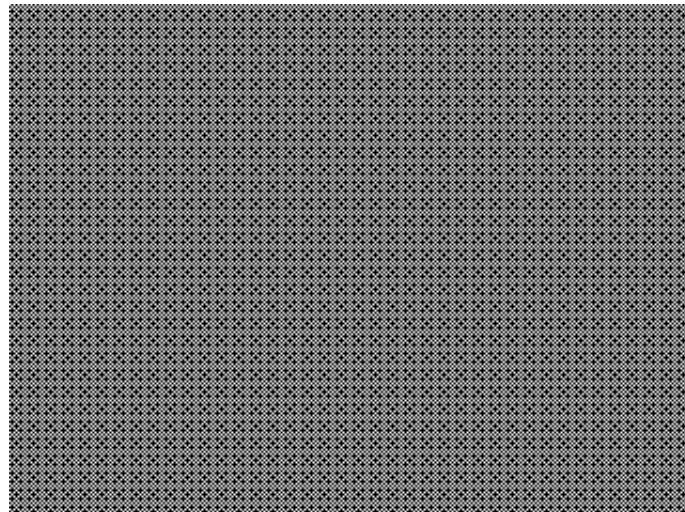


Figura 7: pattern.bmp

3.6.7. color_bars

Cuando una imagen contiene líneas verticales, estas líneas se traducen en coeficientes DCT con valores altos en las frecuencias verticales. En otras palabras, estas líneas generan energía concentrada en las componentes de alta frecuencia vertical de la imagen. Como resultado, muchos coeficientes en esas frecuencias tras realizar la DCT tendrán valores significativos, mientras que los coeficientes en otras frecuencias pueden tener valores cercanos a cero. Además, como solamente hay 8 líneas verticales anchas, los valores deberían concentrarse todavía más en algunos pocos componentes. De esta manera, durante la cuantización se podrá asignar a estos componentes etiquetas más pequeñas, para conseguir resultados de compresión muy buenos.

Mediante este ejemplo, busco evaluar un caso extremadamente favorable para los compresores. Se espera que esta imagen alcance las tasas de compresión más elevadas entre todas las que se están probando, con una diferencia visual prácticamente imperceptible en comparación con la imagen original para la mayoría de los factores de calidad evaluados

La imagen es de 640×480 píxeles y ocupa 900KB.

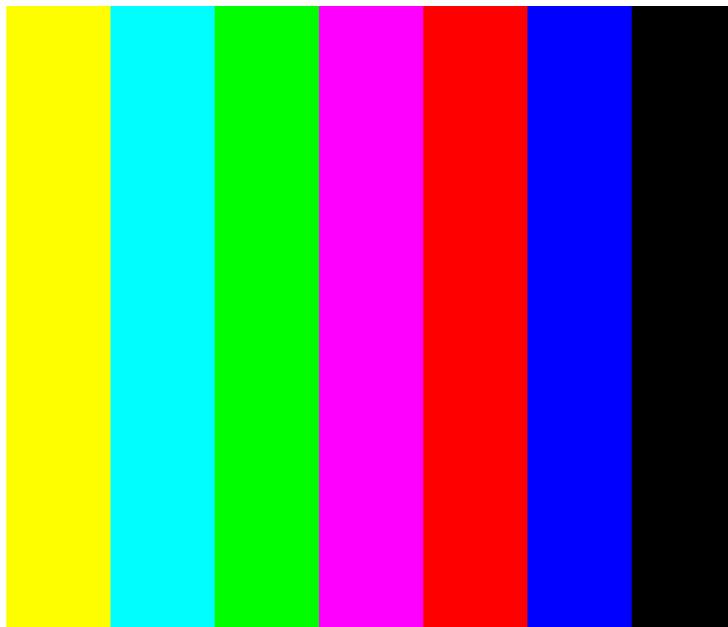


Figura 8: color_bars.bmp

3.6.8. explorer

Esta imagen muestra un fragmento del explorador de archivos de Microsoft Windows 10, que contiene tanto texto como iconos. En nuestro análisis, nos enfocaremos en evaluar si el texto, los botones y otros elementos como los iconos se mantienen legibles y conservan un nivel adecuado de detalle.

Mediante este ejemplo, buscamos examinar el comportamiento de los compresores en una imagen que presenta texto sobre un fondo plano. Las diferencias visuales pueden tener un impacto significativo en la legibilidad del texto, por lo que resulta especialmente interesante estudiar cómo los compresores afectan a esta característica. Además, la imagen también contiene iconos pequeños, los cuales deseamos preservar con un alto nivel de detalle.

A primera vista, el fondo plano nos podría llevar a pensar que se obtendrán buenos resultados de compresión. Sin embargo, debido a la fuerte condición de que el texto debe ser legible, resulta difícil realizar una estimación precisa sobre los resultados obtenidos.

La imagen es de 720×320 píxeles y ocupa 675KB.

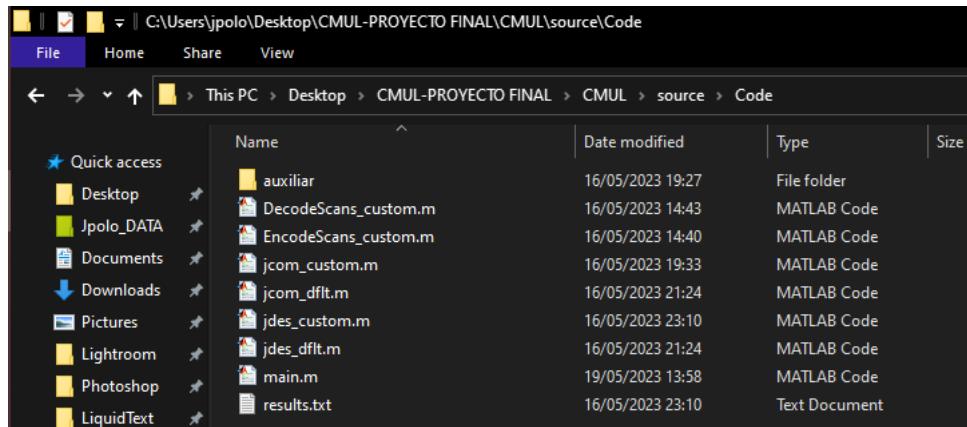


Figura 9: explorer.bmp

3.6.9. gradient

En la imagen vemos una progresión de colores que se produce de forma suave y sin cambios bruscos entre zonas de distinto color.

Creo que esta característica es relevante e interesante, ya que se encuentra comúnmente en muchas fotografías del mundo real, como los tonos del cielo o los elementos grandes que presentan distintos tonos de color según la iluminación. En estas situaciones, los cambios de color se producen de manera gradual en lugar de brusca. Como hipótesis inicial, creo que se irán formando distintas regiones del mismo color, es decir, que los colores se “unifiquen” en diferentes regiones. Sin embargo, debido a que el cambio de color es suave, considero que se podrán obtener buenos resultados de compresión sin sacrificar calidad visual.

Mediante esta imagen, tenemos como objetivo analizar el fenómeno que ocurre al comprimir un patrón de gradación de color de forma aislada, sin la presencia de otros elementos en la imagen que puedan condicionar este análisis. Los resultados obtenidos en este análisis nos permitirán extrapolarlos posteriormente a otras imágenes que contengan regiones de color con características similares.

La imagen es de 300×300 píxeles y ocupa 263KB.



Figura 10: gradient.bmp

3.6.10. color_shapes

En esta imagen, podemos observar grandes zonas que comparten el mismo color, separadas por fronteras diagonales. Considero que esta imagen resulta interesante para analizar el comportamiento de los compresores con respecto a las líneas diagonales que delimitan las diferentes regiones. Esto se debe a que, al eliminar las componentes menos significativas durante el proceso de transformada de coseno discreta (DCT), se puede perder detalle visual, lo que podría ocasionar que las líneas diagonales presenten una apariencia “escalonada”.

Basado en esta consideración, planteo la hipótesis de que los cambios drásticos entre las regiones en la imagen provocarán una pérdida significativa de calidad visual en las líneas diagonales, incluso con factores de compresión bajos. En consecuencia, se espera que los resultados de la compresión no sean muy buenos en este caso.

La imagen es de 320×240 píxeles y ocupa 225KB.

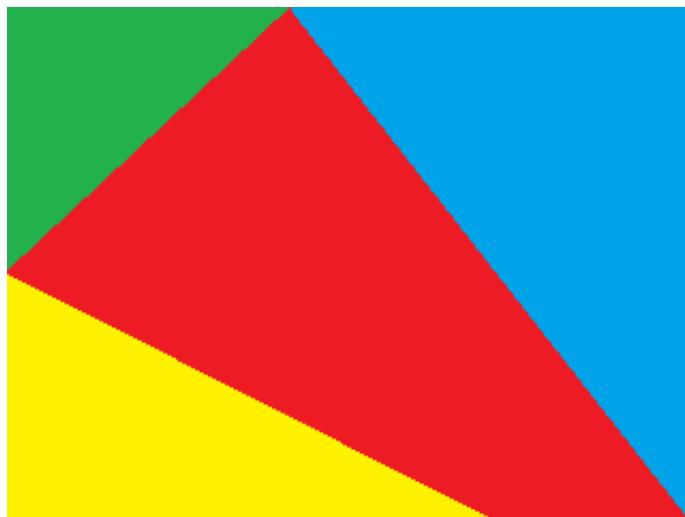


Figura 11: cshapes.bmp

3.6.11. x-ray

La siguiente imagen es una radiografía que he digitalizado yo mismo.

Como planteaba en la introducción, uno de los ámbitos donde la compresión de imágenes eficaz puede tener mucha importancia es en el ámbito médico-sanitario.

En este ejemplo destacan por un lado la gran cantidad de detalles de los huesos que deben preservarse de la mejor manera posible, y por otro la información complementaria y la escala de medida indicadas como texto. Además, las imágenes de los huesos presentan un regiones con color uniforme y cambios suaves en contraste con la región negra prácticamente uniforme del fondo.

De esta manera, la imagen combina muchas de las características interesantes para el análisis mencionadas en las imágenes anteriores en un caso de ejemplo que podemos considerar bastante realista.

