

TRABALHO PRÁTICO 1

MULTIMÉDIA

RELATÓRIO



Alunos:

Stefanos Panagoulas Lucena, 2021115090

Ricardo Rafael Dias Gomes Simões, 2019231869

Rodrigo Santos de Oliveira, 2018305489

João Guilherme Guedes Krichanã, 2019112913

Introdução:

Neste relatório, apresentamos os resultados do nosso trabalho prático focado na compressão de imagem utilizando o codec JPEG, sob orientação dos professores Ricardo Correia Nascimento dos Santos e Rui Pedro Paiva. O projeto foi uma oportunidade para aprofundarmos nosso entendimento sobre os mecanismos e técnicas de compressão de imagem, aplicando conhecimentos teóricos na prática por meio da linguagem Python. Desde a manipulação de imagens no modelo de cor RGB até a implementação de processos complexos como sub-amostragem, DCT (Transformada de Coseno Discreta), quantização e codificação DPCM, buscamos compreender cada etapa do processo de compressão e descompressão.

1. Compressão de imagens bmp no formato jpeg utilizando um editor de imagem.

tabela 1.1

Nature	Q = 25	Q = 50	Q = 75
Qualidade Alta		X	X
Qualidade Média	X		
Qualidade Baixa			

tabela 1.2

Airport	Q = 25	Q = 50	Q = 75
Qualidade Alta			X
Qualidade Média	X	X	
Qualidade Baixa	-		

tabela 1.3

Geometric	Q = 25	Q = 50	Q = 75
Qualidade Alta			
Qualidade Média			X
Qualidade Baixa	X	X	

tabela 1.4

Taxa de compressão	Q = 25	Q = 50	Q = 75
Airport.bmp	1:36	1:23	1:15
Geometric.bmp	1:87	1:65	1:49
Nature.bmp	1:48	1:19	1:14

1.4. Compare os resultados e tire conclusões.

Quanto maior a taxa de compressão, pior poderá ser a qualidade da imagem. Isso pode ser evidenciado pela compressão da imagem `geometric.bmp` para `jpg`, já que essa imagem apresentou as maiores taxas de compressão e, conseqüentemente, as piores avaliações em termos de qualidade visual. (tabelas 1.3 e 1.4).

A imagem `geometric.bmp` teve maior compressão e qualidade visual reduzida devido a sua composição. Ela contém áreas grandes e uniformes, o que facilita a compressão, pois padrões repetitivos e cores consistentes são mais eficientemente codificados pelos algoritmos de compressão. Além disso, a falta de detalhes complexos permite uma maior redução do tamanho do arquivo sem impactar significativamente a percepção visual, uma vez que detalhes visualmente irrelevantes podem ser descartados ou simplificados. Em resumo, a natureza simples dos gráficos e a homogeneidade da imagem contribuem para taxas de compressão mais altas com menor impacto aparente na qualidade.

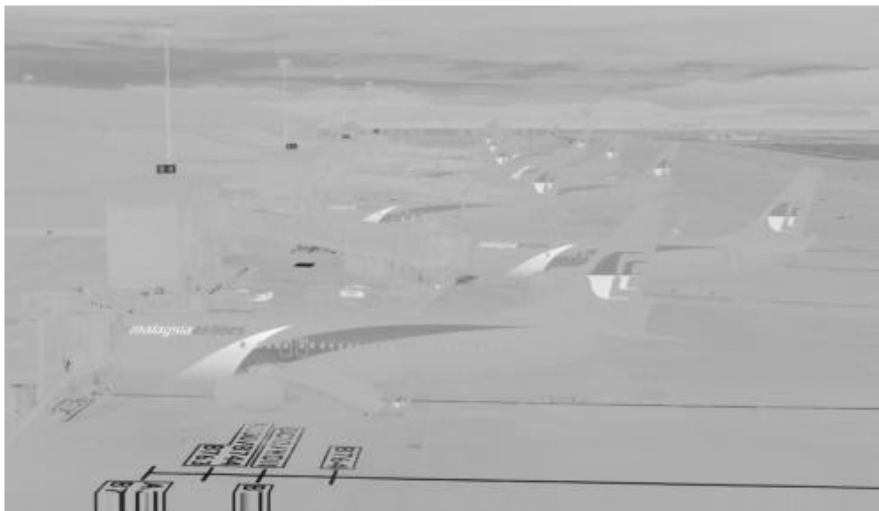
5. Conversão para o modelo cor YCbCr.

5.5. Compare a imagem de Y com R, G e B e com Cb e Cr. Tire conclusões.

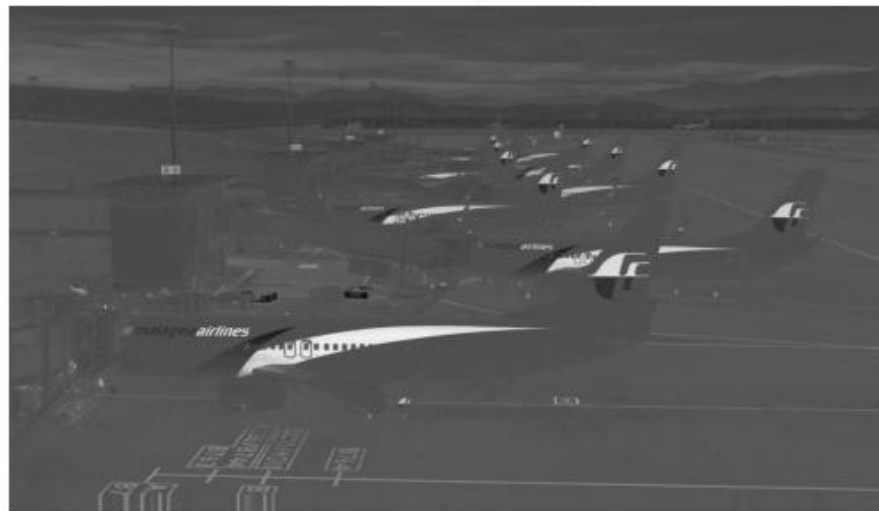
Canal Y airport.bmp



Canal Cb airport.bmp



Canal Cr airport.bmp



Img Red Padded: airport.bmp



Img Green Padded: airport.bmp



Img Blue Padded: airport.bmp



Y (Luminância) vs. R, G, B (Componentes de Cor):

Y (Luminância): Representa a intensidade de luz (brilho) da imagem, sem considerar a cor. Imagens baseadas na luminância mostram variações de claro e escuro, permitindo que observemos detalhes, texturas e formas sem a influência da cor.

