

Multimédia 2023/24

Trabalho prático 1



Docente: Rui Pedro Pinto de Carvalho e Paiva

Da autoria de: Ana Patrícia Fernandes Gonçalves (2002009904)

Cíntia Dalila Luís Cumbane (2020244607)

Cristiana de Brito Gonçalves (2019239753)

Stela Ângelo Macuácua (2019239770)

Índice

1.	Introdução	3
2.	Compressão de imagens bmp no formato jpeg	3
3.	Conversão do modelo de cor RGB para o modelo de cor YCbCr	5
3.1	Comparação do canal Y com o modelo RGB	5
3.2	Comparação do modelo YCbCr e modelo RGB	6
4.	Resultados da compressão para as variantes de downsampling.....	6
5.	Transformada de Cosseno Discreta (DCT)	8
6.	DCT em blocos 8x8	8
7.	Quantização	9
8.	Codificação DPCM dos coeficientes DC.....	11
9.	Métricas de distorção (MSE, RMSE, SNR e PSNR) com os diferentes fatores de qualidade.....	12
10.	Conclusão	16

1. Introdução

No contexto da disciplina de Multimédia, foi sugerida a implementação de um algoritmo de compressão com perda (JPEG Lossy), com o objetivo de compreender o comportamento deste género de codecs e aprofundar conhecimento dos princípios fundamentais da compressão de imagem.

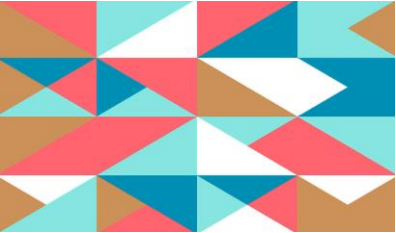
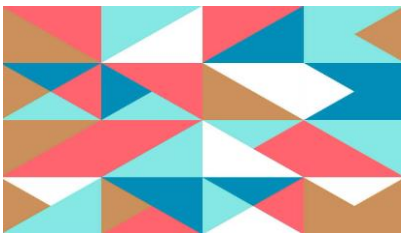
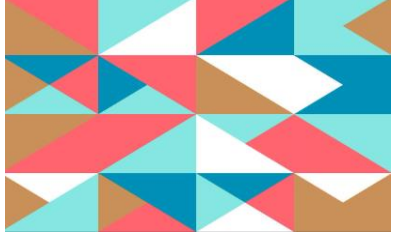
2. Compressão de imagens BMP no formato JPEG

Antes de prosseguir com o projeto, a primeira etapa consistiu em realizar a compressão das imagens fornecidas no formato BMP para o formato JPEG. Essa compressão foi realizada utilizando o editor de imagem GIMP. O objetivo foi obter uma variedade de resultados experimentais para posterior comparação, a fim de entender como a escolha da qualidade de compressão afeta o tamanho das imagens resultantes. Adicionalmente, realizamos o cálculo das taxas de compressão. Para isso, foi usada a fórmula: $Tx = \frac{\text{tamanho da imagem original}}{\text{tamanho da imagem comprimida}}$.

- Imagem original: airport.bmp, apresenta tamanho de: 2,31 MB (2 430 122 bytes)

Taxas de Compressão		
+25	+50	+75
 Tamanho: 66,3 KB (67 916 bytes) Taxa de Compressão: 35,7812887	 Tamanho: 102 KB (105 143 bytes) Taxa de Compressão: 23,112542	 Tamanho: 154 KB (158 287 bytes) Taxa de Compressão: 15,352632

- Imagem original: geometric.bmp, apresenta tamanho de: 2,86 MB (3 003 122 bytes)

<i>Taxas de Compressão</i>		
+25	+50	+75
 <p>Tamanho: 33,6 KB (34 506 bytes) Taxa de Compressão: 87,0318785</p>	 <p>Tamanho: 45,2 KB (46 387 bytes) Taxa de Compressão: 64,7405954</p>	 <p>Tamanho: 59,2 KB (60 681 bytes) Taxa de Compressão: 49,4903182</p>

- Imagem original: nature.bmp, apresenta tamanho de: 5,49 MB (5 760 122 bytes)

<i>Taxas de Compressão</i>		
+25	+50	+75
 <p>Tamanho: 117 KB (119 910 bytes) Taxa de Compressão: 48,037044</p>	 <p>Tamanho: 303 KB (310 963 bytes) Taxa de Compressão: 18,5235</p>	 <p>Tamanho: 401 KB (411 381 bytes) Taxa de Compressão: 14,001916</p>