La imagen es de 1169×654 píxeles y ocupa 2,18MB

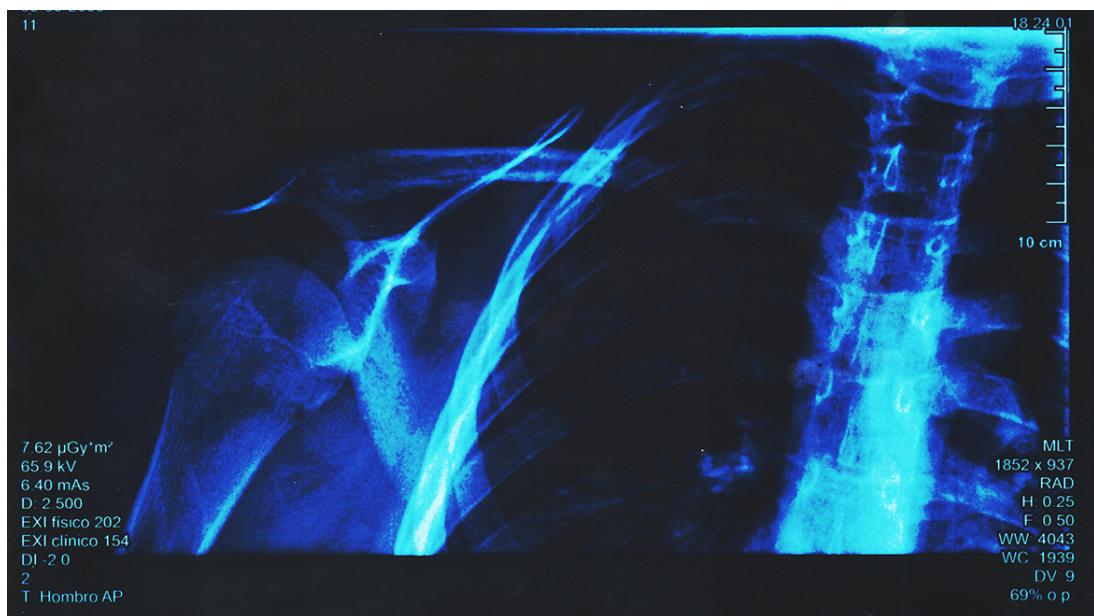


Figura 12: x-ray.bmp

4. Resultados experimentales

En esta sección, analizaremos los resultados obtenidos a partir de los experimentos realizados. Comenzaremos con el estudio cualitativo para obtener información sobre el grado de parecido percibido entre la imagen comprimida y la original, según las consideraciones expuestas en el desarrollo de la metodología (3.3).

Posteriormente, realizaremos el estudio cuantitativo, el cual se basa en las métricas definidas en la sección 3.4. Calcularemos el Ratio de Compresión (RC), el Signal-to-Noise Ratio (SNR) y el Error Cuadrático Medio (MSE) para cada imagen y para ambos compresores utilizados. Estas métricas nos proporcionarán una medida cuantitativa del desempeño de los compresores y nos permitirán comparar los resultados obtenidos.

Es importante tener en cuenta que, aunque el análisis cuantitativo nos brinda información objetiva, no debemos descartar la importancia del análisis cualitativo. Ambos enfoques son complementarios y nos ayudarán a evaluar de manera integral los resultados obtenidos en nuestros experimentos.

4.1. Estudio cualitativo

Debido a la cantidad considerable de datos generados durante los experimentos, hemos realizado un análisis preliminar de los resultados y decidido focalizar nuestra discusión en los valores del parámetro de calidad 50, 100 y 200. Estos valores representan niveles de compresión baja, media y alta, respectivamente y consideramos que proporcionan una visión representativa de los resultados obtenidos. No obstante, en casos específicos donde sea relevante, también mencionaremos otros valores.

Cabe destacar que el análisis se realiza sobre los ficheros de las imágenes. En este documento simplemente se muestran en un tamaño reducido con fines orientativos, se recomienda consultar los ficheros correspondientes para ver mejor los resultados. Los ficheros se encuentran (organizados en subdirectorios para facilitar su organización) en `|PoloGambin-Javier|source|Images|decoded_dflt` (compresor por defecto), y en `|PoloGambin-Javier|source|Images|decoded_custom` (compresor customizado).

4.1.1. candados

Primero se muestra la imagen original para tener una referencia:



Figura 13: Imagen original de referencia.

Imágenes resultantes:



(a) CaliQ=50.



(b) CaliQ=100.



(c) CaliQ=200.

Figura 14: Resultados experimentales de la compresión de la imagen *candados.bmp*.

En la figura 13, vemos la imagen original que servirá de referencia. En la figura 14 vemos las imágenes obtenidas como resultado de la compresión. Este formato será el mismo para el resto de imágenes.

Centrándonos en la imagen, con factor de calidad 50 no se aprecian diferencias muy significativas respecto a la original, como diferencia más notable se aprecia la textura de los candados, que empieza a verse más pixelada, así como los bordes de las marcas y las letras. En cualquier caso estos detalles son muy sutiles.

Con factor de calidad 100 las diferencias anteriores se acentúan de forma moderada, y se ven claramente zonas en las que se forman “manchas” de píxeles. Esto es más notable en las zonas que no tienen un color uniforme. Las letras escritas sobre el candado grande (azul sobre fondo blanco) se distinguen pero con dificultad.

En el factor de calidad 200 las diferencias son muy grandes. La imagen adquiere una textura pixelada (se aprecian grandes bloques de píxeles que componen la imagen). Hay además fallos de color, porque aparecen colores y tonalidades donde antes parecía un color casi uniforme, y las letras no son distinguibles de ninguna manera. Consideramos que este

resultado no es ya adecuado visualmente.

En este caso considero que el factor de calidad 100 es el más adecuado, puesto que aunque presenta más diferencias que el factor 50, la imagen se sigue viendo natural y los elementos como las letras continúan siendo visibles.

4.1.2. lennon

Primero se muestra la imagen original para tener una referencia:



Figura 15: Imagen original de referencia.

Imágenes resultantes:



(a) CaliQ=50.



(b) CaliQ=100.



(c) CaliQ=200.

Figura 16: Resultados experimentales de la compresión de la imagen *lennon.bmp*.

En esta imagen no aprecio diferencias muy notables entre los resultados con factor 50 y con factor 100. Los elementos principales de la imagen se siguen apreciando prácticamente igual, quizás debido a que la imagen original no tenía una gran cantidad de enfoque y una textura granulosa característica de las imágenes analógicas más antiguas (cuando más se notan las diferencias es cuando la imagen original tiene un enfoque muy nítido y texturas muy regulares, porque es de lo primero que se va perdiendo al comprimir la imagen). El fondo de la imagen empieza a unificarse en bloques del mismo color, pero al ser muy regular en la imagen original apenas sí se aprecia. Con factor 200 la imagen empeora notablemente. Adquiere una textura pixelada , aparecen manchas en la cara y el fondo se ve como grandes manchas de píxeles con una transición abrupta entre regiones. Elementos como el micrófono han perdido casi todos sus detalles, adquiriendo una textura plana.

Por estas razones escogeré el factor 100 como el más adecuado.

4.1.3. lena

Primero se muestra la imagen original para tener una referencia:



Figura 17: Imagen original de referencia.

Imágenes resultantes:



(a) CaliQ=50.



(b) CaliQ=100.



(c) CaliQ=200.

Figura 18: Resultados experimentales de la compresión de la imagen *lena.bmp*.

En esta imagen los resultados son bastante buenos, y la diferencia entre el factor 50 y el factor 100 es mínima, apenas se pierden detalles respecto a la imagen original. Con factor 200 la imagen adquiere la textura pixelada que ya hemos comentado en las imágenes anteriores y en el fondo aparecen manchas uniformes muy apreciables.

En este caso elegiré también el factor 100.

4.1.4. mandrill

Primero se muestra la imagen original para tener una referencia:

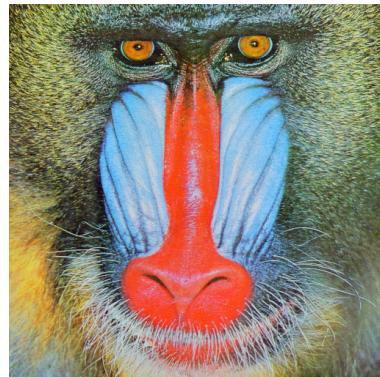
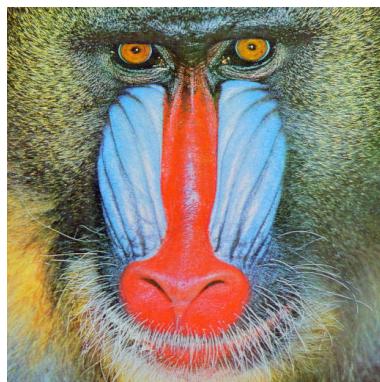
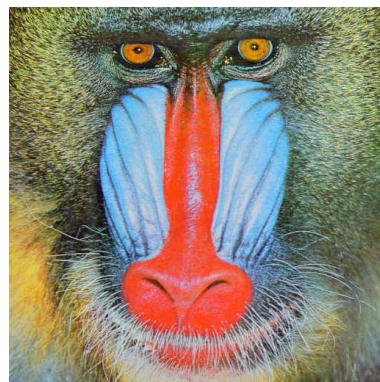


Figura 19: Imagen original de referencia.

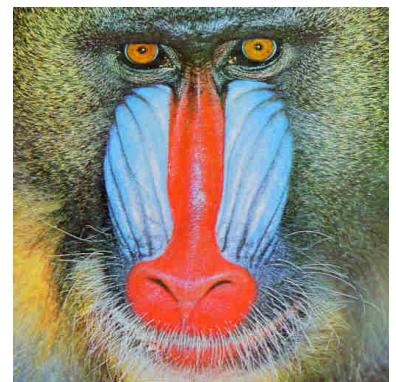
Imágenes resultantes:



(a) CaliQ=50.



(b) CaliQ=100.



(c) CaliQ=200.

Figura 20: Resultados experimentales de la compresión de la imagen *mandrill.bmp*.

En esta imagen enseguida se aprecia la pérdida de calidad incluso en los niveles bajos de compresión. Con CaliQ=100, las texturas de la imagen pierden mucho detalle, especialmente en las zonas del pelo en las esquinas, y la imagen se ve en general “desenfocada” respecto a la original. Los colores, sin embargo, se preservan sorprendentemente bien, incluso en las zonas con cambios bruscos de color.

En este caso elegiré el factor CaliQ=50, por la pérdida de detalle en niveles superiores.

4.1.5. noise

Primero se muestra la imagen original para tener una referencia:

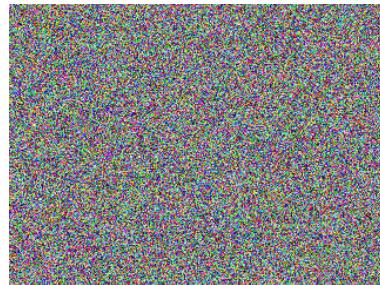
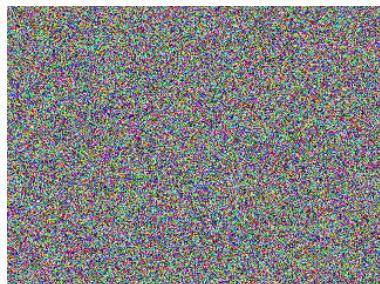
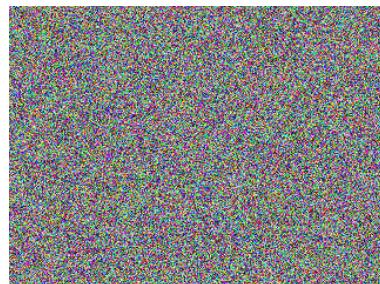


Figura 21: Imagen original de referencia.

Imágenes resultantes:



(a) CaliQ=50.



(b) CaliQ=100.



(c) CaliQ=200.

Figura 22: Resultados experimentales de la compresión de la imagen *noise.bmp*.

Esta imagen es un poco difícil de analizar debido a que no hay formas ni contornos reconocibles. No aprecio ninguna diferencia entre el factor 50 y el factor 100. En el factor 200 ya noto como la imagen se agrupa en bloques de píxeles bien diferenciados. Será interesante ver las diferencias cuantitativas entre los dos primeros factores de calidad.

Elijo el factor 100.

4.1.6. pattern

Primero se muestra la imagen original para tener una referencia:

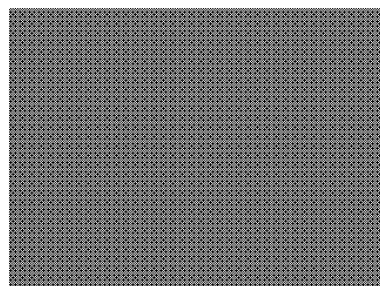


Figura 23: Imagen original de referencia.