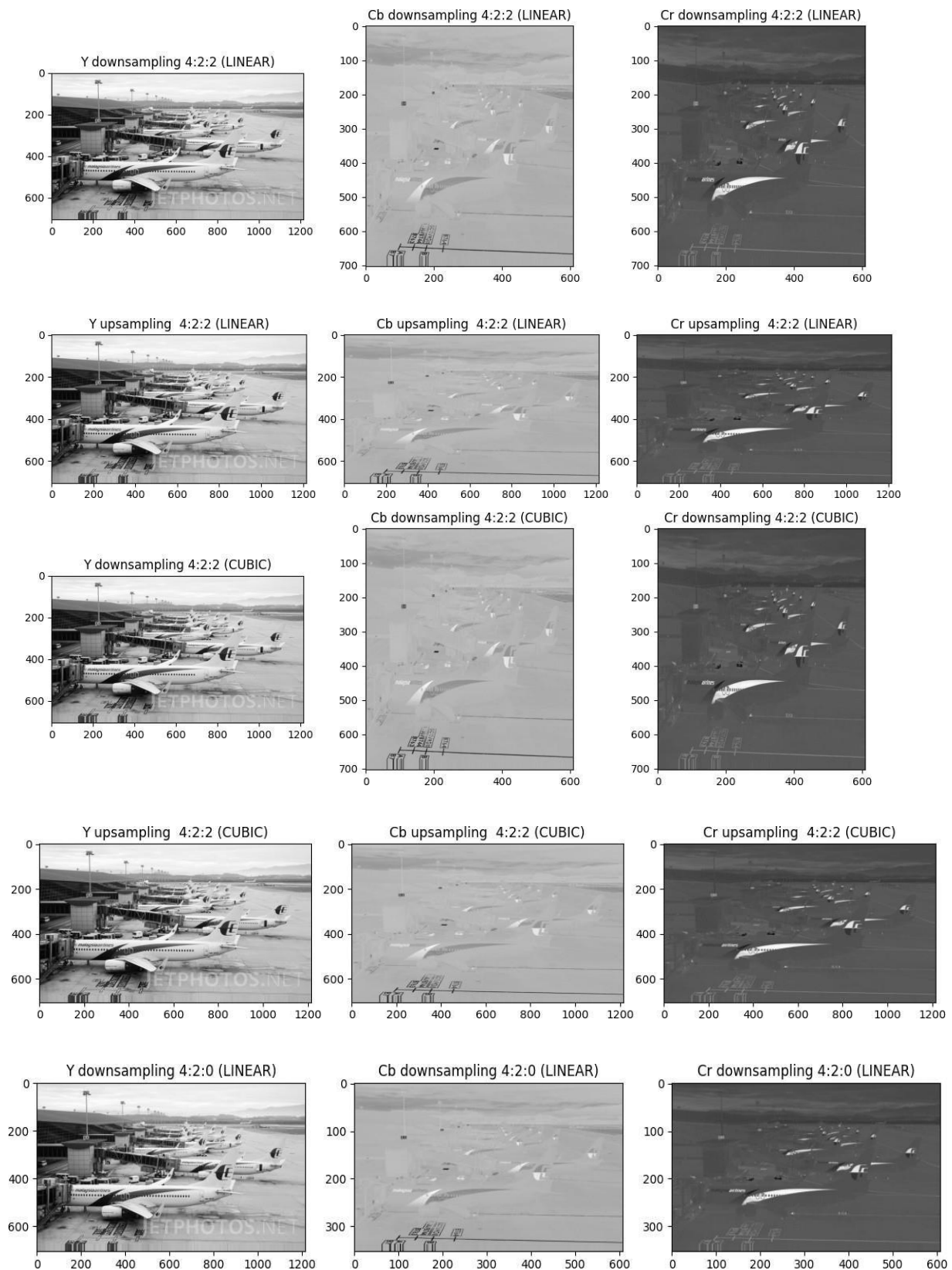
R, G, B (Componentes de Cor): Cada um desses canais carrega informações específicas sobre a quantidade de vermelho, verde e azul em cada pixel da imagem. A combinação desses três canais produz a imagem colorida que percebemos. Comparar Y com R, G e B individualmente pode mostrar como cada componente de cor contribui para o brilho percebido da imagem. Por exemplo, o canal verde geralmente carrega mais informações de luminância devido à maior sensibilidade do olho humano a esta cor.

Y (Luminância) vs. Cb, Cr (Componentes de Crominância)

Cb e Cr: Representam as informações de cor sem o brilho, especificamente as diferenças de azul e vermelho em relação à luminância. Isso permite codificar a cor da imagem de forma mais eficiente, aproveitando a menor sensibilidade do olho humano às variações de cor em comparação com variações de brilho.

6. Sub-amostragem:

6.5. Apresente e analise o resultado da compressão para as variantes de downsampling 4:2:2 e 4:2:0 (taxa de compressão, destrutividade, etc.)



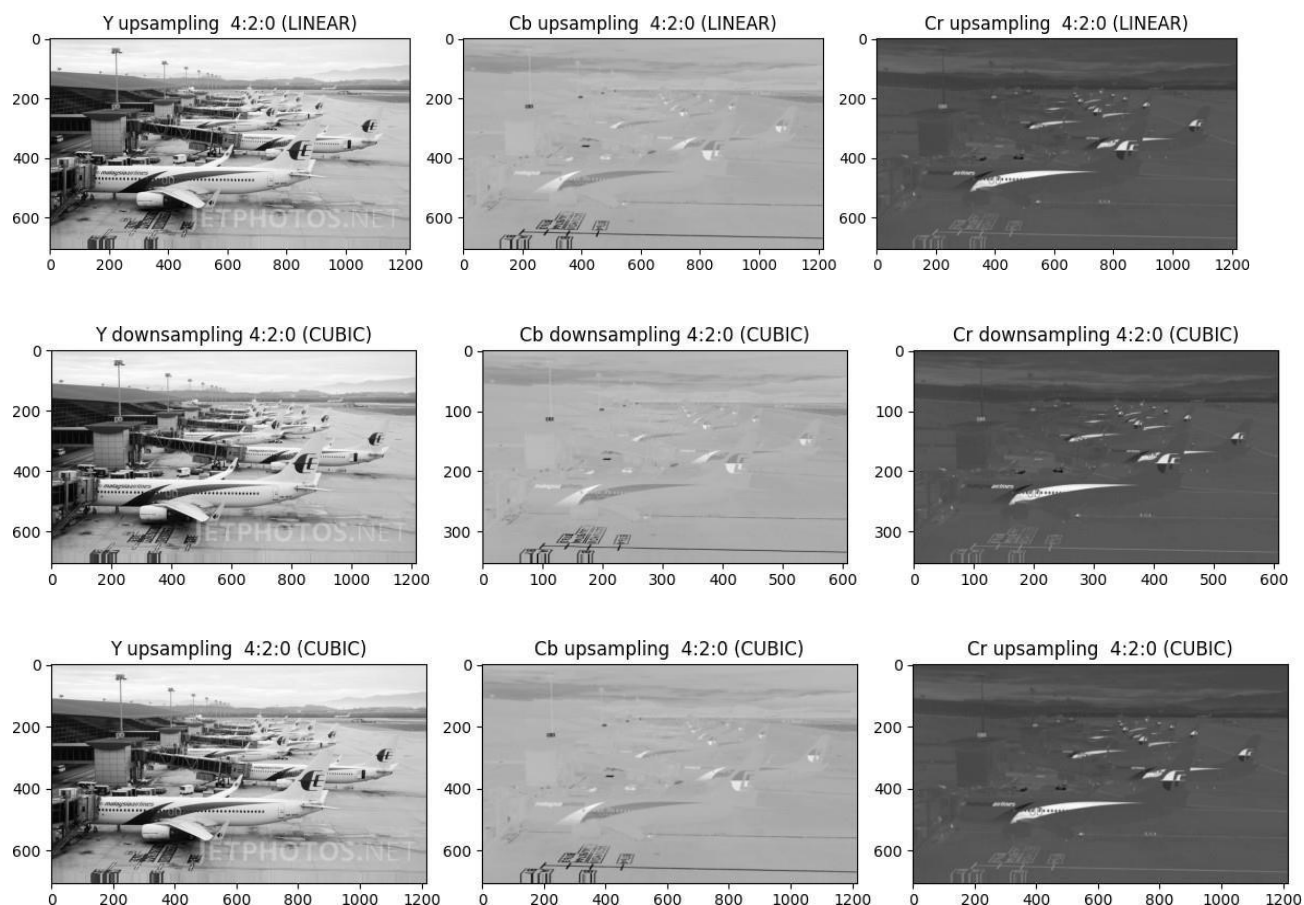


Tabela 6.1 – Dimensões Y, Cb, Cr downsampld

	Y	Cb	Cr
4:4:4	(704, 1216)	(704, 1216)	(704, 1216)
4:2:2	(704, 1216)	(704, 608)	(704, 608)
4:2:0	(704, 1216)	(352, 608)	(352, 608)

Na análise dos impactos do downsampling 4:2:2 e 4:2:0 com interpolação Linear e Cúbica, não se observou qualquer alteração visível em função da interpolação. Sendo assim, concluímos que é mais benéfico adotar o método de interpolação que seja computacionalmente mais eficiente, ou seja, o método linear.