Com base nas taxas de compressão calculadas para cada uma das diferentes imagens, pode concluir-se que existe uma relação entre compressão e qualidade, ou seja quanto maior for a qualidade da imagem comprimida, menor é a taxa de compressão. Analisando cada uma das imagens individualmente podemos observar que:

➤ **Airport:**

Apresenta uma taxa de compressão razoável, embora não muito elevada. Isso sugere que, dada a presença de transições abruptas entre pixels adjacentes na imagem, o codec JPEG Lossy pode não estar a atuar de forma tão eficaz quanto poderia.

➤ **Geometric:**

Mostrou-se ser altamente compressível, com taxas de compressão significativamente superiores às das outras imagens, possivelmente devido à sua natureza geométrica e padrões repetitivos.

➤ **Nature:**

Tem uma taxa de compressão menos eficiente em comparação com as outras. Isso pode ser devido à complexidade da imagem, com muitos detalhes e variações de cor.

3. Conversão do modelo de cor RGB para o modelo de cor YCbCr

Procedemos à conversão do modelo de cor RGB para YCbCr com o intuito de aderir ao princípio da compressão perceptual. Este princípio visa reduzir a redundância ao nível da luminescência e da crominância ao tirar partido das limitações e características da visão humana.

3.1 Comparação do canal Y com o modelo RGB

O canal Y, responsável por capturar a luminescência e o brilho de uma imagem, revela uma notável discrepância de detalhes em relação aos canais R e G, quando comparado com o canal B. Este fenómeno decorre da maior sensibilidade do olho humano aos tons de verde e vermelho. Especificamente no canal G, podemos observar uma qualidade de detalhamento significativamente superior. Com o objetivo de reduzir a redundância a partir da luminescência e aplicar downsampling, prosseguimos com o processo.

As imagens obtidas passaram anteriormente por um processo de padding, com o objetivo de torná-las compatíveis com dimensões múltiplas de 32 bits. Este procedimento é necessário para garantir que as imagens estejam alinhadas de forma adequada às exigências de processamento do Codec JPEG, para otimização de desempenho.



Figura. 1 rgb

Pela observação da figura 1, fica evidente que cada componente RGB apresenta considerável redundância de luminância e crominância, o que traduz a alta correlação entre os três canais. No entanto, ao realizar a conversão para YCbCr, os canais tornam-se descorrelacionados. Essa descorrelação é essencial para aumentar a capacidade de compressão, pois agora a luminância é representada por apenas um canal (Y), enquanto as informações de crominância são distribuídas em dois canais (Cb e Cr). A fórmula $Y \approx 0.3R +$

$0.6G + 0.1B$ é utilizada porque o sistema visual humano é mais sensível em primeiro ao verde e, depois ao vermelho e por fim ao azul.

3.2 Comparação do modelo YCbCr e modelo RGB

Os canais Cb e Cr são responsáveis pela crominância da imagem, onde o Cb indica a variação de azul em relação ao canal Y, enquanto o Cr reflete a variação de vermelho em relação ao canal Y. Essas relações podem ser apreciadas ao compararmos os canais Cb com o canal B e Cr com o canal R.

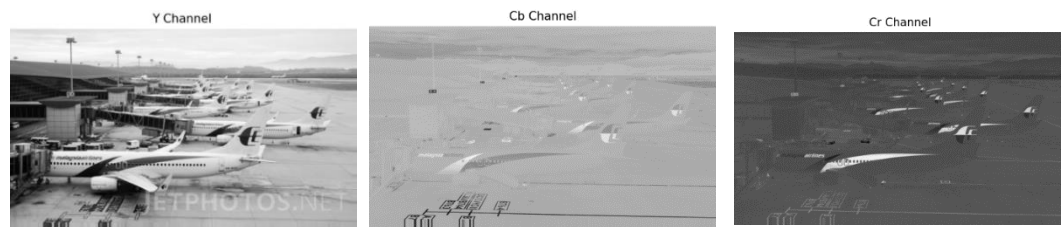


Figura 2. Canais Y Cb Cr

Na representação YCbCr, a crominância é capturada pelos canais Cb e Cr, como demonstrado na figura 2. É perceptível que há uma menor quantidade de detalhes nas componentes cromáticas em comparação com o canal Y, o que se deve ao fato de o olho humano ser mais sensível à informação de luminância do que às informações de cor. O modelo YCbCr beneficia dessa característica ao descartar informações não perceptíveis, otimizando assim o processo de compressão.