Imágenes resultantes:

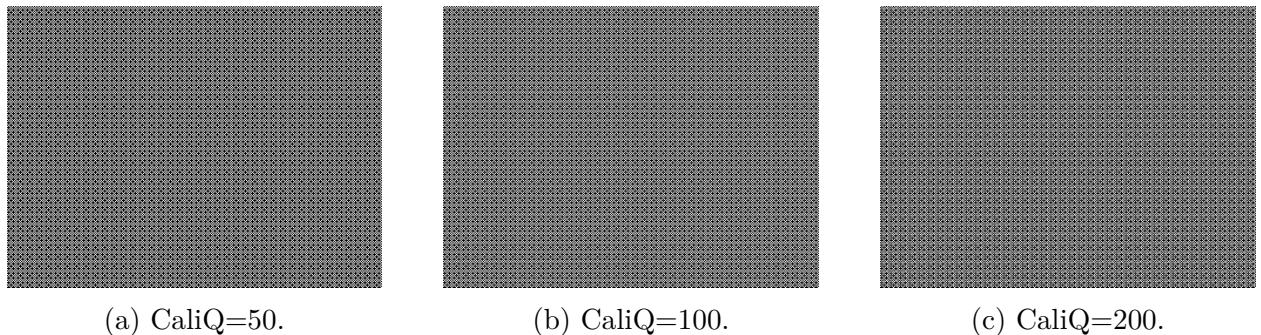


Figura 24: Resultados experimentales de la compresión de la imagen *pattern.bmp*.

Con el factor 50 vemos que el patrón de la imagen se mantiene casi inalterado (salvo algunos píxeles que adquieren una tonalidad de gris casi inapreciable). Con factor 100 aparecen ciertas tonalidades de gris de forma regular en algunas zonas del patrón. Con factor 200 aparte de lo anterior muchos píxeles blancos adquieren distintas tonalidades de gris. Mientras que se aprecian los elementos más importantes del patrón original, en general se diferencia bastante.

Es interesante destacar que, debido a la presencia de un patrón que se repite periódicamente, los cambios respecto a la imagen original se producen de forma uniforme dentro de cada elemento constitutivo base del patrón. Es decir, los cambios son “locales” a nivel de cada elemento del patrón.

En este caso escogeré el factor 100, puesto que visualmente no aprecio mucha diferencia respecto a 50.

4.1.7. color_bars

Primero se muestra la imagen original para tener una referencia:

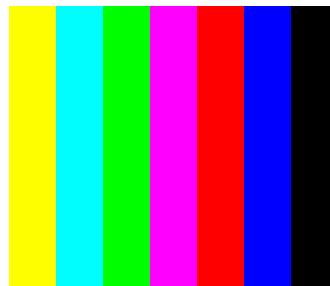


Figura 25: Imagen original de referencia.

Imágenes resultantes:

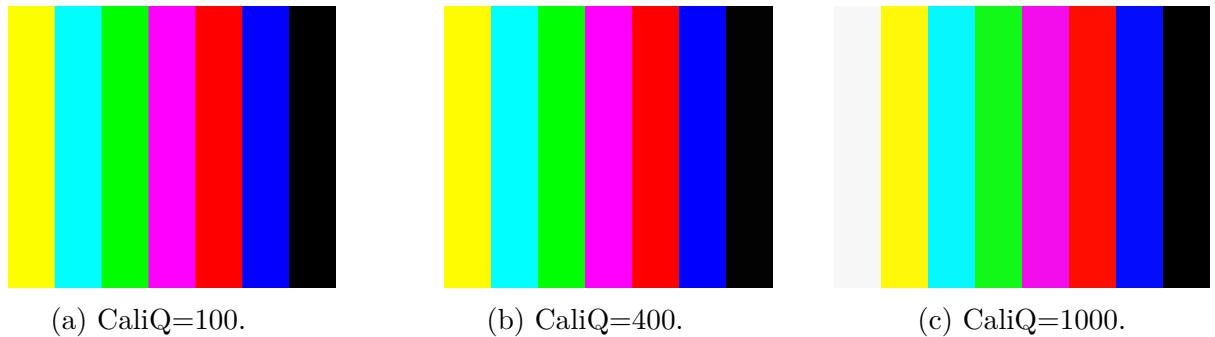


Figura 26: Resultados experimentales de la compresión de la imagen *color_bars.bmp*.

Como hemos señalado en secciones anteriores, este caso es especial puesto que estamos en un caso muy favorable para estos compresores que estamos utilizando. Por esa razón analizaremos en este caso factores de calidad mucho más altos.

Con factor 100 las diferencias son inexistentes, pero es que sucede lo mismo con el factor 400. De hecho, al pasar al factor 1000 la única diferencia notable la encontramos en la tonalidad de la primera barra, la de color blanco, que adquiere un tono grisáceo. Es muy destacable que, pese a los cambios bruscos de color entre las distintas regiones, los contornos no se han visto nada modificados, ni en el factor de calidad más elevado. Esto concuerda con nuestra hipótesis previa basada en el funcionamiento del compresor DCT.

A falta de analizar los resultados cuantitativos, podemos decir que la calidad visual que se preserva es increíble. Por estas razones, escogeré el factor 400.

4.1.8. explorer

Primero se muestra la imagen original para tener una referencia:

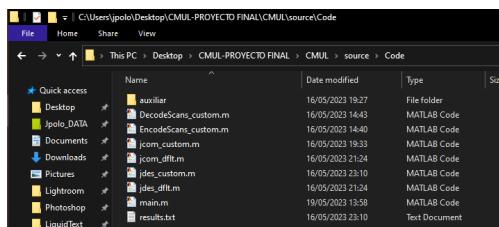


Figura 27: Imagen original de referencia.

Imágenes resultantes:

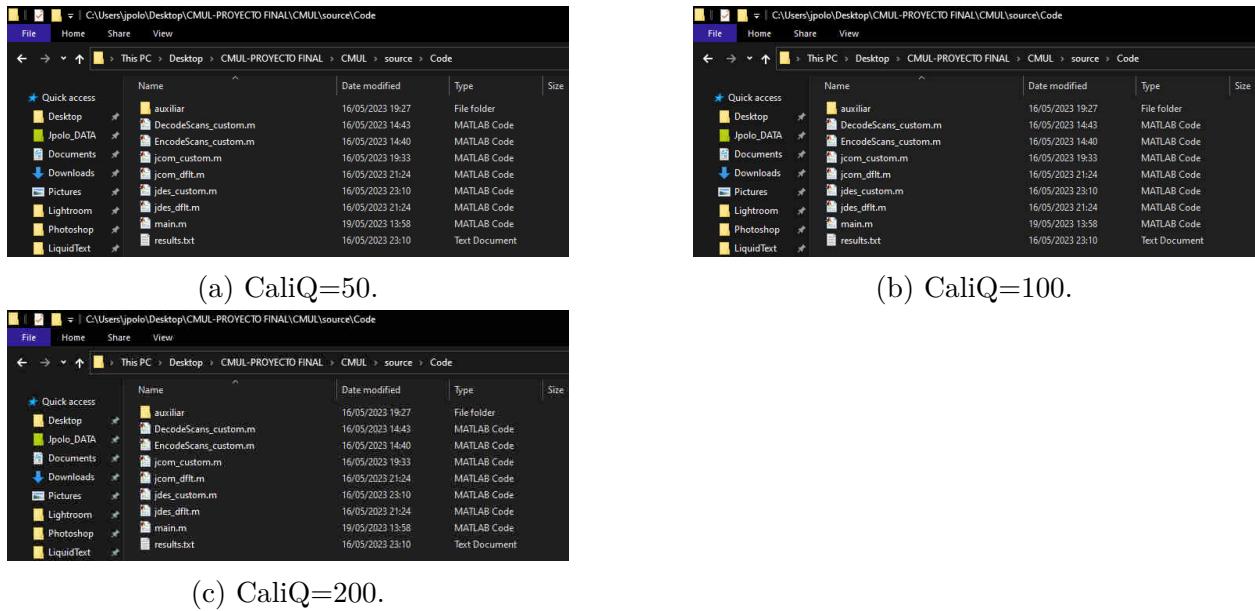


Figura 28: Resultados experimentales de la compresión de la imagen *explorer.bmp*.

Con factor 50 vemos ya cómo el texto se emborrona y aparece una especie de “halo de píxeles” alrededor de las letras. El fondo se preserva plano, pero en los elementos como el texto o los iconos se aprecia bastante este efecto que puede dificultar la lectura. Con factor 100 se acentúa este efecto y algunas líneas como la de la barra de direcciones se empiezan a ver dobles. Las palabras podrían ser difícilmente legibles para personas con algún problema de visión. Con factor 200 cuesta distinguir algunas palabras, y elementos gráficos como la chincheta de la barra de acceso rápido son totalmente indistinguibles.

Debido a que aquí lo importante es la legibilidad del texto, y vemos que empeora rápidamente, me quedaré con el factor 50.

4.1.9. gradient

Primero se muestra la imagen original para tener una referencia:



Figura 29: Imagen original de referencia.

Imágenes resultantes:



(a) CaliQ=50.

(b) CaliQ=100.

(c) CaliQ=200.

Figura 30: Resultados experimentales de la compresión de la imagen *gradient.bmp*.

En este caso vemos cómo los resultados coinciden en gran medida con la hipótesis planteada. De la transición gradual y suave de la imagen original se pasa a transiciones de color cada vez más abruptas en regiones escalonadas del mismo color. Es decir, colores parecidos se agrupan en una sola región de un color único, y el tamaño de esta región crece conforme lo hace Q.

Con factor 50 la transición sigue siendo suave y casi imperceptible. Con factor 100 se empieza a ver una textura un poco pixelada y una transición un poco menos suave, pero el efecto se nota poco. Con factor 200 las transiciones son muy abruptas y el resultado no es adecuado visualmente.

Elegiré en este caso el factor 100.

4.1.10. color_shapes

Primero se muestra la imagen original para tener una referencia:

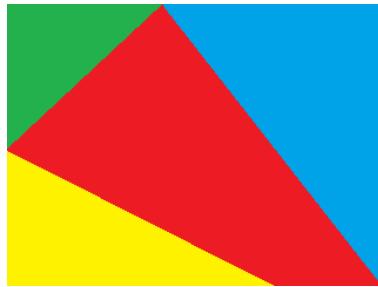


Figura 31: Imagen original de referencia.

Imágenes resultantes:

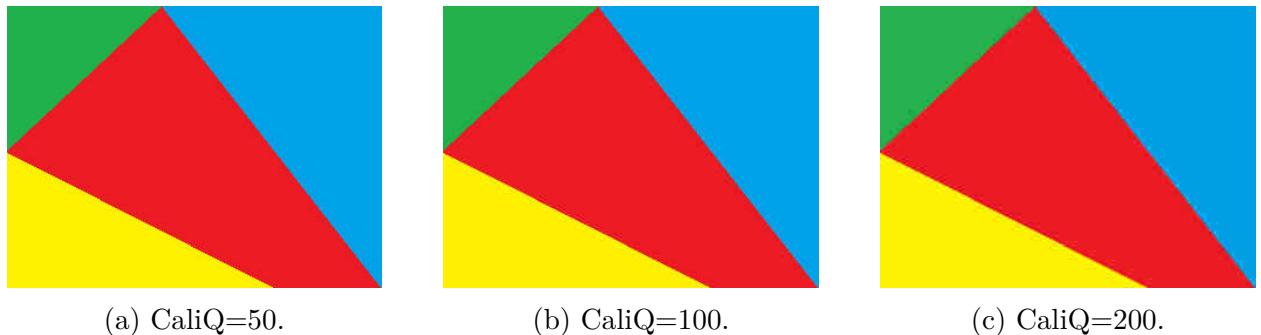


Figura 32: Resultados experimentales de la compresión de la imagen *cshapes.bmp*.

