



FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE DE
COIMBRA



Multimédia

2024/2025 - 2º Semestre

**Trabalho Prático N°1:
Compressão de Imagem**

Carlos Pereira (2022232042)

Diogo Costa (2022213506)

Fernando Santos (2019224933)

Rafael Marques (2022208522)

1. Introdução

A compressão de imagens é algo que acompanha o nosso dia a dia, permitindo reduzir o tamanho de uma imagem sem que haja uma perda significativa de qualidade. Neste relatório apresentamos uma análise detalhada do processo de compressão de imagem JPEG, implementado como parte do Trabalho Prático 1.

Este trabalho prático tem como objetivo a implementação do codec JPEG, dando-nos competências, conhecimento e sensibilidade para as questões fundamentais de Compressão de Imagem.

2. Metodologia

O algoritmo foi implementado em Python, utilizando bibliotecas como NumPy, OpenCV e Matplotlib. As principais etapas do processo incluem:

1. Compressão de imagens BMP
2. Leitura e visualização da imagem
3. Padding
4. Conversão de RGB para YCbCr
5. Downsampling
6. Transformada Discreta do Cosseno (DCT)
7. Quantização
8. Codificação DPCM dos coeficientes DC

3. Análise das Etapas

3.1 Compressão de imagens BMP

A compressão de imagens BMP para o formato JPEG utilizando um editor de imagem como o GIMP permite avaliar o impacto de diferentes níveis de qualidade na compressão.

Ao comprimir as imagens com qualidade alta (Q=75), média (Q=50) e baixa (Q=25), é possível observar uma progressiva redução no tamanho do arquivo (de 2374KB originalmente para 155KB, 103KB e 67KB respetivamente), acompanhada por uma degradação gradual da qualidade visual.

Imagem	Original	75%	50%	25%
airport.bmp	<i>Muito Alta</i>	<i>Muito Alta</i>	<i>Alta</i>	<i>Média</i>
geometric.bmp	<i>Alta</i>	<i>Alta</i>	<i>Média</i>	<i>Média</i>
nature.bmp	<i>Muito Alta</i>	<i>Muito Alta</i>	<i>Alta</i>	<i>Alta</i>

Tabela 1 – Análise qualitativa das imagens após compressão

Após uma análise qualitativa dos resultados da compressão acima descritos foi-nos possível tirar algumas conclusões. Apesar de ter sido efetuada uma compressão destrutiva, a maioria

das imagens parecem perfeitamente bem à primeira vista, nomeadamente a nature.bmp. O maior problema aparece quando focamos ou damos zoom, apercebendo-nos de uma maior inconsistência e falta de qualidade na imagem, onde vemos “blocos” de píxeis. Na imagem geometric.bmp na zona de contacto entre duas cores diferentes elas misturam-se mostrando uma maior perda de qualidade.

3.2 Leitura e Visualização da Imagem

A imagem original é lida e visualizada em seus componentes RGB. Cada canal é exibido separadamente usando mapas de cores específicos (vermelho, verde e azul).



Figura 1 - Visualização dos canais RGB usando colormaps

3.3 Padding

O pré-processamento da imagem através do padding é uma etapa crucial no processo de compressão JPEG para garantir que as dimensões da imagem sejam múltiplas de 32x32, facilitando o processamento em blocos nas etapas subsequentes. É importante que as dimensões da imagem sejam múltiplas de 32 porque, em primeiro lugar, fazemos downsampling que, na hipótese mais destrutiva (4:1:0), transforma blocos 4x4 em apenas um pixel. Deste modo, blocos 32x32 serão transformados em blocos 8x8 que é o tamanho geralmente utilizado para os cálculos da DCT.

O padding é alcançado replicando a última linha e coluna conforme necessário, utilizando o modo 'edge' do NumPy para uma transição suave. No decoder, a função inversa remove o padding adicionado, recuperando os canais RGB com as dimensões originais da imagem.

Esta abordagem assegura que a integridade da imagem é mantida durante todo o processo de compressão e descompressão, permitindo uma comparação precisa entre a imagem original e a reconstruída.

3.4 Conversão de RGB para YCbCr

A conversão do espaço de cor RGB para YCbCr é uma etapa fundamental para a compressão JPEG, pois explora as características da percepção visual humana.

A implementação da função de conversão garante que os valores de luminância (Y) e crominância (Cb e Cr) sejam calculados corretamente, permitindo a separação da informação de brilho da informação de cor.

A função inversa, que converte YCbCr de volta para RGB, assegura que os valores resultantes estejam dentro do intervalo (0-255), garantindo a integridade da imagem. No modelo de cor YCbCr o canal Y contém a maior parte da informação visual, enquanto os canais Cb e Cr

representam as diferenças de cor em relação ao cinza. Ao comparar o canal Y com os canais R, G e B, é evidente que Y é o que se assemelha mais à média ponderada dos canais RGB, refletindo a importância da luminância na percepção humana. Os canais Cb e Cr, por outro lado, contêm informações sobre a intensidade do azul e vermelho, respectivamente. Esta etapa prepara o caminho para o downsampling dos canais Cb e Cr, permitindo uma maior compressão sem uma perda significativa na qualidade visual percebida.