Além disso, com base nos dados apresentados na [Tabela 6.1](#), é possível determinar a taxa de compressão teórica utilizando os cálculos a seguir:

a) Para a taxa de sub-amostragem de 4:4:4, temos
 $704 \times 1216 \times 3 = 2.568.192$ pixels, o que corresponde a uma **compressão de 0%**, ou seja, não há compressão.

b) Com a taxa de sub-amostragem de 4:2:2, o cálculo é
 $(704 \times 608 \times 2) + (704 \times 1216) = 1.712.128$ pixels, resultando em uma **compressão de 33,33%**.

c) Finalmente, para a taxa de sub-amostragem de 4:2:0, obtemos
 $(352 \times 608 \times 2) + (704 \times 1216) = 1.284.096$ pixels, o que significa uma **compressão de 50%**.

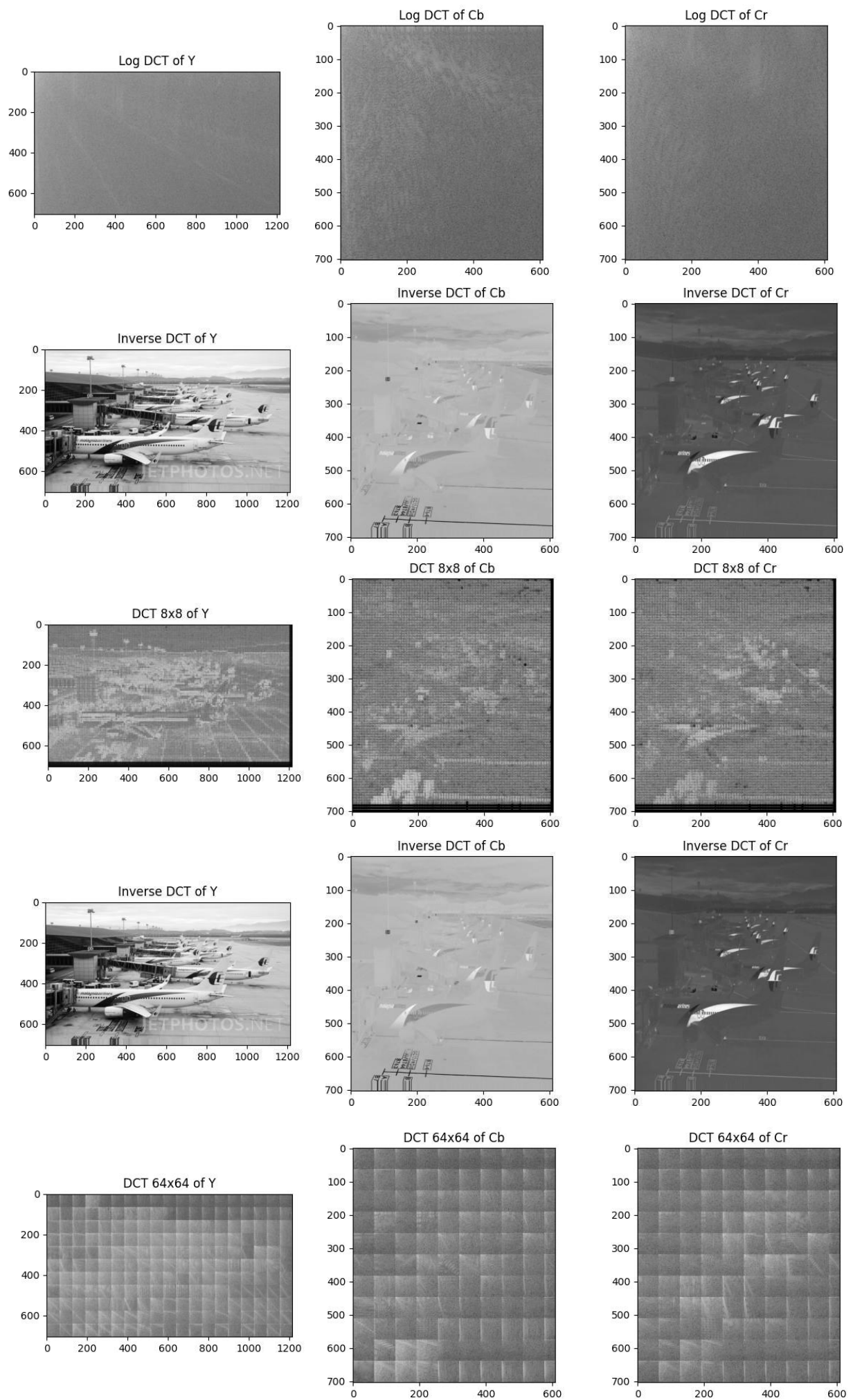
Esses cálculos demonstram como diferentes taxas de sub-amostragem afetam o tamanho dos dados de crominância, com uma redução significativa no número de pixels e, conseqüentemente, no tamanho do arquivo.

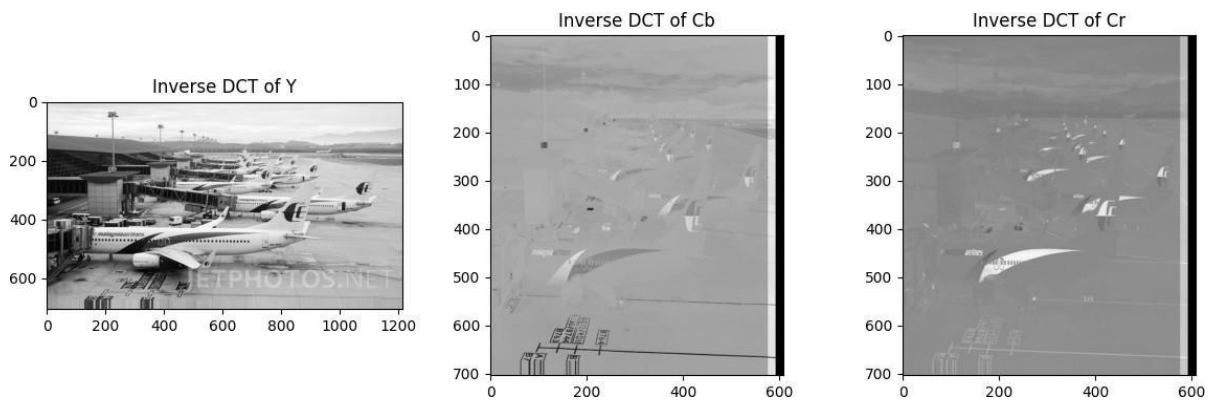
Por fim, ao aplicar a compressão nos canais Cb e Cr, a destrutividade se refere à perda de qualidade de imagem resultante do processo.