En este ejemplo la clave es ver cómo se comportan los bordes de unión entre regiones de distinto color. El problema principal es que las líneas de unión son diagonales. En las imágenes resultantes vemos cómo los bordes se encuentran difuminados, tanto en la forma como en los colores, que se mezclan y aparecen tonalidades nuevas que no existían en la imagen original. También se pueden apreciar claramente los bloques que conforman la imagen en las líneas diagonales.

Sin embargo, al tratarse de una imagen de tamaño pequeño el efecto no es tan apreciable y considero que con el factor CaliQ=200 todavía es un resultado aceptable, puesto que el resto de la imagen se conserva perfectamente.

4.1.11. x-ray

Primero se muestra la imagen original para tener una referencia:

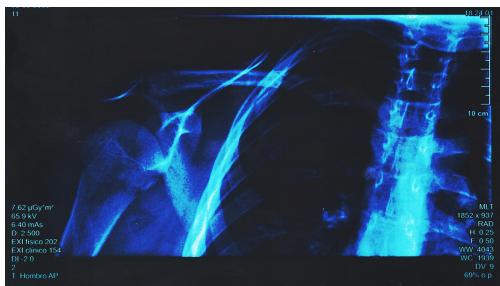


Figura 33: Imagen original de referencia.

Imágenes resultantes:

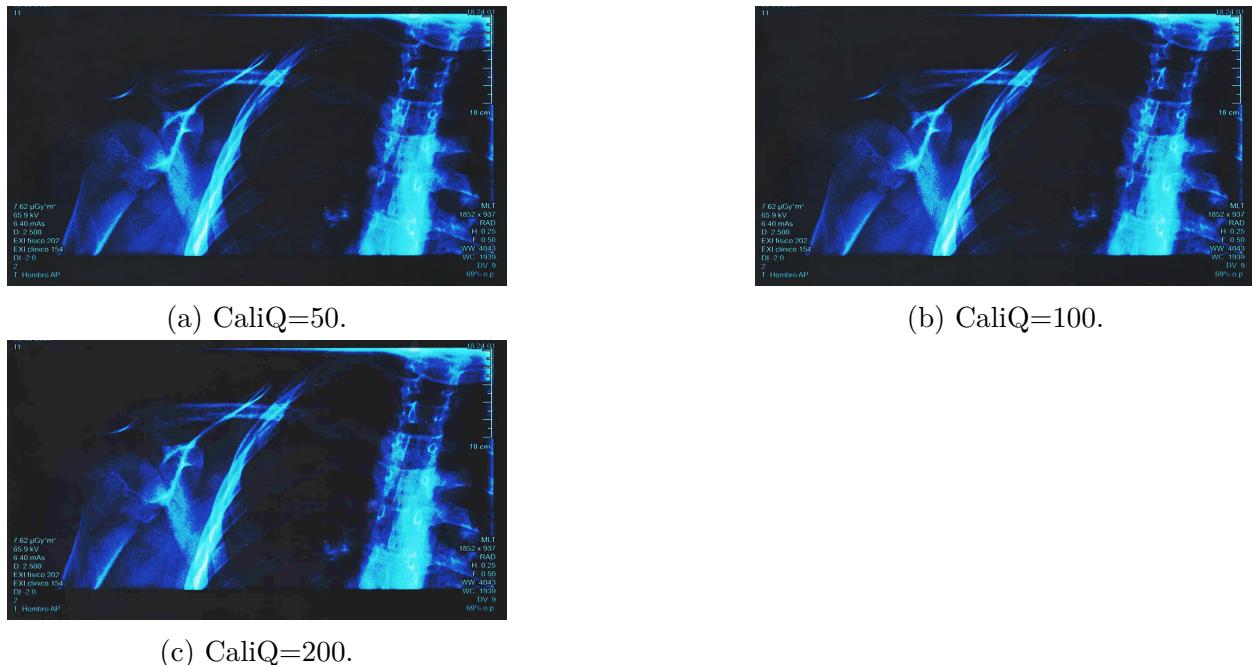


Figura 34: Resultados experimentales de la compresión de la imagen *x-ray.bmp*.

En esta imagen vemos como con factor CaliQ=100 la imagen aparece bastante pixelada, sobre todo el fondo negro. Esto podría dificultar la interpretación de la radiografía y dar lugar a resultados erróneos (aparecen cosas que no están y viceversa), lo que es totalmente inasumible en este contexto. Por esta razón escogeré el factor más bajo de los propuestos, en este caso CaliQ=50, porque las diferencias que presenta respecto a la original no son tolerables en cualquier otro factor de compresión.

Como observación, el texto y la escala de medida se conservan relativamente bien, posiblemente porque las letras no son demasiado pequeñas.

Para CaliQ=200 el fondo se encuentra totalmente pixelado y los detalles más finos de los huesos se pierden.

4.1.12. Conclusiones del estudio cualitativo

Se ha comprobado cómo los resultados considerados “adecuados” varían en las diferentes imágenes, no existe un solo criterio único que sea aplicable a todas.

Sin embargo, sí que se pueden extraer algunas conclusiones:

El análisis parece indicar que para las imágenes con mayor cantidad de detalles pequeños y diferentes, especialmente elementos que deben ser legibles (explorer, mandrill, x-ray) solo se obtienen buenos resultados si se consideran valores de CaliQ bajos (50 en este caso). Por otro lado, para una imagen que podemos considerar promedio (candados, lennon, lena, gradient) el factor 100 ofrece buenos resultados. Hay que destacar que en casi ningún caso se ha considerado el factor 200 como adecuado, debido a que la pérdida de información con

este factor es muy notable.

Es importante recordar que este análisis es plenamente subjetivo, y por tanto podría variar según el observador. También hay que destacar que cada imagen se ha analizado en su contexto, dando importancia por ejemplo a la legibilidad del texto si es que lo había o a la importancia de mantener detalles en la radiografía mostrada en *x-ray*.

De esta manera, el análisis indica que el valor 100 del factor de calidad CaliQ es el que ofrece mejores resultados visuales subjetivos en la mayoría de las situaciones.

4.2. Estudio cuantitativo

Los resultados obtenidos del estudio cualitativo parecen indicar que un valor adecuado para el factor de calidad podría ser el valor 100.

En el estudio cuantitativo, buscaremos responder esta misma pregunta, determinándolo de manera objetiva un valor adecuado para el factor de calidad basándonos en las métricas MSE, RC y SNR mencionadas anteriormente (3.4), con ayuda de distintas gráficas y tablas con los resultados experimentales obtenidos. Comenzaremos discutiendo los resultados obtenidos para cada imagen en particular y luego intentaremos extraer una conclusión global.

Es importante recordar que las métricas cuantitativas nos brindan una medida objetiva de la calidad de compresión, pero no siempre reflejan completamente el parecido visual entre la imagen original y la comprimida.

Como observación, también calculamos para cada imagen la métrica SSIM, **esto es extra** y se explicará en profundidad en su sección correspondiente (4.2.13).

Además, como aspecto **extra**, se incluye en la entrega el directorio `\PoloGambin-Javier\source\Data` con los datos y las gráficas en formato de hoja de cálculo para complementar este informe y facilitar su visualización en caso de que fuese necesario.

4.2.1. candados

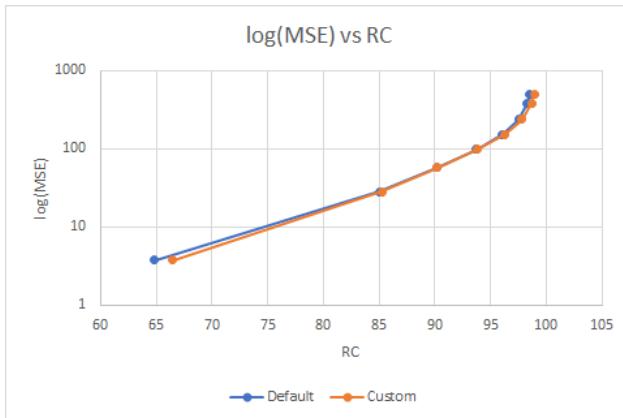


Figura 35: $\log(\text{MSE})$ vs RC - candados

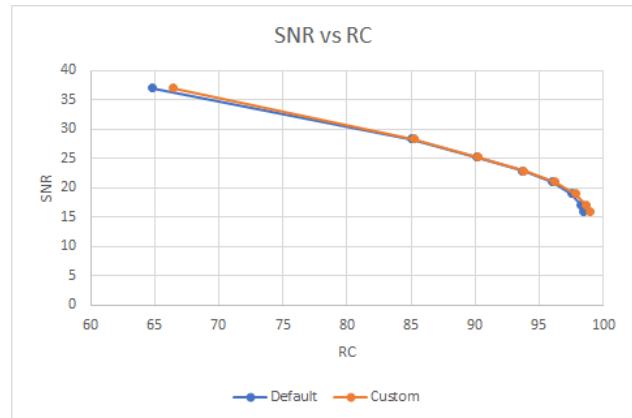


Figura 36: SNR vs RC - candados

DEFAULT					CUSTOM			
MSE	RC	SNR	SSIM	CaliQ	MSE	RC	SNR	SSIM
3.802	64.847	37.086	0.9957	5	3.802	66.491	37.086	0.9957
28.584	85.011	28.325	0.97449	25	28.584	85.228	28.325	0.97449
57.307	90.127	25.304	0.95432	50	57.307	90.225	25.304	0.95432
98.405	93.7	22.956	0.92719	100	98.405	93.791	22.956	0.92719
153.45	96.061	21.027	0.89121	200	153.45	96.265	21.027	0.89121
240.54	97.506	19.074	0.83544	400	240.54	97.827	19.074	0.83544
385.67	98.271	17.024	0.75271	750	385.67	98.662	17.024	0.75271
501.02	98.498	15.888	0.69149	1000	501.02	98.922	15.888	0.69149

Tabla 1: Tabla de resultados experimentales

En la figura 35 se muestran los valores del MSE en función de los valores de RC. Cada par de valores (MSE,RC) es el que se ha obtenido como resultado de compresión con un valor determinado de calidad CaliQ. Como resulta esperable, cuanto mayor es el valor de CaliQ, mayor es RC (se conservan menos detalles por tanto la tasa de compresión crece, como cabría esperar), por esa razón basta considerar el par de valores descrito, entendiendo que el primer punto se corresponde con lo obtenido para el primer valor de CaliQ, el segundo con el segundo valor y así con todos los demás (esto nos permitirá visualizar los resultados más fácilmente en una gráfica bidimensional, en vez de una de 3 dimensiones).

Además, para representar el MSE se ha empleado escala logarítmica, debido a que los valores crecen muy rápidamente y dificultaría visualizarlo en escala decimal usual.

Una vez realizadas las consideraciones iniciales, vamos a analizar las gráficas obtenidas.