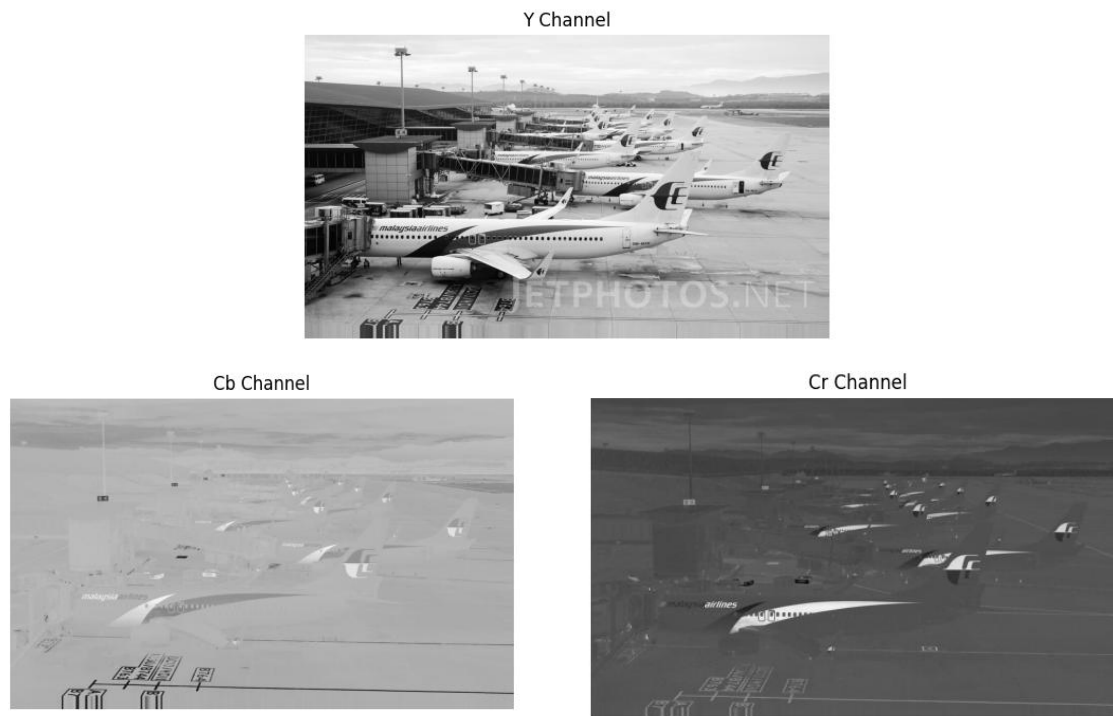


Figura 2 - Visualização dos canais YCbCr

3.5 Downsampling e Upsampling

O downsampling ou subamostragem é um método usado para reduzir a resolução dos canais da crominância (Cb e Cr), pois o olho humano apresenta menor sensibilidade a detalhes nessas informações de cor do que no canal da luminância (Y). Assim, o objetivo é diminuir o tamanho dos dados sem perda significativa de qualidade visual.

O upsampling consiste no processo inverso, para a reestruturação da imagem onde, idealmente, não se nota essa perda de qualidade.

Para a visualização do seu efeito, neste trabalho, vamos analisar a compressão para os seguintes níveis de subamostragem:

Níveis	Y	Cb	Cr
4:2:2	Original	Metade da largura	Metade da largura
4:2:0	Original	Metade da largura e altura	Metade da largura e altura

Tabela 2 – Tamanhos das imagens de subamostragem

Resultados Downsampling 4:2:2

Nesta secção discutimos o resultado do downsampling com fator 4:2:2.

Em primeiro lugar, constatamos que os tamanhos dos canais Y, Cb e Cr, obtidos comprovam a expectativa teórica. Os canais Cb e Cr são reduzidos em 50% apenas na horizontal.

- Dimensões após downsampling 4:2:2 - Y: (704, 1216), Cb: (704, 608), Cr: (704, 608)

- **Canal Y:**



Figura 3 - Visualização do canal Y após Downsampling 4:2:2 com interpolação linear

- **Canal Cb:**



Figura 4 - Visualização do canal Cb após Downsampling 4:2:2 e interpolação linear

- **Canal Cr:**



Figura 5 - Visualização do canal Cr após Downsampling 4:2:2 e interpolação linear

Nestas figuras conseguimos comprovar que os canais Cb e Cr sofreram uma compressão na horizontal, ao ter apenas metade da dimensão da imagem original, existindo perda de informação, já que é feita uma média. Apresentam também menos detalhes do que o canal Y, que contém quase toda a informação da imagem.