4. Resultados da compressão para as variantes de downsampling

É neste processo de compressão JPEG que ocorre a redução da resolução espacial nos canais de crominância Cb e Cr, enquanto o canal de luminância Y permanece inalterado. Isso é feito utilizando um método de interpolação que calcula a média dos valores de uma vizinhança de pixels 2x2 ao redor de cada pixel na imagem original.

Na variação 4:2:2, a crominância é amostrada a cada dois pixels horizontais, resultando em uma redução de 50% no número de colunas da imagem. Já na variação 4:2:0, a crominância é amostrada a cada dois pixels horizontal e verticalmente, levando a uma redução de 75% no número de colunas e linhas da imagem original.

Por conseguinte, como a variação 4:2:2 mantém mais informação nos canais de cor, a degradação da qualidade é menor em comparação com a variação 4:2:0. Na amostragem 4:2:0, a quantidade de informação de cor é reduzida pela metade em relação à amostragem 4:2:2, o que pode levar a uma perda de qualidade perceptível, especialmente em imagens com detalhes finos ou texturas.

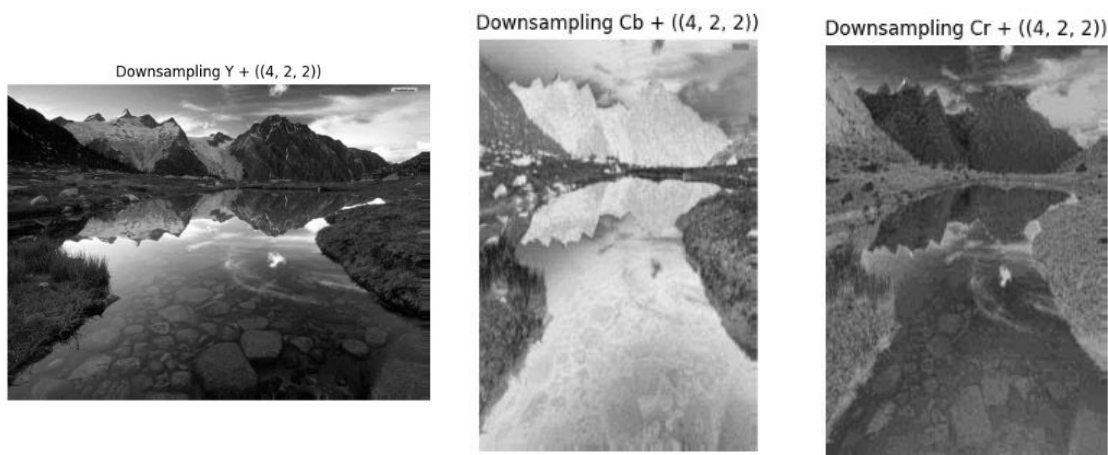


Figura 3. Downsampling 4,2,2

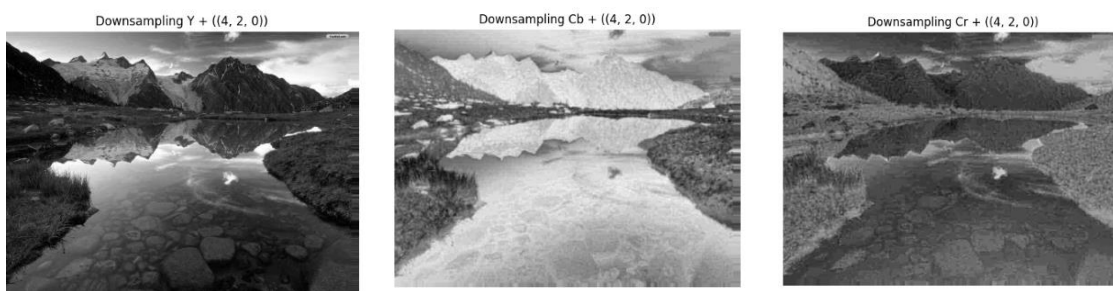


Figura 4. Downsampling 4,2,0

Apresentamos de seguida as tabelas das dimensões dos canais cb e cr com a amostragem 4:2:2 e 4:2:0, do upsampling. Isto permitiu validar as conclusões obtidas nas tabelas 1 e 2.