- a) 4:4:4: Não há destrutividade, pois não há compressão. A imagem mantém a totalidade da informação de cor.
- b) 4:2:2: Há uma compressão moderada. A qualidade da imagem é ligeiramente afetada, mas geralmente é aceitável para o olho humano.
- c) 4:2:0: Compressão mais agressiva, com redução de 50% das informações de cor. Isso pode levar a uma perda de qualidade mais perceptível, principalmente em imagens com detalhes de cor críticos.

Quanto maior a compressão, maior a destrutividade, afetando potencialmente a fidelidade da cor na imagem reconstruída.

7. Transformada de Coseno Discreta (DCT)



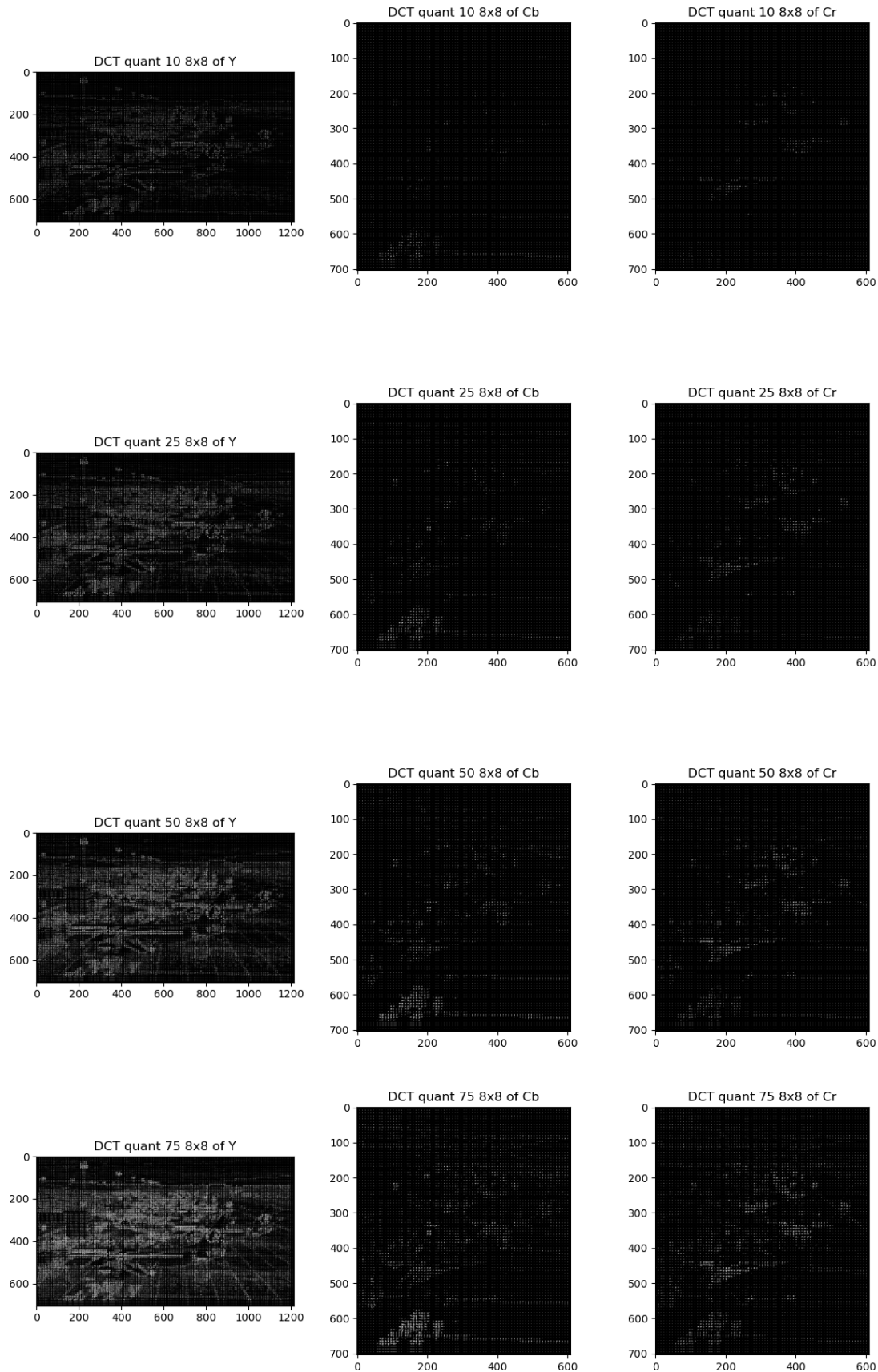


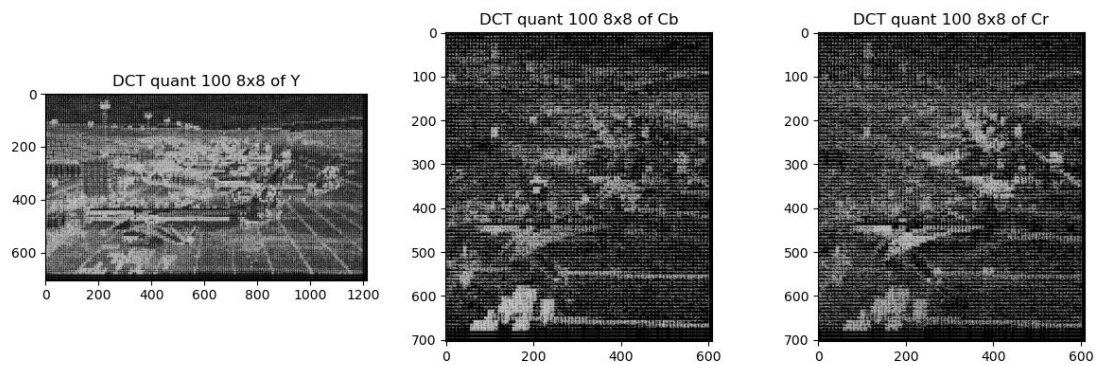
7.4. Compare e discuta os resultados obtidos em 7.1, 7.2 e 7.3 em termos de potencial de compressão.

A escolha do método de Transformada Discreta de Cosseno (DCT) — seja em canais completos, blocos de 8x8 ou de 64x64 — é influenciada pelo objetivo de compressão, a qualidade da imagem desejada e a disponibilidade de recursos computacionais. Blocos de 8x8 são um equilíbrio eficaz entre compressão e qualidade, comum no padrão JPEG. Blocos de 64x64 se adequam a situações que demandam compressão intensa com perdas de detalhes toleráveis, enquanto a DCT em canais completos é ideal para a máxima preservação de qualidade sem restrições de recursos. A eficácia da compressão varia com a uniformidade da imagem, com blocos de 8x8 preferíveis para detalhes e 64x64 para áreas homogêneas, permitindo uma compressão mais forte. Portanto, a seleção do método depende da finalidade específica, da qualidade visada e das características da imagem, visando alinhar a capacidade de processamento à preservação da qualidade.

8. Quantização.

8.5. Compare os resultados obtidos com os vários factores de qualidade e discuta-os em termos de potencial de compressão.