Resultados Downsampling 4:2:0

Dimensões após downsampling 4:2:0 - Y: (704, 1216), Cb: (352, 608), Cr: (352, 608)

- **Canal Y:**



Figura 6 - Visualização do canal Y após Downsampling 4:2:0 com interpolação linear

- **Canal Cb:**

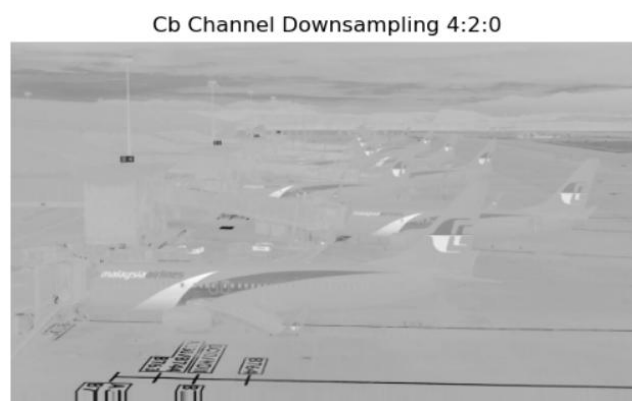


Figura 7 - Visualização do canal Cb após DownSampling 4:2:0 com interpolação linear

- **Canal Cr:**



Figura 8 - Visualização do canal Cr após Downsampling 4:2:0 com interpolação linear

No downsampling de fator 4:2:0 conseguimos visualizar uma perda de informação e de detalhes horizontal e verticalmente.

Análise 4:2:2 vs 4:2:0

Cálculo da taxa de compressão

A taxa de compressão pode é a razão entre o tamanho original e o tamanho comprimido dos dados, ou seja, após o downsampling.

$$\text{Taxa de Compressão} = \text{Tamanho Original} \div \text{Tamanho após Downsampling}$$

- **Para 4:2:2**

1. O canal Y mantém o tamanho original.
2. Cb e Cr são reduzidos para metade
3. Tamanho total: $Y + (Cb/2) + (Cr/2)$
4. Taxa de compressão: $\frac{Y+Cb+Cr}{Y+(\frac{Cb}{2})+(\frac{Cr}{2})} = \frac{3}{2} = 1.5$

- **Para 4:2:0**

1. O canal Y mantém o tamanho original.
2. Cb e Cr são reduzidos para $\frac{1}{4}$ do tamanho original (redução horizontal e vertical)
3. Tamanho total = $Y + (Cb/4) + (Cr/4)$
4. Taxa de compressão: $\frac{Y+Cb+Cr}{Y+(\frac{Cb}{4})+(\frac{Cr}{4})} = \frac{3}{1.5} = 2$

Isto deve-se ao facto de os canais Cb e Cr serem reduzidos duas vezes para metade (horizontalmente e verticalmente).

Método	Tamanho Relativo	Taxa de Compressão
4:4:4	1x	1.0x
4:2:2	2/3	1.5x
4:2:0	1/2	2.0x

Tabela 3 – Tamanho relativo e taxa de compressão dos níveis de subamostragem considerados

Efeitos visuais e Destrutividade

A destrutividade inerente à compressão JPEG, que se manifesta na perda de informação e qualidade visual, está intrinsecamente ligada à taxa de compressão. O nível de subamostragem 4:2:0, ao reduzir mais drasticamente a informação de crominância, resulta numa taxa de compressão superior, mas também numa maior destrutividade em comparação com o nível 4:2:2. Neste último, a imagem preserva mais detalhes, especialmente na vertical.

Comparação de resultados das interpolações

A mudança de interpolação afeta diretamente a qualidade da imagem reconstruída após o upsampling. Através da função `cv2.resize()` podemos implementar as diferentes interpolações.

Para efeitos de simplificação, fizemos comparações apenas no nível 4:2:2 de downsampling. Dos quatro métodos de interpolação que utilizámos, o que nos proporcionou melhores resultados de reconstrução foi o de interpolação cúbica, com um MSE de aproximadamente 86.59. Já a interpolação linear 87.29, nearest 99.71 e area 92.84 todas apresentaram maiores distorções na imagem reconstruída.

- **Interpolação Linear**

Faz a média dos píxeis vizinhos mais próximos (consecutivos).

Nota: Resultados demonstrados acima.

- **Interpolação Nearest**

Cb Channel Downsampling 4:2:2 (Nearest) Cr Channel Downsampling 4:2:2 (Nearest)

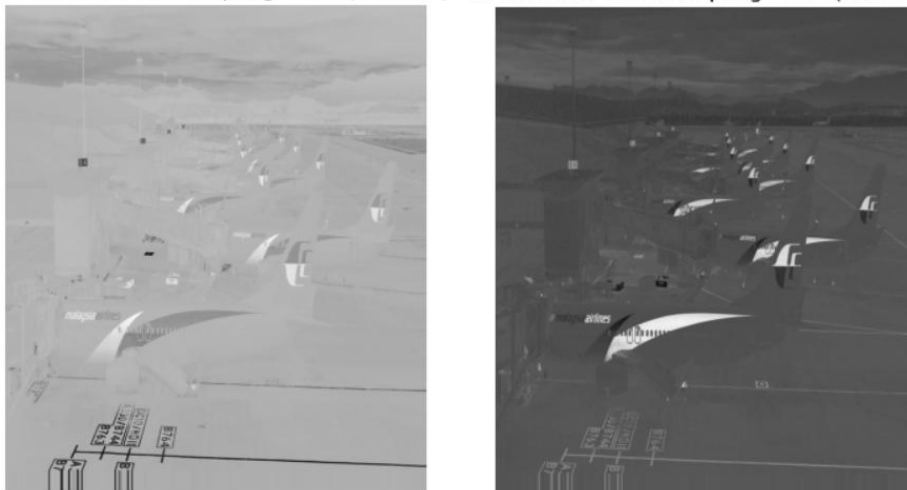


Figura 9 - Visualizações dos canais Cb e Cr após Downsampling 4:2:2 com interpolação Nearest

Seleciona o valor do pixel mais próximo na imagem original para preencher o pixel correspondente na imagem redimensionada, podendo criar bordas visíveis e efeito de “blocos”.