	4:2:2	
Ficheiros	Cb	Cr
<i>airport.bmb</i>	(704,608)	(704,608)
<i>geometric.bmp</i>	(1024,512)	(1024,512)
<i>nature.bmp</i>	(1216,816)	(1216,816)

Tabela 1 Tabela de Resultados do upsampling (4:2:2)

	4:2:0	
Ficheiros	Cb	Cr
<i>airport.bmb</i>	(352,608)	(352,608)
<i>geometric.bmp</i>	(512,512)	(512,512)
<i>nature.bmp</i>	(608,816)	(608,816)

Tabela 2 Tabela de Resultados do upsampling(4:2:0)

5. Transformada de Cosseno Discreta (DCT)

No processo de compressão JPEG, é aplicada a Transformada Discreta do Cosseno (DCT) para converter a imagem do domínio espacial para o domínio de frequência. Nesse novo domínio, a imagem é representada por uma matriz de coeficientes de frequência. A DCT é capaz de comprimir a imagem devido à sua propriedade de concentrar a energia dos dados nos coeficientes de frequência de baixa ordem, enquanto os coeficientes de alta ordem tendem a ser muito menores.

Essa transformação aumenta consideravelmente a eficiência da compressão, pois permite que algoritmos de compactação como Run Length Encoding(RLE) ou a Codificação de Huffman sejam aplicados com maior eficácia.

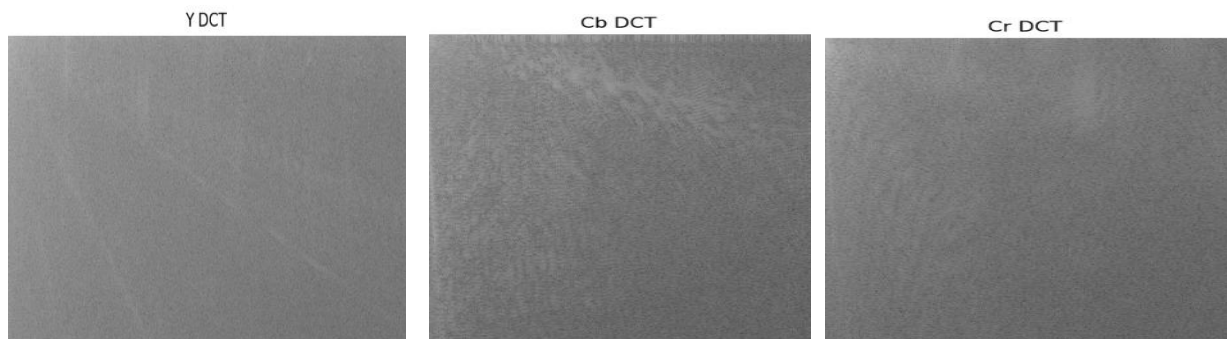


Figura 5. DCT Y, Cb e Cr

Observando a figura 5, pode-se verificar uma maior presença de branco no canto superior esquerdo por se tratar de valores mais elevados (relativos as baixas frequências espaciais).

6. DCT em blocos 8x8

Ao aplicar a Transformada Discreta do Cosseno (DCT) às imagens, observou-se que usar blocos de tamanho 8x8, figura 6, é mais vantajoso do que não usar blocos ou utilizar blocos maiores, como 64x64, figura 7. Isso deve-se ao fato de que a DCT em blocos 8x8 consegue equilibrar melhor a resolução espacial e a resolução de frequência. Consequentemente ela pode garantir que não haja transições abruptas em blocos adjacentes ou que a probabilidade de acontecerem seja mínima.