Lo primero que podemos observar es que para un mismo factor de calidad, los valores de MSE obtenidos son iguales. Esto resulta esperable si tenemos en cuenta que el MSE cuantifica cuánto se aleja numéricamente una imagen de la imagen original y que cuando se usa un mismo factor de calidad, se descartan los mismos coeficientes de la imagen en el espacio de las frecuencias.

Analizaremos por tanto cuál de los dos algoritmos tiene una mayor relación de compresión (RC) para un mismo nivel de pérdida de información. Visualmente, será mejor la gráfica que esté más a la derecha de las dos.

En este primer ejemplo, ambos compresores obtienen niveles de compresión prácticamente iguales para cada valor de calidad, con el compresor custom siendo un poco mejor en niveles muy bajos o muy altos de compresión. Observamos cómo a partir del cuarto punto (CaliQ=100) el MSE empieza a crecer muy rápido superando el valor 100 (recordemos que estamos en escala logarítmica), por lo que se espera que a partir de este valor el resultado empeore rápidamente.

La gráfica 36 muestra los pares (SNR,CR) para cada valor de calidad CaliQ (el orden de los puntos se corresponde con el valor de CaliQ, como en el caso anterior). Igual que en el caso anterior, para cada valor de calidad, los valores de SNR son los mismos en ambos compresores, de forma que buscaremos cuál es el compresor que obtiene mayor RC para un mismo SNR (más a la derecha visualmente). Esta gráfica tendrá pendiente negativa, puesto que cuando aumenta la pérdida de información, aumenta el RC pero también aumenta el ruido (y por tanto disminuye SNR).

En esta imagen vemos también cómo ambos compresores tienen un desempeño casi idéntico en función del SNR. Se puede observar cómo a partir del quinto punto (CaliQ=200) el valor empieza a decrecer de forma más rápida, empeorando rápidamente a partir de este punto.

Observamos cómo parece haber una relación entre el MSE y el SNR en términos de a partir de qué valor de calidad la imagen comienza a empeorar exponencialmente. Con el resto de imágenes comprobaremos si esta relación se mantiene.

4.2.2. lennon

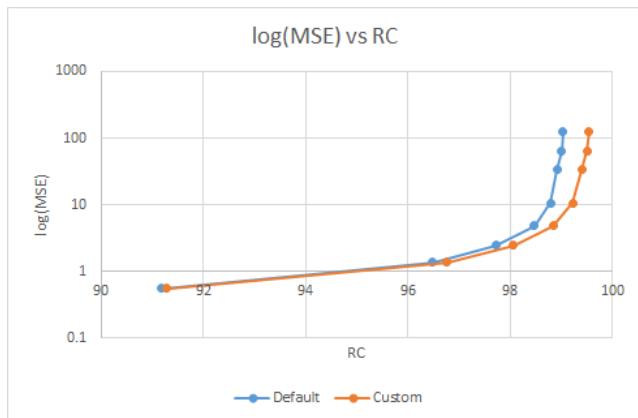


Figura 37: log(MSE) vs RC - lennon

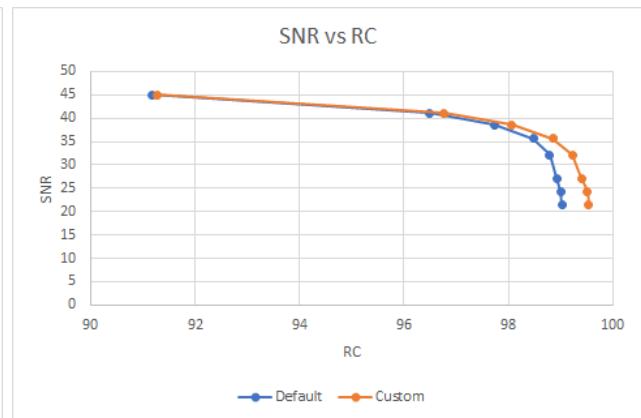


Figura 38: SNR vs RC - lennon

DEFAULT					CUSTOM			
MSE	RC	SNR	SSIM	CaliQ	MSE	RC	SNR	SSIM
0.54644	91.185	45.034	0.99751	5	0.54644	91.275	45.034	0.99751
1.3686	96.489	41.047	0.99426	25	1.3686	96.763	41.047	0.99426
2.4305	97.726	38.553	0.99048	50	2.4305	98.059	38.553	0.99048
4.7607	98.472	35.633	0.98233	100	4.7607	98.844	35.633	0.98233
10.533	98.787	32.184	0.96606	200	10.533	99.229	32.184	0.96606
33.501	98.922	27.159	0.9269	400	33.501	99.405	27.159	0.9269
63.916	99.005	24.354	0.86175	750	63.916	99.507	24.354	0.86175
127.42	99.03	21.357	0.80121	1000	127.42	99.538	21.357	0.80121

Tabla 2: Tabla de resultados experimentales

En la primera gráfica vemos que el MSE crece exponencialmente a partir del tercer punto (CaliQ=50), pero se mantiene bastante bajo, superando el valor 100 de MSE únicamente en el último punto. Esto nos podría indicar que para esta imagen podríamos aplicar valores de CaliQ elevados (CaliQ=200 o incluso CaliQ=400) sin tener mucha pérdida de información, pero apenas se mejora el RC.

Lo mismo ocurre para el SNR, donde la pendiente negativa aumenta tras los primeros puntos pero no alcanza valores muy bajos (por encima de 25 en todos los puntos menos el último).

En ambos casos la gráfica del compresor custom se encuentra más a la derecha, obteniendo resultados bastante mejores que el default.

4.2.3. lena

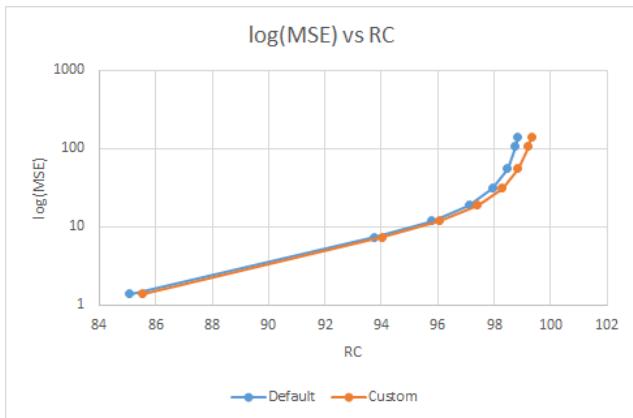


Figura 39: $\log(\text{MSE})$ vs RC - lena

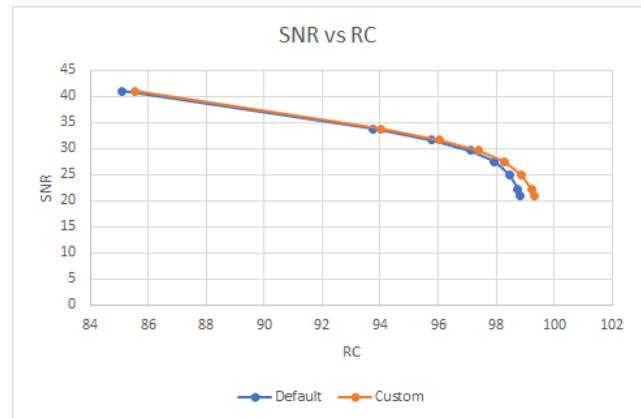


Figura 40: SNR vs RC - lena

DEFAULT						CUSTOM			
MSE	RC	SNR	SSIM	CaliQ	MSE	RC	SNR	SSIM	
1.3829	85.09	41.067	0.98906	5	1.3829	85.517	41.067	0.98906	
7.3942	93.763	33.786	0.95456	25	7.3942	94.042	33.786	0.95456	
12.006	95.784	31.68	0.93624	50	12.006	96.046	31.68	0.93624	
19.072	97.115	29.67	0.91202	100	19.072	97.397	29.67	0.91202	
31.137	97.944	27.542	0.8782	200	31.137	98.286	27.542	0.8782	
55.747	98.459	25.012	0.82541	400	55.747	98.849	25.012	0.82541	
105.91	98.75	22.225	0.75271	750	105.91	99.209	22.225	0.75271	
141.81	98.844	20.957	0.72186	1000	141.81	99.321	20.957	0.72186	

En esta imagen vemos un comportamiento casi idéntico al de la imagen anterior, aunque en ambos compresores se alcanzan valores de ratio de compresión (RC) un poco menores en los valores de CaliQ más pequeños. Esto tiene sentido puesto que ambas imágenes son

similares, mostrando un sujeto con detalles complejos sobre un fondo con pocos elementos. Las dos tienen también valores de MSE y SNR muy parecidos.

El MSE empieza a aumentar exponencialmente a partir del cuarto punto. El SNR empieza a disminuir exponencialmente en el mismo punto.

4.2.4. mandrill

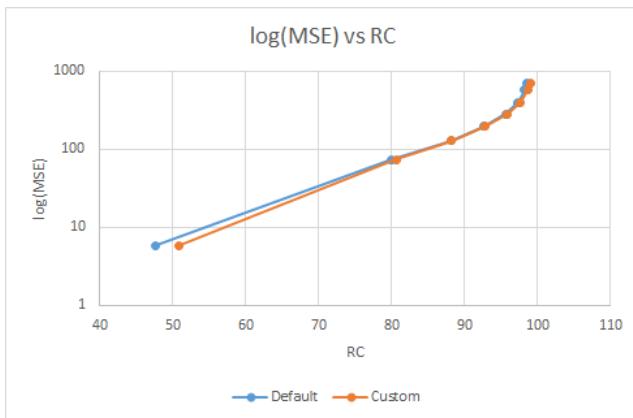


Figura 41: log(MSE) vs RC - mandrill

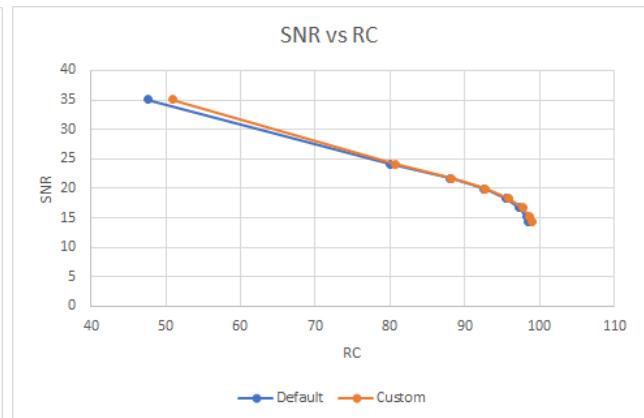


Figura 42: SNR vs RC - mandrill

DEFAULT								CUSTOM			
MSE	RC	SNR	SSIM	CaliQ	MSE	RC	SNR	SSIM	CaliQ	RC	SNR
5.8864	47.632	35.123	0.99594	5	5.8864	50.876	35.123	0.99594	5	50.876	35.123
73.4	80.002	24.164	0.95431	25	73.4	80.657	24.164	0.95431	25	80.657	24.164
129.86	88.1	21.687	0.92457	50	129.86	88.246	21.687	0.92457	50	88.246	21.687
197.33	92.645	19.869	0.89093	100	197.33	92.785	19.869	0.89093	100	92.785	19.869
283.54	95.596	18.295	0.84647	200	283.54	95.872	18.295	0.84647	200	95.872	18.295
404.12	97.324	16.756	0.78649	400	404.12	97.718	16.756	0.78649	400	97.718	16.756
581.31	98.262	15.177	0.70036	750	581.31	98.711	15.177	0.70036	750	98.711	15.177
698.92	98.564	14.377	0.63966	1000	698.92	99.026	14.377	0.63966	1000	99.026	14.377

Tabla 3: Tabla de resultados experimentales

Este ejemplo no ofrece resultados muy buenos. Destacar que se necesita alcanzar un valor de CaliQ=100 para alcanzar un ratio de compresión superior al 90 %, pero entonces el MSE se sitúa alrededor de 200 y el SNR por debajo de 20dB. A partir de aquí apenas se consigue mejorar el RC, a costa de aumentar el ruido y el error de la imagen comprimida.