- **Interpolação Cúbica**

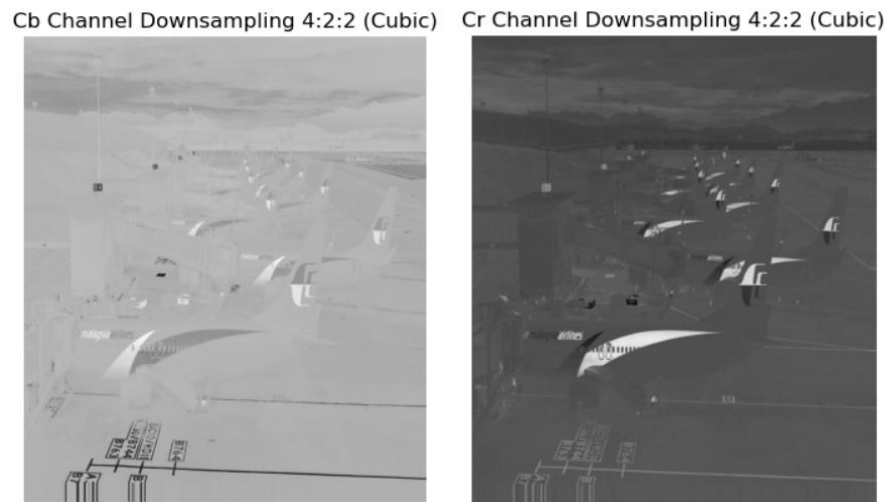


Figura 10 - Visualizações dos canais Cb e Cr após Downsampling 4:2:2 e interpolação cúbica

Considera blocos de 16 píxeis vizinhos para suavizar a imagem. Apresenta resultados mais suaves e com qualidade superior à interpolação linear.

- **Interpolação Área**

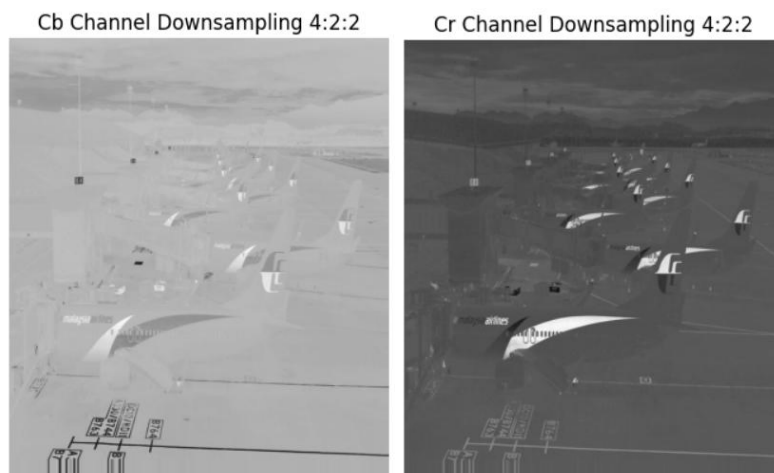


Figura 11 - Visualizações dos canais Cb e Cr após Downsampling 4:2:2 e interpolação Area

Calcula a média dos pixels numa área para criar um pixel menor.

3.6 Transformada de Cosseno Discreta (DCT)

A implementação da Transformada de Cosseno Discreta (DCT) no algoritmo de compressão JPEG é realizada em três configurações diferentes:

1. Blocos 1x1 (imagem inteira)
2. Blocos 8x8
3. Blocos 64x64

Para cada configuração, são criadas funções para calcular a DCT e sua inversa, utilizando as funções `scipy.fftpack.dct` e `scipy.fftpack.idct`.

A visualização dos resultados é feita através de uma transformação logarítmica, que permite uma melhor representação da ampla gama de valores resultantes da DCT. Esta abordagem permite uma análise detalhada do comportamento da DCT em diferentes escalas e sua eficácia na concentração de energia nos coeficientes de baixa frequência, especialmente no canto superior esquerdo dos blocos.

A comparação entre as três configurações revela diferenças significativas em termos de potencial de compressão. A DCT aplicada aos canais completos oferece uma visão global da distribuição de frequências na imagem.

A DCT em blocos 8x8, que é o padrão no JPEG, proporciona um bom equilíbrio entre compressão eficiente e preservação de detalhes locais. Já a DCT em blocos 64x64 oferece maior compressão, mas pode resultar em maiores e mais visíveis distorções, especialmente em áreas com muitos detalhes (transições mais bruscas). Esta análise comparativa é crucial para entender como o tamanho do bloco afeta o compromisso entre compressão e qualidade da imagem no processo de codificação JPEG.