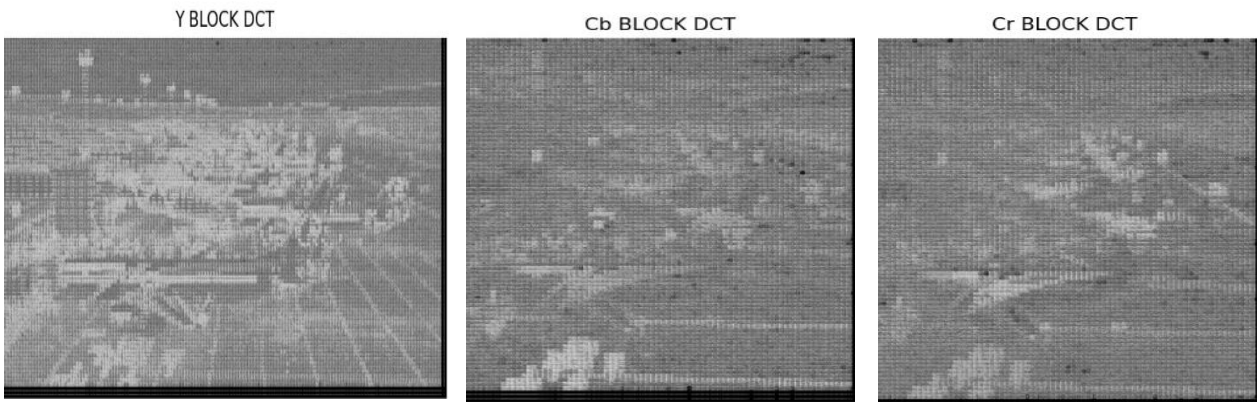


Figura 6. Blocos 8 x8 Y, Cb e Cr

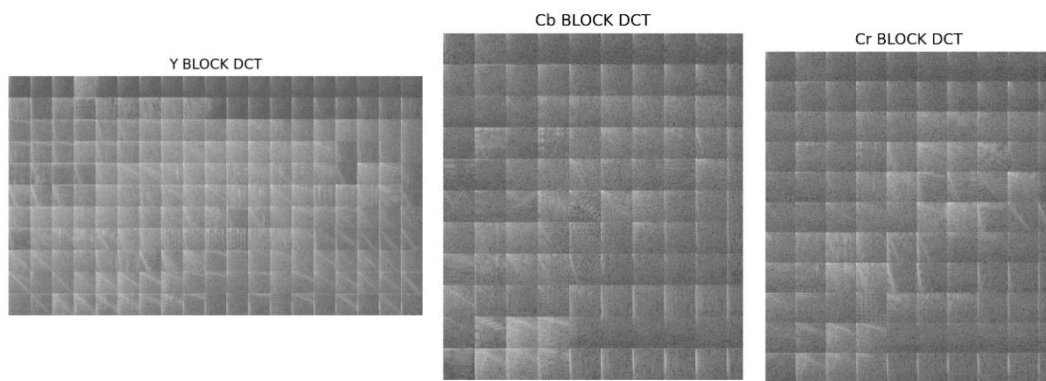


Figura 7. Blocos 64 x 64, Y, Cb e Cr

A imagem resultante com aplicação de blocos de 64x64, figura 7, é praticamente impercetível em comparação com a aplicação de blocos de 8x8, figura 6.

7. Quantização

De seguida fez-se a quantização. Para isso é aplicada à matriz de coeficientes de frequência obtida na etapa anterior, a tabela de quantização. Isso é feito para reduzir a quantidade de informação necessária para armazenar a imagem. Ao dividir os coeficientes de frequência pelos valores correspondentes na tabela de quantização, muitos desses coeficientes são arredondados para zero ou valores muito pequenos, resultando numa redução significativa na quantidade de informação a ser armazenada.

Esses fatores de escala na tabela de quantização são ajustados de forma que as frequências mais altas sejam mais fortemente reduzidas do que as frequências mais baixas. Isso ocorre porque o olho humano tem uma menor sensibilidade à variação exata de intensidade em componentes de alta frequência espacial.

Assim, a qualidade da imagem é controlada pela taxa de compressão, que é determinada pelas matrizes de quantização. Quanto menor for o valor do fator de qualidade, maiores serão os valores na matriz de quantização, resultando em menos bits necessários para representar os coeficientes, o que leva a uma maior taxa de compressão. No entanto, isso também significa que haverá uma redução na qualidade da imagem, com menos detalhes visíveis.