Ambos compresores actúan prácticamente igual en este caso.

4.2.5. noise

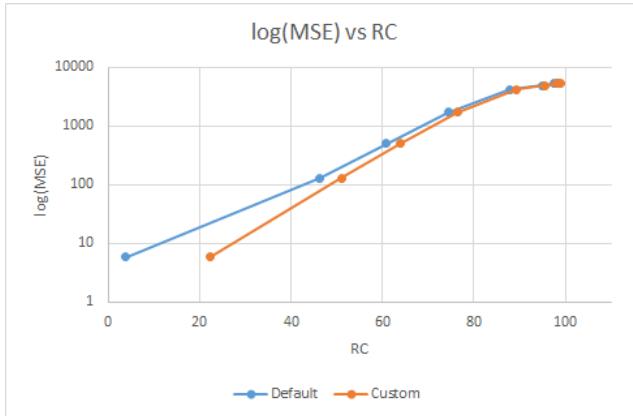
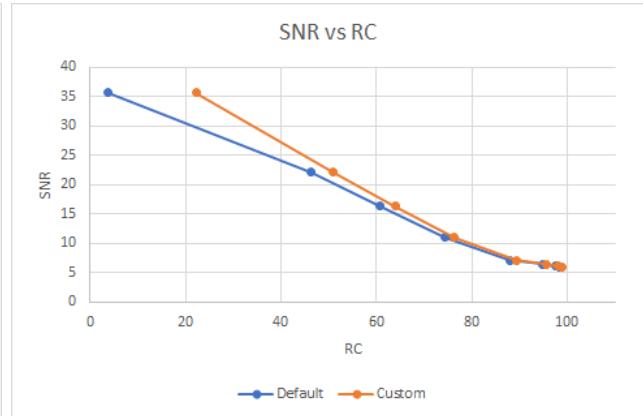
Figura 43: $\log(\text{MSE})$ vs RC - noise

Figura 44: SNR vs RC - noise

DEFAULT					CUSTOM			
MSE	RC	SNR	SSIM	CaliQ	MSE	RC	SNR	SSIM
5.8758	3.8042	35.67	0.9994	5	5.8758	22.159	35.67	0.9994
131.16	46.367	22.182	0.98674	25	131.16	51.023	22.182	0.98674
500.76	60.833	16.364	0.95033	50	500.76	63.897	16.364	0.95033
1730.4	74.386	10.979	0.83479	100	1730.4	76.425	10.979	0.83479
4179.8	87.913	7.1488	0.51253	200	4179.8	89.372	7.1488	0.51253
4884.5	94.898	6.4722	0.29067	400	4884.5	95.575	6.4722	0.29067
5316.6	97.646	6.104	0.16772	750	5316.6	98.203	6.104	0.16772
5432.2	98.427	6.0106	0.096689	1000	5432.2	98.918	6.0106	0.096689

Tabla 4: Tabla de resultados experimentales

En esta imagen destaca cómo el MSE alcanza valores extremadamente altos. Esto es algo esperable debido a que los valores son aleatorios (no hay correlación), de manera que al eliminar componentes de la transformada inevitablemente hay muchos cambios respecto a la original y aumenta el error. Por eso, si queremos mantener un valor relativamente pequeño del MSE (en torno a 100), conseguimos tasas de compresión muy pequeñas (alrededor de 50 %), un resultado muy inferior al resto de ejemplos.

Con el SNR ocurre lo mismo, disminuye extremadamente rápido debido a la aparición de ruido con cualquier mínimo cambio que se produzca.

Una observación interesante es que para los valores de calidad más altos, se alcanzan cotas de RC cercanas al 100 %. Esto se debe a que la imagen no se puede representar con unos pocos componentes de la transformada con magnitud alta, y se desvirtúa hasta que se asemeja cada vez más a una imagen plana (por eso el error y el ruido se disparan).

El compresor custom parece ofrecer mejores resultados con valores de calidad bajos, pero en cuanto aumenta ambos tienen un desempeño prácticamente igual de malo.

4.2.6. pattern

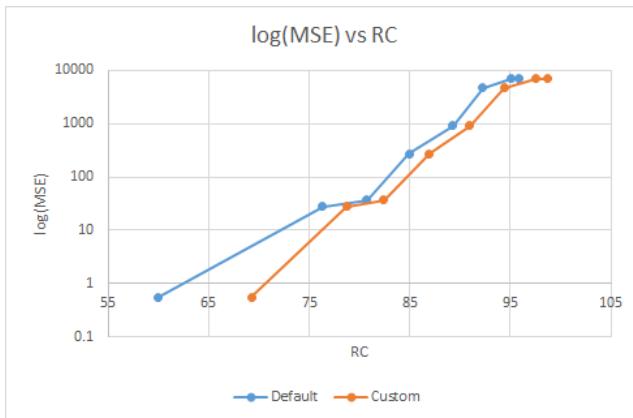


Figura 45: $\log(\text{MSE})$ vs RC - pattern

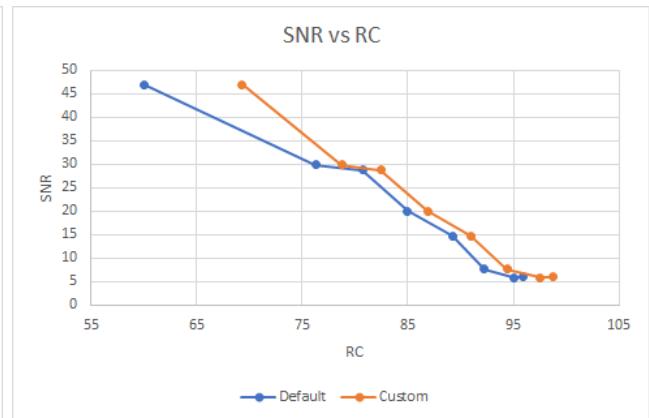


Figura 46: SNR vs RC - pattern

DEFAULT					CUSTOM			
MSE	RC	SNR	SSIM	CaliQ	MSE	RC	SNR	SSIM
0.54688	59.961	47.004	0.99998	5	0.54688	69.28	47.004	0.99998
27.781	76.299	29.945	0.99909	25	27.781	78.72	29.945	0.99909
36.609	80.725	28.747	0.99882	50	36.609	82.425	28.747	0.99882
270.2	84.955	20.066	0.99084	100	270.2	86.916	20.066	0.99084
928.56	89.251	14.705	0.96695	200	928.56	90.955	14.705	0.96695
4767.7	92.245	7.5996	0.81055	400	4767.7	94.406	7.5996	0.81055
7067.2	95.044	5.8902	0.70365	750	7067.2	97.536	5.8902	0.70365
6980	95.89	5.9441	0.71897	1000	6980	98.714	5.9441	0.71897

Tabla 5: Tabla de resultados experimentales

Este caso resulta curioso porque a priori la regularidad del patrón debería ofrecer buenos resultados en la compresión. Sin embargo, vemos que aumenta el MSE rápidamente hasta alcanzar valores elevados (MSE 900 con CaliQ=200). Con el SNR sucede lo mismo, a partir del cuarto punto los valores son muy bajos como para poder considerarlos satisfactorios.

Este es un caso en el que la diferencia visual subjetiva no es muy alta pero el MSE sí.

El compresor custom parece funcionar un poco mejor que el default.

4.2.7. color_bars

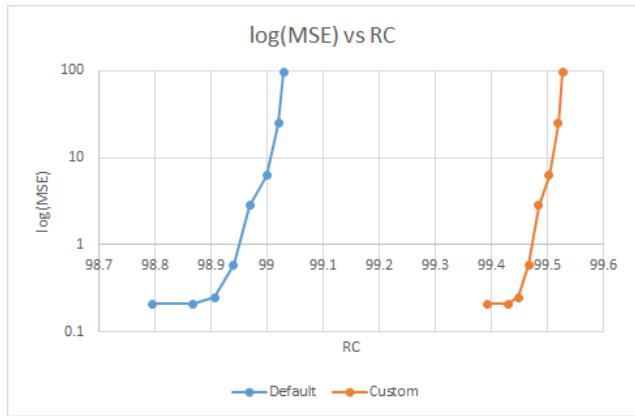


Figura 47: log(MSE) vs RC - color_bars

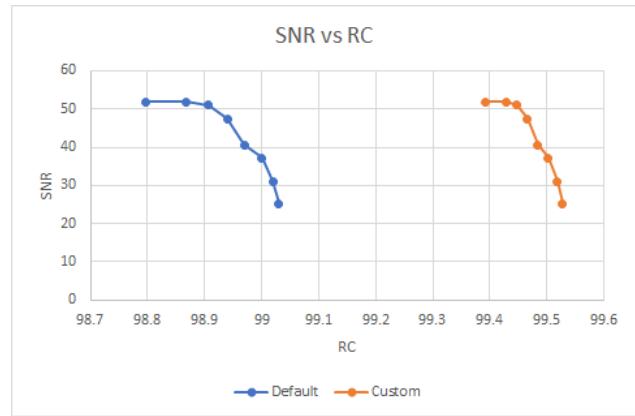


Figura 48: SNR vs RC - color_bars

DEFAULT					CUSTOM			
MSE	RC	SNR	SSIM	CaliQ	MSE	RC	SNR	SSIM
0.20833	98.797	51.899	0.99999	5	0.20833	99.393	51.899	0.99999
0.20833	98.868	51.899	0.99999	25	0.20833	99.429	51.899	0.99999
0.25	98.907	51.107	0.99999	50	0.25	99.448	51.107	0.99999
0.58333	98.94	47.427	0.99992	100	0.58333	99.466	47.427	0.99992
2.8333	98.97	40.563	0.99989	200	2.8333	99.484	40.563	0.99989
6.25	99	37.128	0.99965	400	6.25	99.502	37.128	0.99965
25	99.02	31.107	0.88895	750	25	99.518	31.107	0.88895
97.583	99.03	25.193	0.99374	1000	97.583	99.527	25.193	0.99374

Tabla 6: Tabla de resultados experimentales

Este es el caso más interesante. La imagen se había considerado con la intención de ofrecer un caso mucho más favorable para el compresor custom que para el compresor default debido a que existen muy pocos tipos de píxeles que aparecen muchas veces al ser todo líneas verticales de unos pocos colores (y el compresor custom asigna códigos más bajos en función de la frecuencia de aparición).

Aunque ambos compresores ofrecen unos resultados excelentes (errores muy bajos, SNR alto y ratios de compresión de casi el 100 %), el compresor custom es claramente superior, con resultados de compresión para CaliQ=5 mejores que el compresor default para CaliQ=1000.