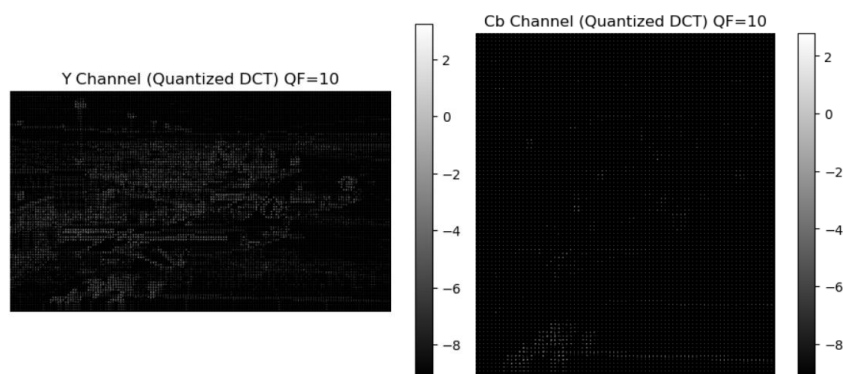
3.7 Quantização

A quantização é o processo de reduzir a precisão dos coeficientes da DCT. Elimina os detalhes menos perceptíveis para o olho humano (altas frequências) e reduz o tamanho do arquivo.

No contexto do JPEG, a quantização ocorre através da divisão dos coeficientes DCT por uma matriz de quantização específica, controlada por um fator de qualidade (QF). Quanto menor este fator, maior a compressão e maior a perda de detalhes, quanto maior, menor a compressão e melhor a qualidade.

Comparação de resultados com diferentes QFs

- **QF = 10**



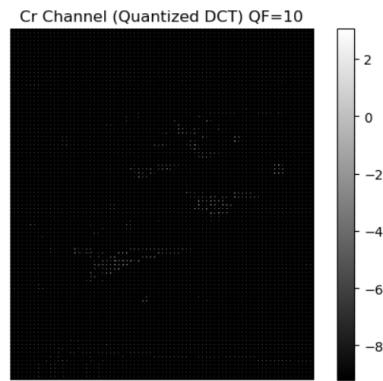


Figura 12 - Visualização dos canais Y, Cb e Cr quantizados com QF = 10

Com um fator de qualidade tão reduzido, podemos verificar os canais com muita compressão e perda de detalhes surgindo distorções perceptíveis.

- **QF = 25**

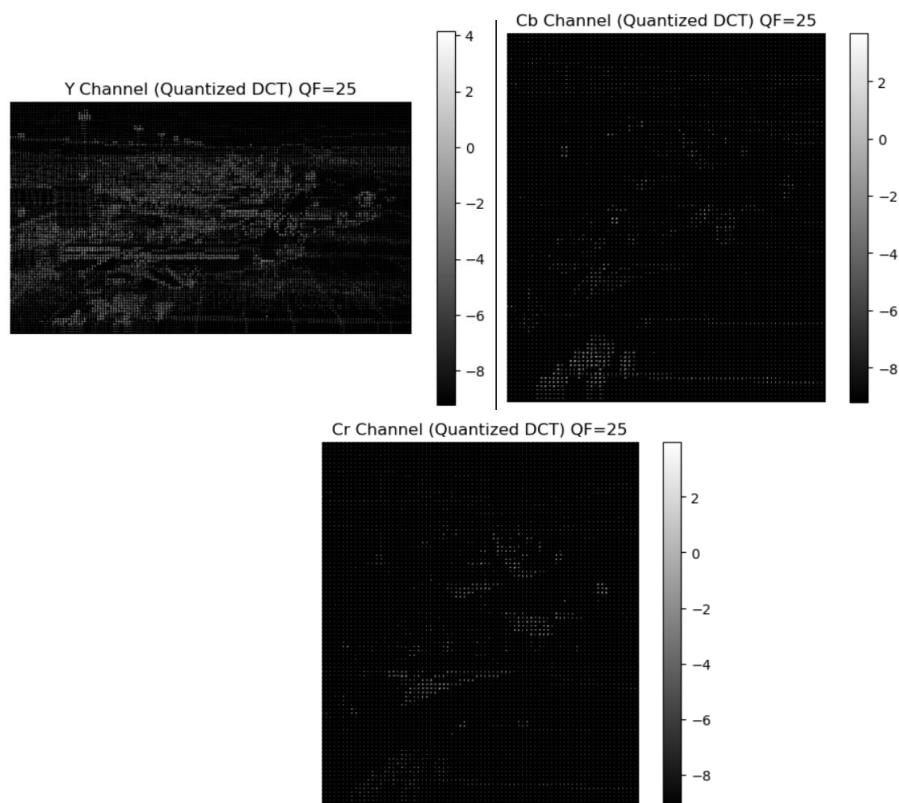


Figura 13 - Visualizações dos canais Y, Cb e Cr quantizados com QF = 25

Com este fator de qualidade conseguimos visualizar um pouco mais de detalhes, mas ainda com muita perda de texturas e detalhes.