Baixos Fatores de Qualidade (10, 25): A aplicação de baixos fatores de qualidade resulta em uma quantização mais agressiva, o que significa uma maior perda de informações. Isso leva a uma alta taxa de compressão, mas também pode resultar em uma qualidade de imagem perceptivelmente mais baixa, especialmente em detalhes finos e texturas.

Fatores de Qualidade Médios (50): Um fator de qualidade médio busca um equilíbrio entre compressão e preservação da qualidade da imagem. Isso é útil quando é necessário reduzir o tamanho dos dados, mas ainda é importante manter uma boa qualidade visual.

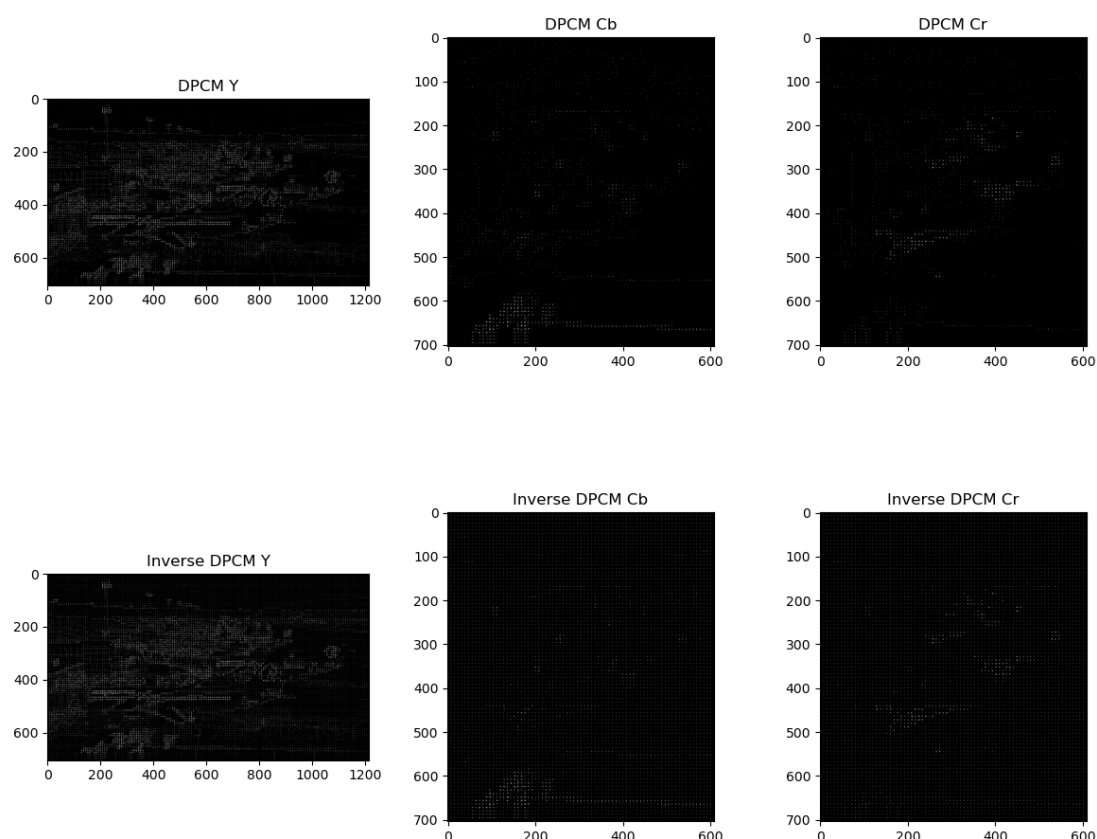
Altos Fatores de Qualidade (75, 100): Altos fatores de qualidade resultam em uma quantização menos agressiva, preservando mais informações dos coeficientes DCT originais. Isso reduz a taxa de compressão, mas mantém uma alta qualidade de imagem, sendo ideal quando a qualidade da imagem é prioritária.

8.6. Compare os resultados obtidos com os da alínea 7 (DCT) e tire conclusões.

DCT sem Quantização: A aplicação da DCT sem qualquer quantização (ou com uma quantização muito leve) permite uma transformação reversível, onde os dados originais podem ser recuperados quase perfeitamente. No entanto, isso não resulta em uma compressão significativa dos dados, pois todos os coeficientes DCT são mantidos.

DCT com Quantização: A quantização altera significativamente o potencial de compressão da DCT, permitindo reduzir o tamanho dos dados ao custo de perder informações. O grau de perda de informação e compressão depende do fator de qualidade utilizado na quantização.

9. Codificação DPCM dos coeficientes DC.



9.5. Analise os resultados e tire conclusões.

A conclusão principal é que a técnica de codificação DPCM está sendo eficaz na representação e compressão dos coeficientes DC, pois consegue realizar isso com uma quantidade mínima de informação adicionada (a diferença entre os valores DC consecutivos). Isso implica que a codificação e decodificação DPCM estão mantendo a qualidade visual do conteúdo de imagem, enquanto reduzem a quantidade de dados necessários para armazenamento ou transmissão, o que é fundamental em sistemas de compressão de imagem.

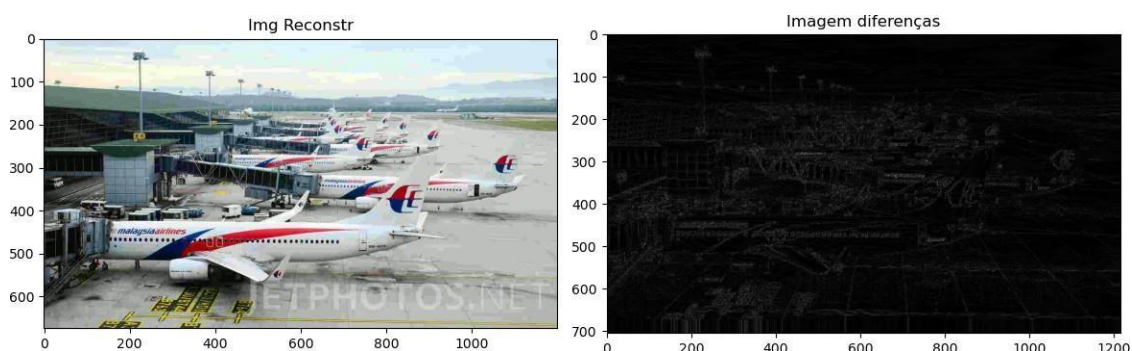
10. Codificação e decodificação end-to-end.

10.5. Visualize as imagens decodificadas. Visualize também a imagem das diferenças entre o canal Y de cada uma das imagens originais e da imagem decodificada respectiva para cada um dos factores de qualidade testados. Calcule as várias métricas de distorção (imagem RGB: MSE, RMSE, SNR, PSNR; canal Y: max_diff e avg_diff) para cada uma das imagens e factores de qualidade. Tire conclusões.