4.2.8. explorer

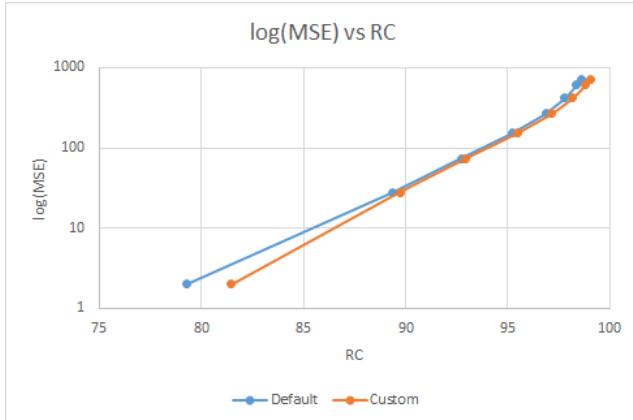


Figura 49: log(MSE) vs RC - explorer

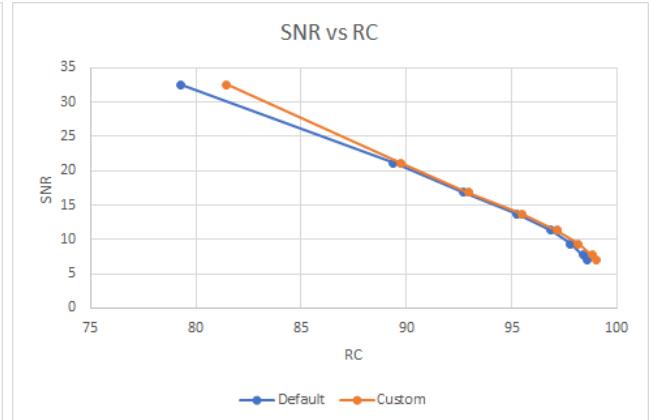


Figura 50: SNR vs RC - explorer

DEFAULT		CUSTOM						
MSE	RC	SNR	SSIM	CaliQ	MSE	RC	SNR	SSIM
2.0174	79.265	32.544	0.99554	5	2.0174	81.468	32.544	0.99554
27.963	89.365	21.126	0.96151	25	27.963	89.732	21.126	0.96151
74.153	92.714	16.891	0.93686	50	74.153	92.953	16.891	0.93686
153.32	95.221	13.736	0.90553	100	153.32	95.461	13.736	0.90553
268.1	96.862	11.309	0.87495	200	268.1	97.144	11.309	0.87495
421.09	97.786	9.3483	0.82631	400	421.09	98.172	9.3483	0.82631
611.57	98.369	7.7276	0.70553	750	611.57	98.806	7.7276	0.70553
721.35	98.575	7.0106	0.72691	1000	721.35	99.031	7.0106	0.72691

Tabla 7: Tabla de resultados experimentales

En esta imagen se obtienen resultados buenos de compresión para factores de calidad bajos (alrededor de 90 % para CaliQ=25), pero luego (a partir de CaliQ=200) empeora rápidamente, aumentando el MSE por encima de 250 y disminuyendo el SNR debajo de 15dB.

El comprresor custom parece que funciona un poco mejor que el default, pero no podemos decir que esta diferencia sea muy significativa cuando crece el factor CaliQ.

4.2.9. gradient

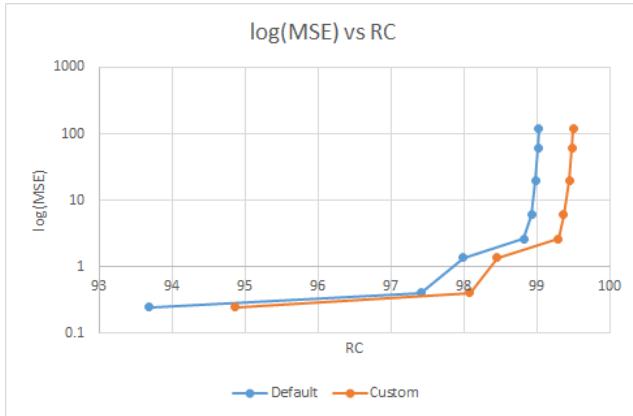


Figura 51: log(MSE) vs RC - gradient

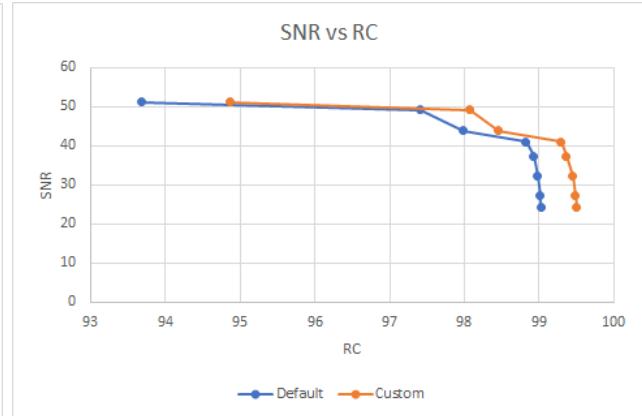


Figura 52: SNR vs RC - gradient

DEFAULT					CUSTOM				
MSE	RC	SNR	SSIM	CaliQ	MSE	RC	SNR	SSIM	
0.24319	93.686	51.303	0.99997	5	0.24319	94.861	51.303	0.99997	
0.39748	97.409	49.17	0.99996	25	0.39748	98.076	49.17	0.99996	
1.3449	97.984	43.876	0.99986	50	1.3449	98.455	43.876	0.99986	
2.6264	98.82	40.969	0.99975	100	2.6264	99.289	40.969	0.99975	
5.9551	98.926	37.414	0.99943	200	5.9551	99.364	37.414	0.99943	
19.772	98.983	32.202	0.99824	400	19.772	99.446	32.202	0.99824	
61.904	99.014	27.246	0.99478	750	61.904	99.484	27.246	0.99478	
121.31	99.022	24.324	0.99014	1000	121.31	99.5	24.324	0.99014	

Tabla 8: Tabla de resultados experimentales

Estos resultados son curiosos debido a que valores pequeños de CaliQ ofrecen ratios de compresión bastante elevados (más del 95 %) con buenos valores de MSE y SNR, pero a partir del 4º punto (CaliQ=100), el MSE se dispara exponencialmente y el SNR disminuye de la misma forma. Esto nos indica que a partir del valor 100, no se gana apenas en RC y los resultados empeoran considerablemente.

Una vez más el compresor custom tiene un desempeño mejor que el compresor default.

4.2.10. color_shapes

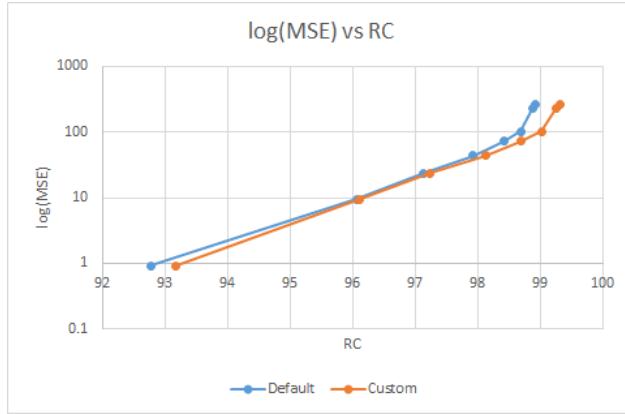
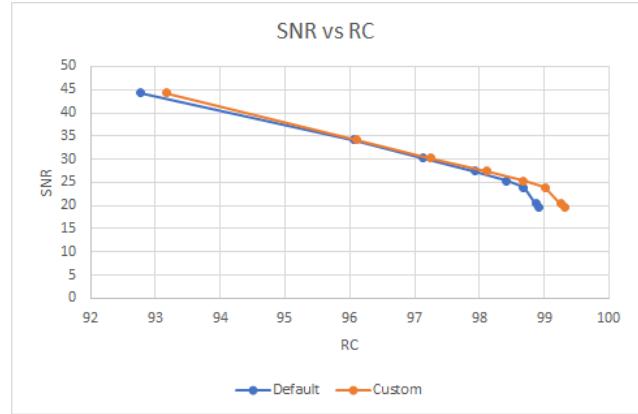
Figura 53: $\log(\text{MSE})$ vs RC - color_shapes

Figura 54: SNR vs RC - color_shapes

DEFAULT					CUSTOM			
MSE	RC	SNR	SSIM	CaliQ	MSE	RC	SNR	SSIM
0.91712	92.775	44.281	0.99995	5	0.91712	93.173	44.281	0.99995
9.4936	96.059	34.131	0.99937	25	9.4936	96.116	34.131	0.99937
23.517	97.135	30.192	0.99844	50	23.517	97.242	30.192	0.99844
43.939	97.928	27.477	0.99716	100	43.939	98.123	27.477	0.99716
72.221	98.417	25.319	0.99509	200	72.221	98.684	25.319	0.99509
101.07	98.685	23.859	0.99315	400	101.07	99.018	23.859	0.99315
225.92	98.868	20.366	0.98443	750	225.92	99.252	20.366	0.98443
270.49	98.916	19.584	0.98396	1000	270.49	99.314	19.584	0.98396

Tabla 9: Tabla de resultados experimentales

En este caso se alcanzan ratios de compresión elevados (más del 95 % a partir de CaliQ=25), lo que tiene sentido teniendo en cuenta que hay grandes áreas del mismo color y no hay detalles excepto en las fronteras entre colores. El MSE aumenta de forma prácticamente lineal en la gráfica hasta CaliQ=200, donde aumenta la pendiente. Con el SNR sucede de manera análoga decreciente. Esto nos indica que el valor 100 puede ser un buen valor para tomar, a partir de ese punto la imagen empeora y no hay tanta ganancia en RC.

4.2.11. x-ray

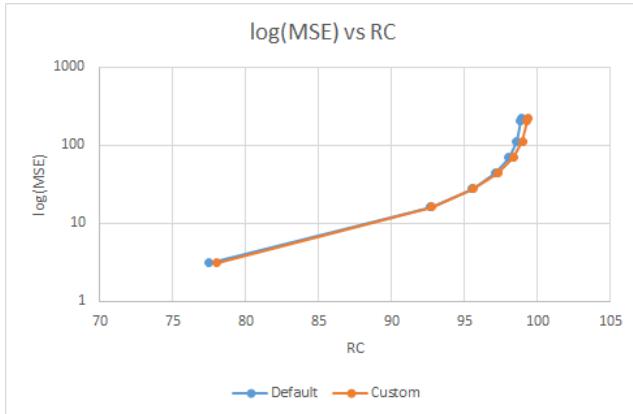


Figura 55: log(MSE) vs RC - x-ray

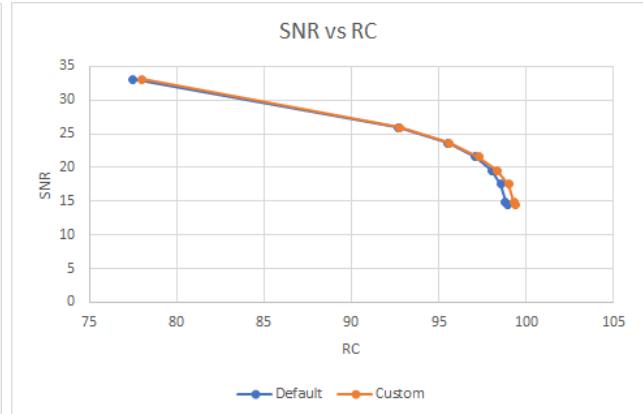


Figura 56: SNR vs RC - x-ray

DEFAULT		CUSTOM						
MSE	RC	SNR	SSIM	CaliQ	MSE	RC	SNR	SSIM
3.1169	77.451	33.107	0.98371	5	3.1169	78.006	33.107	0.98371
16.313	92.677	25.919	0.96718	25	16.313	92.759	25.919	0.96718
27.827	95.521	23.599	0.95885	50	27.827	95.599	23.599	0.95885
44.107	97.089	21.599	0.94761	100	44.107	97.289	21.599	0.94761
70.002	98.03	19.593	0.91369	200	70.002	98.346	19.593	0.91369
113.92	98.574	17.478	0.87198	400	113.92	98.996	17.478	0.87198
207.8	98.83	14.868	0.79916	750	207.8	99.303	14.868	0.79916
223.69	98.914	14.548	0.8113	1000	223.69	99.401	14.548	0.8113

Tabla 10: Tabla de resultados experimentales

En este ejemplo se consiguen ($\text{CaliQ}=100$) tasas de compresión de alrededor del 98% con un MSE bastante bajo de alrededor de 50 y un SNR superior a 20dB (valores mejores que en muchos de los ejemplos anteriores). A partir de este punto la calidad de la imagen resultante empeora sin mejorar la RC significativamente.