- **QF = 50**

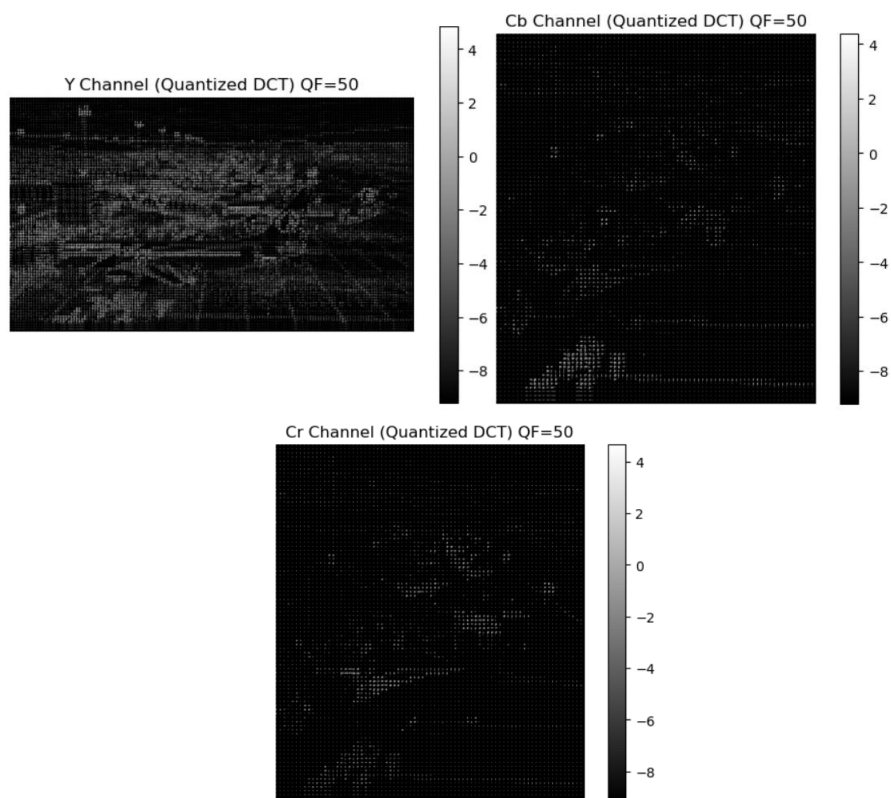


Figura 14- Visualização dos canais Y, Cb e Cr quantizados, com QF = 50

Com um fator de qualidade médio já conseguimos visualizar mais detalhe, com menos “borrões”, especialmente no canal Y da luminância.

- **QF = 75**

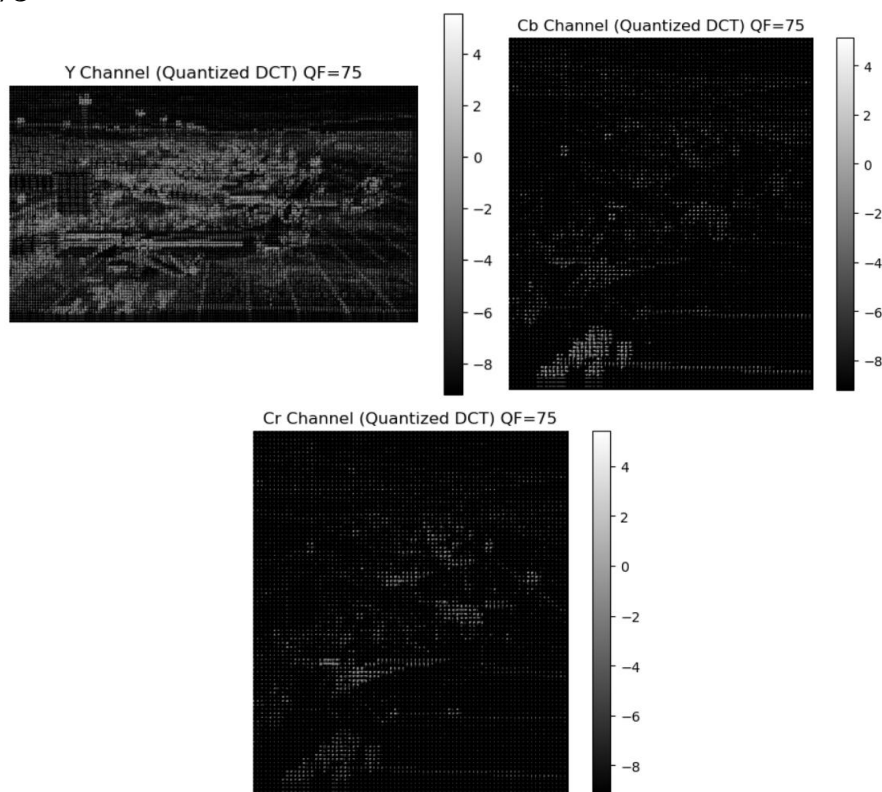


Figura 15 - Visualização dos canais Y, Cb e Cr quantizados com QF = 75

Preservam-se mais detalhes, sendo possível já visualizar a silhueta dos aviões no canal Y, pelo que a perda de detalhes é menor. Este QF corresponde ao padrão de utilização do JPEG ou de algumas fotos online.