AIRPORT:

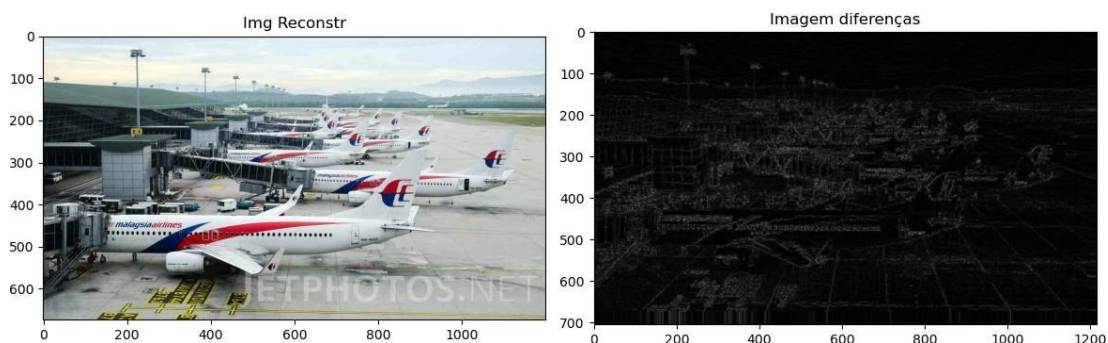
Q = 10

Imagem 10.1



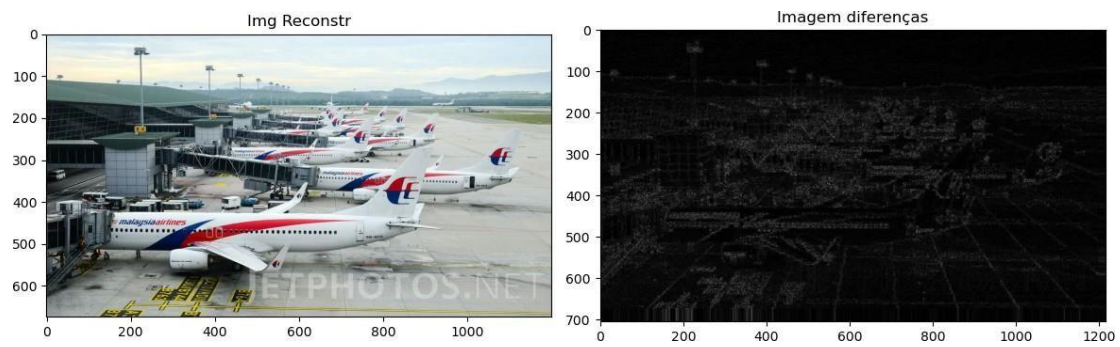
Q = 25

Imagem 10.2



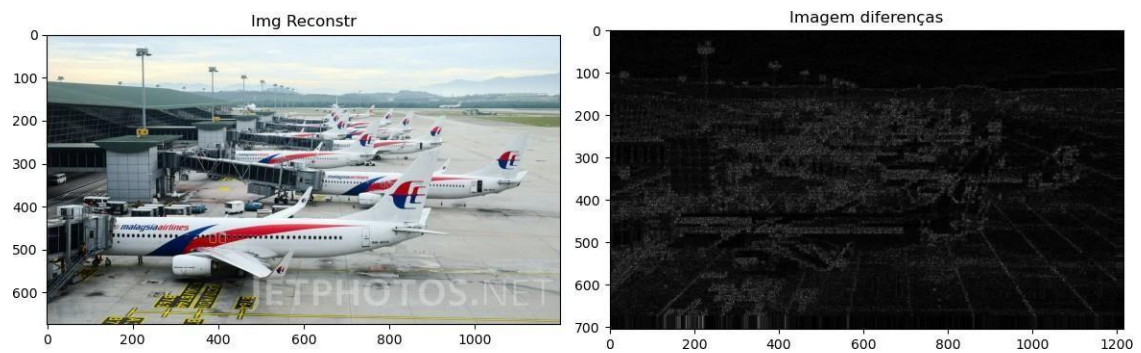
Q = 50

Imagem 10.2



Q = 75

Imagem 10.3



Q = 100

Imagem 10.4

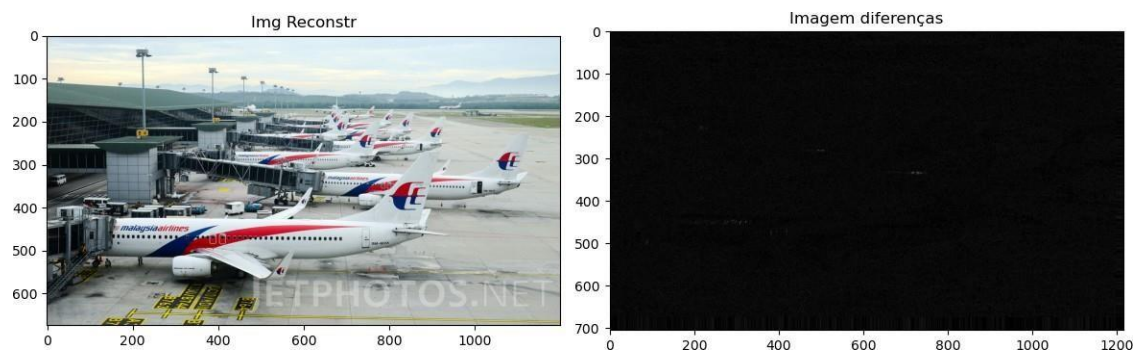


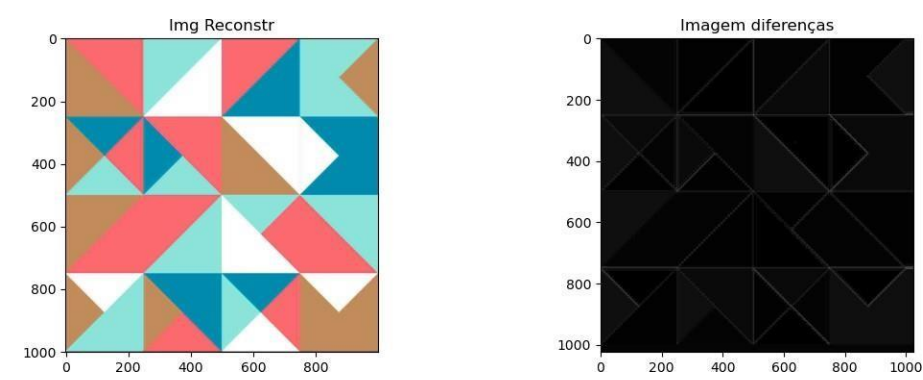
Tabela 10.1 – métricas de distorção (Airport)

AIRPORT	Q = 10	Q = 25	Q = 50	Q = 75	Q = 100
MSE	572.4591024691358	290.6768493827166	167.68662962962964	89.5693561728394	15.68877777777778
RMSE	23.92611758035841	17.04924775689032	12.94938722988967	9.46418548473224	3.959794575497159
SNR	22.07436659358544	25.0168808587845	27.485928890258112	30.12931999226985	37.71721627252807
PSNR	20.553358955440586	23.4966991659569	25.8858192533216	28.69209353380993	36.19711463262011
Max_diff	188.195	98.59999999999995	76.316	42.812999999999996	15.884000000000003
Avg_diff	7.289554610482959	5.016413863916717	3.725145792837917	2.788736386838385	0.557267158763831

GEOMETRIC:

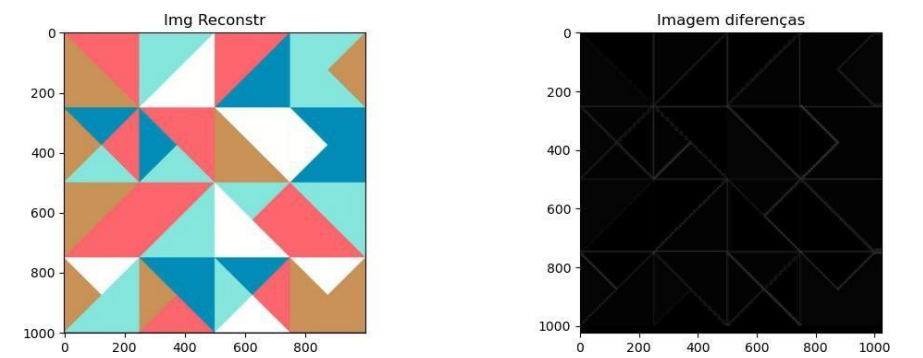
Q = 10

Imagem 10.5



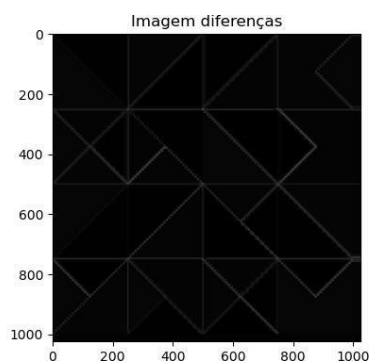
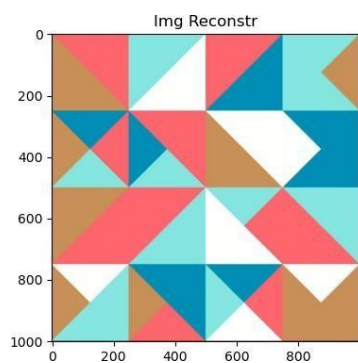
Q = 25

Imagem 10.6



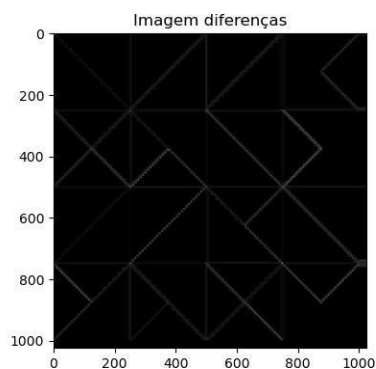
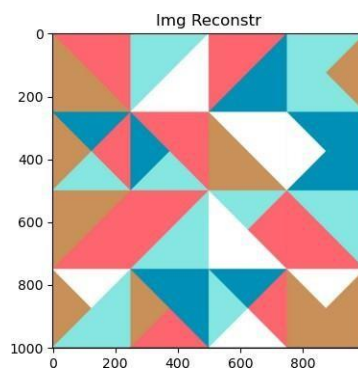
Q = 50

Imagem 10.7



Q = 75

Imagem 10.8



Q = 100

Imagem 10.9

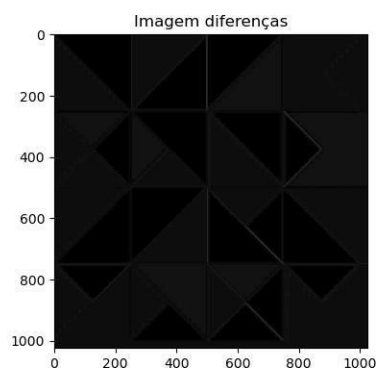
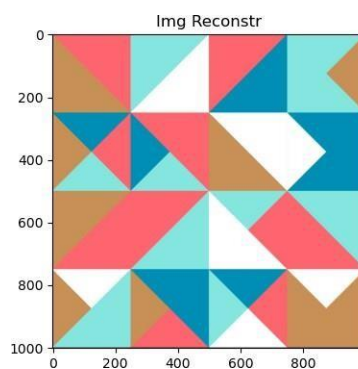


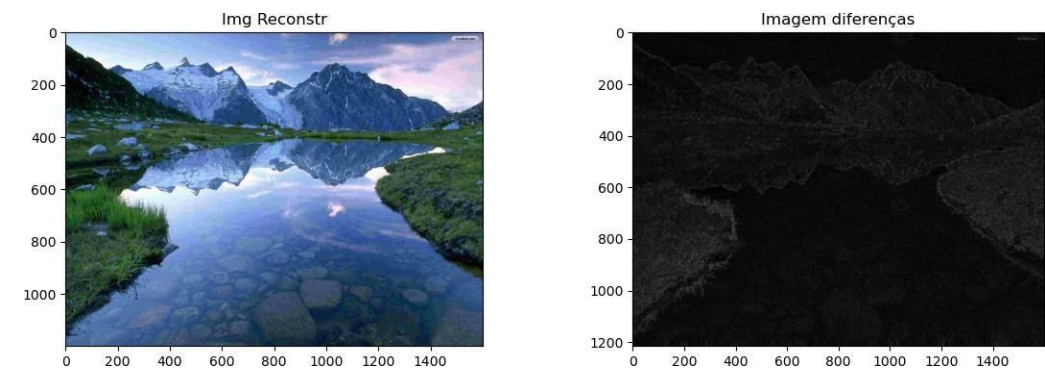
Tabela 10.2 – métricas de distorção (Geometric)

GEOMETRIC	Q = 10	Q = 25	Q = 50	Q = 75	Q = 100
MSE	225.42971428571428	188.32735364635364	67.9163469538469	41.19154345654346	21.985019584015984
RMSE	15.01431697699133	10.48842738497527	8.241134915489777	6.418863840173555	4.68881816282764
SNR	26.60033637893857	29.7836646967853	31.810875844087656	33.98267331119886	36.7091812787951
PSNR	24.60069280267924	27.78342227088527	29.811063128148326	31.98272295493156	34.7953575162019
Max_diff	80.421	60.91700000000003	39.85900000000001	27.435000000000002	9.866999999999962
Avg_diff	2.7364947729110765	1.1622383966445995	0.9773145179748486	0.34965447807312094	0.391803174972534

NATURE:

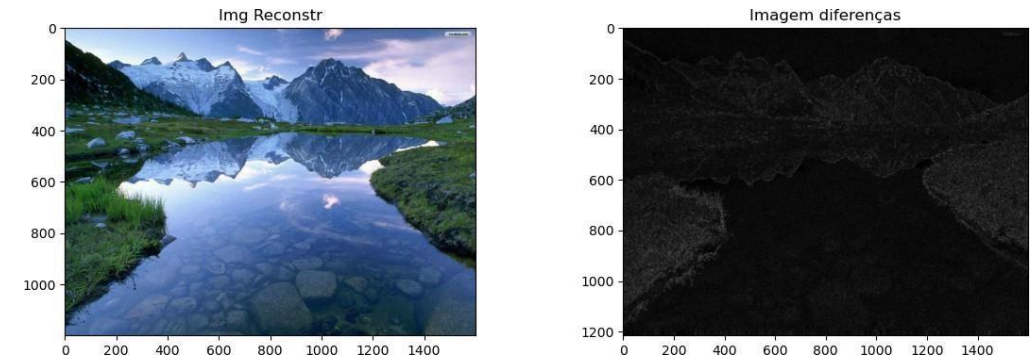
Q = 10

Imagem 10.10



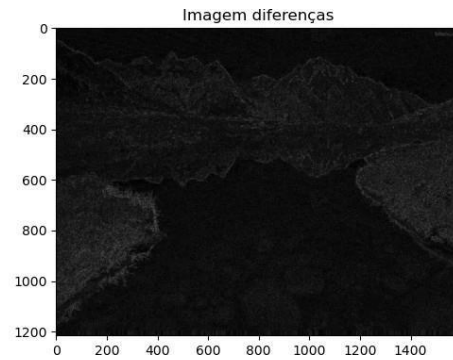
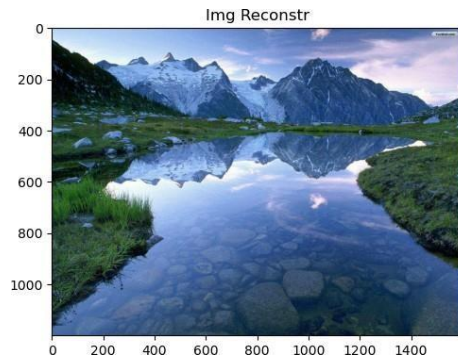
Q = 25