El comprresor custom es ligeramente mejor al default en este ejemplo.

4.2.12. Conclusiones del estudio cuantitativo

En los ejemplos anteriores se ha analizado la evolución de las diferentes métricas (RC,MSE,SNR) en función del valor del factor de calidad CaliQ escogido. Se ha visto en qué ejemplos se conseguían mayores ratios de compresión con un MSE menor y un SNR mayor, en cuáles se disparaba el error y en cuáles se mantenía en niveles más bajos (respectivamente con el SNR).

De aquí podemos extraer varias conclusiones:

Para empezar, los datos obtenidos señalan un mejor desempeño del compresor custom en más de la mitad de los ejemplos probados (concretamente en 6 de los 11 ejemplos). En aquellos ejemplos en los que este compresor no era claramente superior, ambos obtenían resultados prácticamente idénticos. Esto parece indicar que el compresor custom obtiene mejores resultados en general (aunque habría que realizar una prueba más exhaustiva con una batería de imágenes mucho mayor para poder obtener conclusiones más firmes).

También se ha determinado que en 9 de las 11 imágenes analizadas se alcanza claramente un punto a partir del cual el MSE se dispara y el SNR disminuye rápidamente. Esto indicaría que a partir de ese punto la pérdida de información en la imagen comprimida aumenta exponencialmente, sin proporcionar aumentos notables en el ratio de compresión (RC). Para intentar establecer un criterio objetivo basado únicamente en estos datos, podríamos considerar que el valor adecuado del factor de compresión CaliQ para emplear con los compresores es precisamente el anterior a que se produzca este crecimiento. De esta manera se intenta maximizar la compresión sin que la imagen pierda excesiva calidad.

En 5 de las 9 ocasiones en las que el este fenómeno de crecimiento se ha percibido se ha determinado que el punto óptimo es el cuarto, correspondiente al factor de calidad CaliQ=100. Los otros dos valores que más se repiten son CaliQ=50 (2/9) y CaliQ=200 (2/9). Este resultado coincide con el que se obtuvo en el análisis cualitativo, lo que refuerza la fiabilidad de estas conclusiones.

De esta manera, por los datos obtenidos podríamos concluir que la configuración óptima del compresor para un caso general sería utilizar el **compresor Custom con factor de calidad CaliQ=100**.

Como observaciones finales, los ejemplos en los que los dos compresores obtienen resultados prácticamente idénticos son precisamente las imágenes en las que más detalles diferentes había (candados, mandrill, noise) y por tanto menos correlación de la que poder sacar partido, sobre todo el compresor custom. Son también las imágenes en las que peores resultados se obtienen. Es decir, el compresor custom funciona mejor en los casos promedio y en los casos extremos condicionados negativamente los dos funcionan igual.

4.2.13. Extra: Análisis de SSIM (Structural Similarity Index)

Al realizar el análisis cuantitativo solo se está teniendo en cuenta la diferencia “numérica” entre la imagen comprimida y la original, mediante un criterio únicamente matemático bajo la premisa de que las diferencias “reales” que existen entre las imágenes se traducen en diferencias percibidas por el observador. Esto es así en la mayoría de los casos, pero en otros (por ejemplo, en la imagen *pattern* analizada anteriormente), no tiene por qué ser así.

Si investigamos sobre el tema, encontramos una métrica utilizada con frecuencia en el estudio de algoritmos de compresión que permite aunar hasta cierto punto el análisis cuantitativo con la percepción visual de la imagen. Se trata del SSIM (Structural Similarity Index) [1].

A diferencia del MSE y la SNR (las métricas que se han estudiado en la asignatura), el SSIM tiene en cuenta la estructura y la similitud perceptual de las imágenes. Evalúa la similitud de luminancia, el contraste y la estructura y proporciona una evaluación más centrada en cómo los humanos perciben y procesan visualmente las imágenes ya que considera la preservación de detalles, bordes y texturas en lugar de medir solo las diferencias de píxeles (MSE) o la relación señal a ruido (SNR). No deja de ser un análisis numérico pero los datos evaluados parecen estar más relacionados con el resultado final de percepción de un observador humano.

Por esta razón, decidimos ampliar el código de la práctica para calcular la métrica SSIM, para ver si nos aporta alguna información extra que sea relevante.

Para ello se ha utilizado la función de la librería de *Matlab* [3], puesto que la implementación manual sería compleja y escaparía del ámbito de este trabajo. Esta función calcula el valor de la métrica SSIM entre una imagen (en nuestro caso la imagen comprimida) y la imagen de referencia (en nuestro caso la original).

Calculamos este valor para cada una de las combinaciones de compresor-factor de calidad y para todas las imágenes y obtenemos resultados que se han mostrado en las tablas de la sección anterior (columna “SSIM”). El índice SSIM toma valores entre 0 y 1, siendo 1 una reconstrucción perfecta de la imagen original (a partir de 0.97 se considera ya una reconstrucción muy buena). Cabe esperar que el SSIM disminuya conforme aumenta la compresión y disminuya la calidad de las imágenes.

Analizamos los valores obtenidos, fijándonos sobre todo en CaliQ=100, que hemos determinado que era el valor más adecuado en general:

En algunos ejemplos como *candados*, *lena*, *mandrill*, *explorer*, *x-ray*, tenemos valores entre 0.90 y 0.95 para el valor de CaliQ que habíamos seleccionado como adecuado en el análisis anterior. Esto indica un resultado visual que no podemos considerar muy bueno pero sí aceptable, lo que concuerda con la percepción visual que habíamos obtenido en el análisis.

En el caso de *lennon*, habíamos considerado que el factor 200 no era aceptable, pero el índice SSIM da un valor de 0.966 0.97 en ese punto, de forma que se consideraría un resultado aceptable según esta métrica. Aquí vemos que puede haber un poco de disparidad entre el parecido percibido y el calculado, algo esperable, pero no es demasiado discordante.

Quizás el caso más interesante sea el de *gradient* y *color_shapes*, en los que el índice toma valores muy buenos (>0.98) para todos los valores de CaliQ (incluído el valor 1000). En este caso dista mucho de la percepción visual que se puede observar, y es quizás el ejemplo donde más diferencia hay.

Como conclusión, se trata de una métrica bastante acertada (quizás incluso más que el MSE y el SNR) para medir la percepción visual de las imágenes comprimidas, aunque no es perfecta. Sin embargo, considero que puede ser de gran utilidad en combinación con las otras métricas para poder extraer conclusiones cuando no es posible contar con una verificación humana para confirmar los resultados.

5. Conclusión

En este proyecto hemos desarrollado dos compresores de imágenes basados en las técnicas de DCT y codificación Huffman, con un enfoque similar pero con algunas diferencias, como aproximación al método de compresión JPEG.

Hemos llevado a cabo una batería de experimentaciones con los compresores sobre imágenes cuidadosamente seleccionadas, comprimiendo y descomprimiendo para obtener datos sobre cada uno de ellos.

El análisis cualitativo nos ha permitido determinar que en la mayoría de los casos los compresores producían mejores resultados cuando se tomaba el factor de calidad CaliQ=100. También hemos podido observar algunos patrones comunes que se producían en las imágenes conforme aumentaba la compresión (difuminación de texturas, distorsión de colores, etc.) y hemos intentado deducir su causa.

Posteriormente se ha realizado el análisis cuantitativo, para comprobar cómo cambiaban “numéricamente” las imágenes con los distintos parámetros de compresión y para cuantificar el error producido. De manera extra, se ha realizado el cálculo de la métrica SNR para ampliar este análisis. Hemos concluido que el valor más adecuado de CaliQ que minimizaba el error cometido y el ruido introducido, maximizando al mismo tiempo el ratio de compresión es precisamente CaliQ=100, reforzando las conclusiones extraídas con el análisis cualitativo.

Al final del apartado se ha incluido también una breve discusión extra sobre la métrica SSIM y sus virtudes frente a las otras que se han estudiado. Se ha calculado para cada ejemplo y se han analizado brevemente los resultados obtenidos.

En conclusión, hemos realizado un estudio exhaustivo del desempeño de los dos compresores que se han implementado. Esto nos ha permitido obtener información muy relevante acerca de los dos compresores implementados y determinar su configuración de uso óptima para un caso general. El desarrollo del proyecto también nos ha permitido poner en práctica conceptos y técnicas aprendidas en la asignatura, mediante un enfoque más práctico que ha ayudado a su comprensión.

Referencias

- [1] Imatest. *Introduction — The Structural Similarity Index (SSIM)*. Online. Accessed: 2 de Junio 2023. URL: <https://www.imatest.com/support/docs/23-1/ssim/#:~:text=Introduction%5C%20%5C%E2%5C%80%5C%94%5C%20The%5C%20Structural%5C%20Similarity%5C%20Index,by%5C%20losses%5C%20in%5C%20data%5C%20transmission..>
- [2] International Telecommunication Union (ITU). *ITU T.81: Digital compression and coding of continuous-tone still images - Requirements and guidelines*. ITU-T Recommendation. Annex K, Annex C, and Annex F. 1992.
- [3] MathWorks. *ssim*. MathWorks Documentation. Accessed: 3 de Junio 2023. URL: <https://es.mathworks.com/help/images/ref/ssim.html>.
- [4] Universidad de Murcia Profesores de la asignatura Compresión Multimedia. *PROYECTO FINAL “COMPRESIÓN DE IMÁGENES BASADA EN DCT: ESTUDIO EXPERIMENTAL”*. Material proporcionado por los profesores de Compresión Multimedia. Accedido a través del aula virtual el 16 de Abril de 2023.
- [5] USC Signal and Image Processing Institute. *USC-SIPI Image Database*. Online. Accedido: 20 de Mayo 2023. URL: <https://sipi.usc.edu/database/database.php?volume=misc>.
- [6] Wikipedia contributors. *Discrete cosine transform*. Wikipedia. Accessed: 12 de Mayo 2023. URL: https://en.wikipedia.org/wiki/Discrete_cosine_transform.
- [7] Yao Yao. *Sample Data*. NYU Electrical Engineering Department. Accessed: 15 de Mayo 2023. URL: <https://eeweb.engineering.nyu.edu/~yao/EL5123/SampleData.html>.