- **QF = 100**

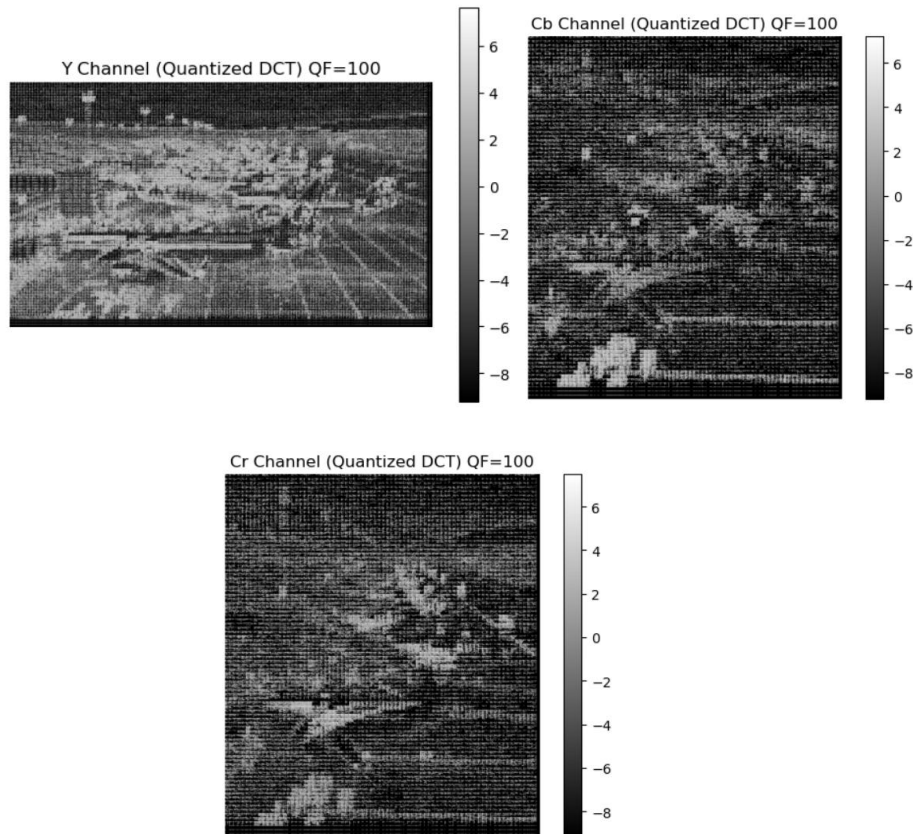


Figura 16 - Visualização dos canais Y, Cb e Cr quantizados com QF = 100

Com a qualidade máxima, mesmo nos canais da cromaticidade já é possível visualizar muitos mais detalhes, em detrimento da compressão de um maior número de dados.

QF	Compressão	Qualidade visual	Preservação de detalhes
10	Forte compressão	Forte suavização, pouca nitidez	Perda acentuada de detalhes
25	Alta compressão	Perda de visibilidade	Poucas texturas
50	Boa compressão	Algumas áreas suavizadas	Pequena perda de texturas
75	Pouca compressão	Leve perda	Mantém uma boa parte de detalhes
100	Muito pouca compressão	Não perde muita visibilidade	Maior parte dos detalhes preservados

Tabela 4 – Impacto na qualidade visual e taxa de compressão dos fatores de qualidade utilizados

Comparação dos resultados da Quantização vs DCT

Para efeitos de simplificação, decidimos comparar com o padrão de JPEG (QF = 75) e DCT aplicada a blocos 8x8.

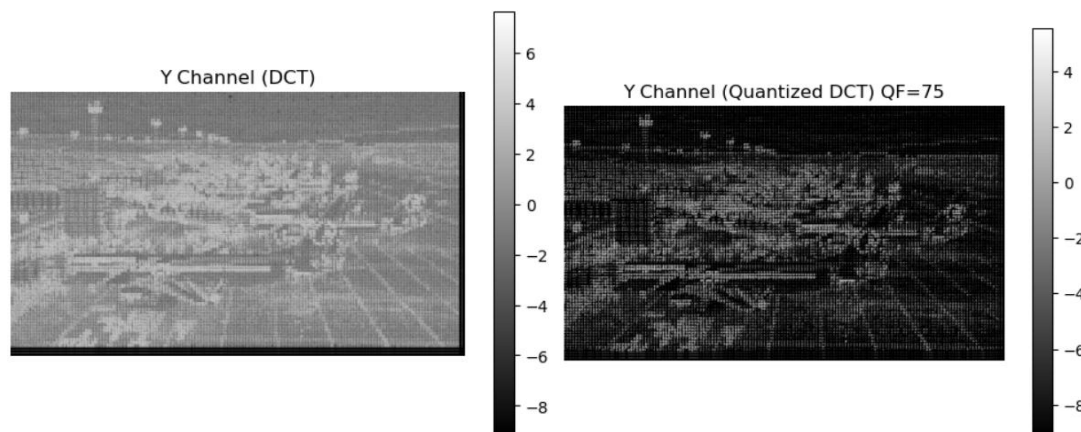


Figura 17 - Canal Y após DCT vs Canal Y quantizado com QF = 75

Aplicar a DCT não comprime a imagem (transforma para o domínio de frequência), é necessária uma posterior quantização, pois este processo reduz a precisão dos coeficientes e elimina os menos relevantes. A DCT transforma os dados, a quantização é responsável por os comprimir.

Esta situação é evidente na Figura 14, na DCT a imagem mantém os detalhes uniformemente e sem suavização visível, pois ainda não houve uma compressão real. Na imagem do canal Y quantizado nota-se perda de detalhes e uma suavização, embora sem blocos visíveis com este fator de qualidade.