Imagem 10.11



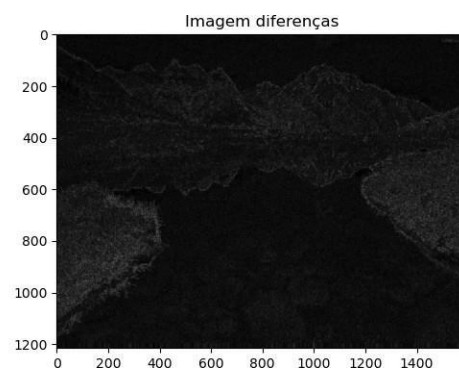
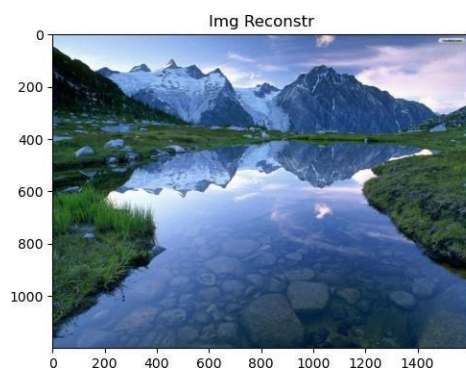
Q = 50

Imagem 10.12



Q = 75

Imagem 10.13



Q = 100

Imagem 10.14

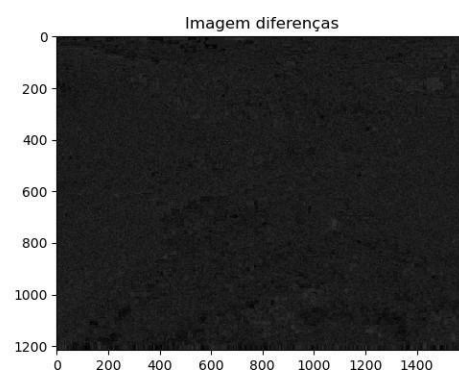
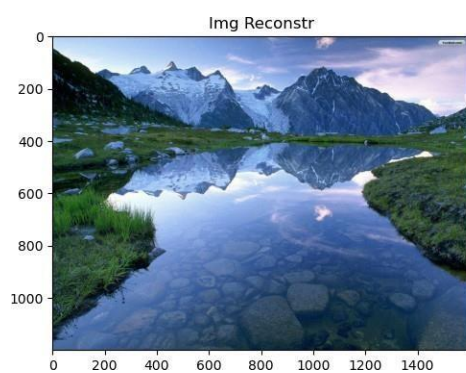


Tabela 10.3 – métricas de distorção (Nature)

NATURE	Q = 10	Q = 25	Q = 50	Q = 75	Q = 100
MSE	565.318503125	382.21681979166667	38.78689695833334	30.312719728833334	5.474191145833333
RMSE	23.77652933214127	19.55345771665518	6.221624377598898	5.505698799501598	2.3396989434184334
SNR	19.44997678196686	21.149758470853822	31.89471757449942	32.1565983517726	39.58964774781182
PSNR	20.69793864575477	22.30771475215861	32.25272785658721	33.3145546386939	40.74768402998661
Max_diff	134.69	117.157	30.179000000000002	27.99999999999997	4.729999999999999
Avg_diff	9.037499262483827	7.74987598381574	2.1918397898592107	2.032639136513156	0.5342299444449812

MSE (Mean Squared Error): Essa métrica mede a média dos quadrados dos erros, ou seja, a média quadrática da diferença entre os valores de pixel originais e comprimidos. Um valor MSE menor indica uma qualidade de imagem mais alta. Nas imagens, o MSE diminui conforme aumentamos a qualidade, o que é esperado.

RMSE (Root Mean Squared Error): É a raiz quadrada do MSE, proporcionando uma métrica de erro na mesma unidade dos dados da imagem. A tendência é semelhante ao MSE; menores valores de RMSE correspondem a uma melhor qualidade de imagem.

SNR (Signal-to-Noise Ratio): Esta é uma medida da relação entre o sinal desejado (a imagem original) e o ruído de fundo (erros introduzidos pela compressão). Valores mais altos de SNR indicam uma imagem de maior qualidade. Os valores de SNR aumentam com a qualidade nas imagens fornecidas.

PSNR (Peak Signal-to-Noise Ratio): Derivado do MSE, o PSNR é utilizado para medir a qualidade de imagens reconstruídas após compressão. Quanto maior o PSNR, melhor é considerada a qualidade da imagem. Os dados obtidos mostram que o PSNR aumenta com a qualidade da imagem.

Max_diff e Avg_diff: "max_diff" refere-se à maior diferença entre os valores de pixel no canal Y da imagem original e da imagem comprimida, enquanto "avg_diff" refere-se à média dessas diferenças. Idealmente, queremos que ambos os valores sejam baixos, indicando que a compressão teve um impacto mínimo na qualidade percebida da imagem. As imagens mostram que, conforme a qualidade melhora, o "max_diff" e o "avg_diff" diminuem.

Conclusões:

A qualidade da imagem melhora à medida que o fator de qualidade aumenta. Isso é demonstrado pelo decréscimo nos valores de MSE e RMSE, e pelo aumento dos valores de SNR e PSNR.

A imagem diferença se torna progressivamente mais escura à medida que aumentamos o fator de qualidade, o que indica menos diferenças entre a imagem original e a comprimida.

No extremo mais alto de qualidade (QUALIDADE 100); [imagens 10.4, 10.9, 10.14](#); as diferenças entre a imagem original e a comprimida são mínimas, com a imagem diferença mais escura e métricas de distorção apresentando os melhores valores (menor MSE e maior PSNR).

Essa análise sugere que uma taxa de compressão mais alta (fator de qualidade mais baixo) resulta em maior perda de detalhes da imagem, o que pode ser aceitável dependendo do uso pretendido da imagem. Em aplicações onde a qualidade da imagem é crítica, uma taxa de compressão mais baixa (fator de qualidade mais alto) seria necessária.

10.6. Volte a analisar a alínea 1, de forma a validar/complementar as conclusões tiradas nesse ponto.

A análise revela uma relação diretamente proporcional entre a qualidade das imagens e a redução da compressão; à medida que a compressão diminui, as imagens aproximam-se mais da sua forma original, evidenciando melhorias nas métricas de distorção. Identificamos que um fator de qualidade de 75 equilibra de forma ideal a redução da compressão sem sacrificar excessivamente a qualidade visual.

Portanto, a diminuição da taxa de compressão resulta na preservação de mais detalhes da imagem original, reduzindo a distorção visual. Assim, a imagem comprimida retém mais de sua qualidade original.

Os dados suportam firmemente esta observação, evidenciando que uma abordagem menos agressiva na compressão favorece uma melhor qualidade de imagem, como mostrado pelas métricas de distorção que melhoram com taxas de compressão mais baixas.