3.8 Codificação DPCM dos Coeficientes DC

A etapa de codificação DPCM (Differential Pulse Code Modulation) é implementada para explorar a redundância entre os coeficientes DC dos blocos adjacentes. Ao substituir o valor DC de cada bloco pela diferença entre o seu valor e o do bloco anterior, a amplitude dos coeficientes é geralmente reduzida, tornando-os mais adequados para a codificação entrópica subsequente. As funções de codificação e decodificação DPCM garantem que os valores originais dos coeficientes DC quantizados (Y_q , Cb_q e Cr_q) possam ser recuperados com precisão. A análise dos resultados revela que a DPCM concentra ainda mais a energia em torno de zero, o que contribui para uma melhor compressão geral, especialmente quando combinada com técnicas de codificação entrópica como a codificação de Huffman ou aritmética.

4. Análise dos Resultados

4.1 Efeito do Fator de Qualidade

Com base nas etapas implementadas no encoder e decoder, a avaliação “end-to-end” do algoritmo de compressão JPEG permite tirar conclusões abrangentes sobre o desempenho do sistema. Ao codificar e decodificar as imagens com fatores de qualidade variados (10, 25, 50, 75 e 100), torna-se possível quantificar o impacto da compressão na qualidade da imagem.

QF	Observações
10	Artefactos visíveis, perda significativa de detalhes
25	Melhoria na qualidade, ainda com distorções perceptíveis
50	Equilíbrio razoável entre compressão e qualidade
75	Boa qualidade, imprecisões mínimas
100	Qualidade próxima à original, compressão mínima

Tabela 5 – Efeito do fator de qualidade

A visualização das imagens decodificadas revela que fatores de qualidade mais baixos (e.g., 10 e 25) introduzem perda de detalhes, enquanto fatores mais altos (e.g., 75 e 100) preservam melhor a qualidade visual.

A imagem das diferenças, calculada entre o canal Y da imagem original e da imagem decodificada, permite identificar as áreas onde a compressão causou as maiores alterações.

4.2 Impacto do Tamanho dos Blocos DCT

Tamanho do Bloco	Observações
1x1	Preserva detalhes finos, mas resulta em menor compressão
8x8	Bom equilíbrio entre compressão e preservação de detalhes
64x64	Maior compressão, mas pode resultar em distorções de bloco mais visíveis

Tabela 6 – Efeito do tamanho do bloco

A DCT em blocos menores (8x8) preserva melhor detalhes finos e adapta-se a variações locais da imagem, enquanto blocos maiores (64x64) promovem maior compressão, mas podem introduzir blocos visíveis. O tamanho ideal do bloco DCT depende do compromisso desejado entre taxa de compressão e preservação da qualidade da imagem.

4.3 Métricas de Distorção

O cálculo das métricas de distorção (MSE, RMSE, SNR e PSNR) fornece uma avaliação quantitativa da qualidade da imagem reconstruída.

Espera-se que, à medida que o fator de qualidade diminui, o MSE e o RMSE aumentem, indicando maior erro, enquanto o SNR e o PSNR diminuem, refletindo uma menor relação sinal-ruído e pior qualidade da imagem.

As métricas *max_diff* e *avg_diff*, calculadas para o canal Y, fornecem informações sobre a máxima e a média diferença de brilho entre a imagem original e a reconstruída.

4.4 Comparação Imagens

Ao analisar em conjunto as imagens visualizadas, a imagem das diferenças e as métricas de distorção, podemos tirar conclusões sobre o efeito do processo de compressão na qualidade das imagens.

A análise comparativa com as conclusões tiradas na alínea 1 permite validar se as observações iniciais sobre o impacto dos fatores de qualidade se mantêm consistentes com a avaliação “end-to-end” do sistema de compressão.

Com base na tabela abaixo, usando QF = 75 e o método de interpolação cúbico, podemos ver que a imagem *nature.bmp* é a que tem um menor MSE, ou seja, em que o quadrado da soma do erro de cada pixel é mais pequeno. Observamos também que a imagem *nature.bmp* fica com uma qualidade aceitável quando restituída.

Imagem	Original	75%	MSE	Avg Diff
airport.bmp	<i>Muito Alta</i>	<i>Muito Alta</i>	<i>86.59</i>	<i>2.62</i>
geometric.bmp	<i>Alta</i>	<i>Alta</i>	<i>37.64</i>	<i>0.45</i>
nature.bmp	<i>Muito Alta</i>	<i>Muito Alta</i>	<i>27.31</i>	<i>2.00</i>

Tabela 7 – Análise qualitativa das imagens após compressão

5. Conclusão

O algoritmo implementado demonstra eficazmente as principais etapas do processo de compressão JPEG. A análise revela que:

1. A escolha do fator de qualidade tem um impacto significativo no equilíbrio entre compressão e qualidade da imagem.
2. O downsampling 4:2:0 oferece maior compressão, enquanto o 4:2:2 preserva mais informações de cor.
3. A DCT em blocos 8x8 proporciona um bom equilíbrio entre preservação da qualidade de imagem e taxa de compressão.
4. A quantização é a etapa mais crítica para o controle da taxa de compressão e qualidade da imagem.
5. A codificação DPCM dos coeficientes DC contribui para a redução adicional do tamanho dos dados comprimidos.

Este trabalho fornece uma base sólida para a compreensão e implementação do algoritmo de compressão JPEG, destacando os trade-offs entre compressão e qualidade da imagem em cada etapa do processo.