➤ Qualidade = 10

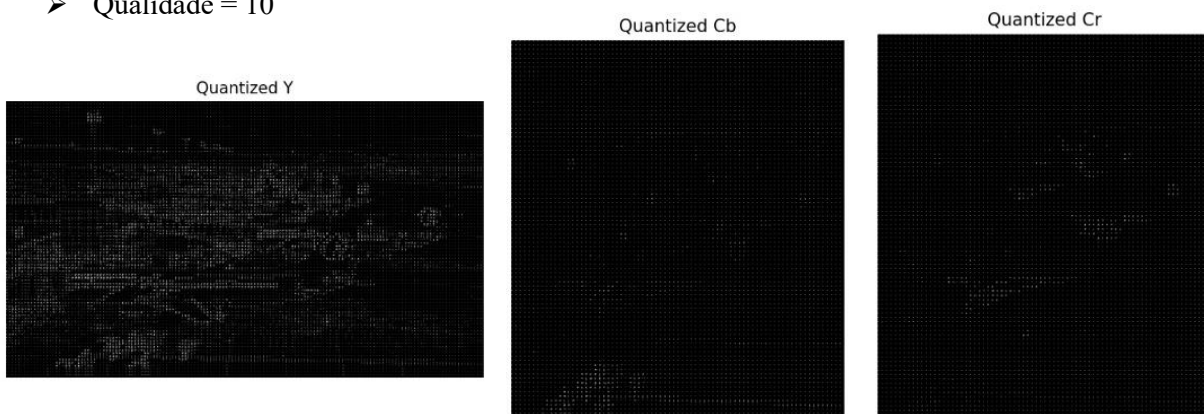


Figura 8. Quantização com Qualidade 10

- Qualidade = 25

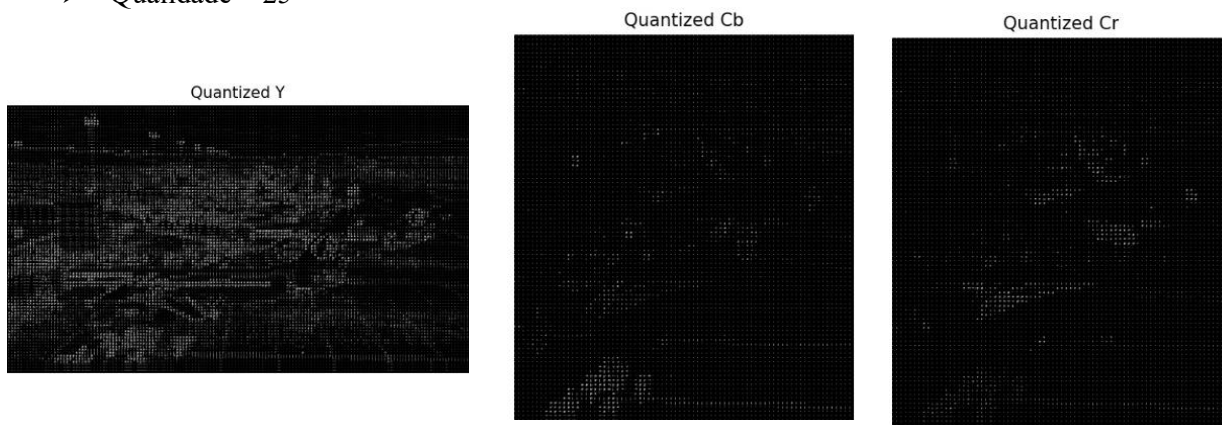


Figura 9. Quantização com Qualidade 25

- Qualidade = 50

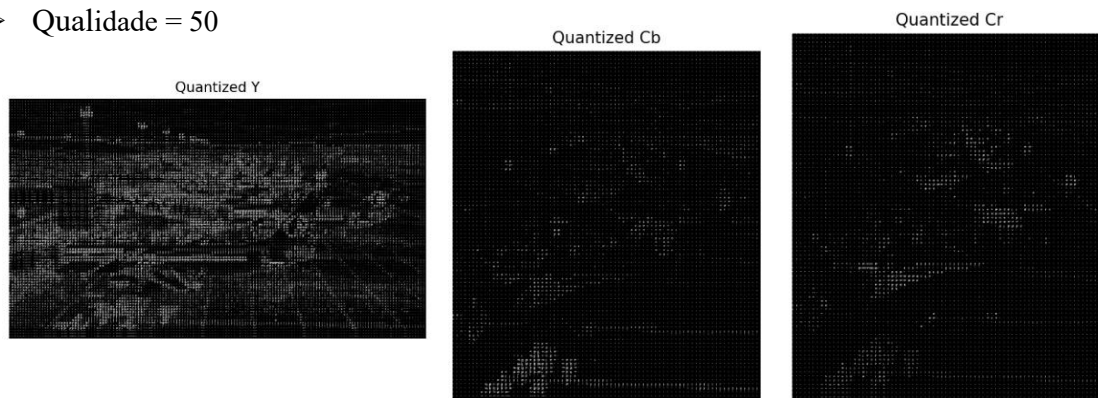


Figura 10. Quantização com Qualidade 50

- Qualidade = 75

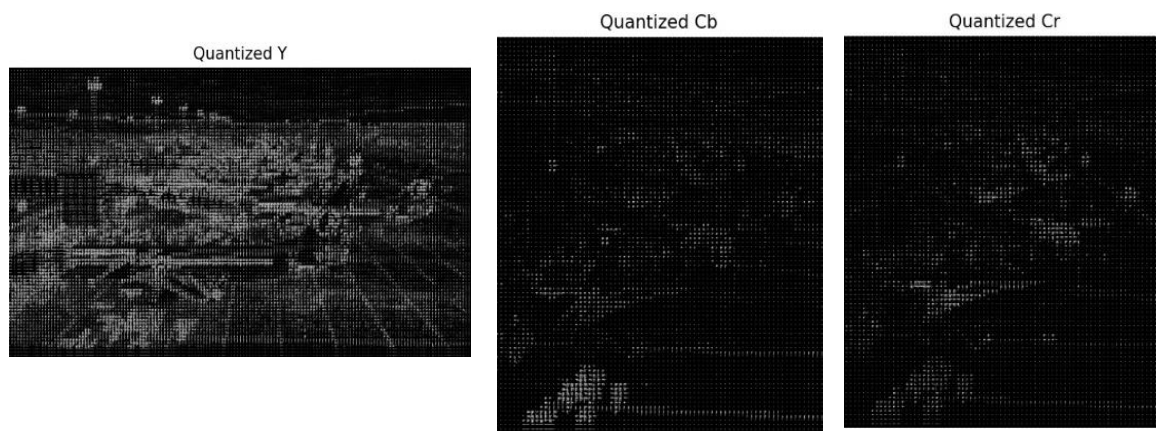


Figura 11. Quantização com Qualidade 75

- Qualidade = 100

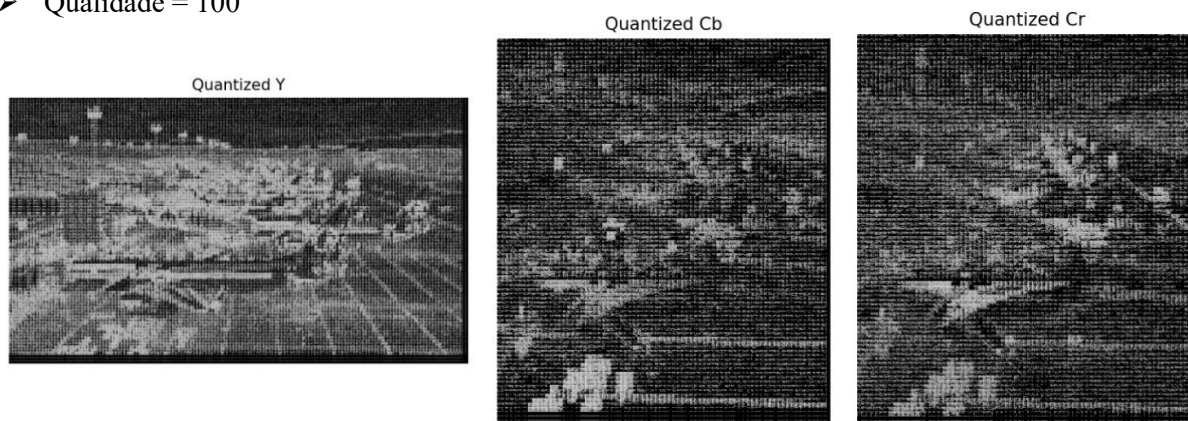


Figura 12. Quantização com Qualidade 100

Após a aplicação da quantização, usando vários fatores de qualidade 10,25,50,75,100 e pode-se observar que quanto maior a qualidade de imagem, mais perceptível ela fica. Esta característica está bem visível nas figuras 8,9,10, 11,12, sendo a figura 8 o resultado da aplicação de um fator de qualidade 10 e a figura 12 o resultado da aplicação de um fator de qualidade 100: na primeira verifica-se uma forte tonalidade escura, enquanto que na segunda já se conseguem distinguir os contornos dos componentes da imagem.

8. Codificação DPCM dos coeficientes DC

Após a quantização dos coeficientes DC, o algoritmo de compressão JPEG aplica a Differential pulse-code modulation (DPCM) nesses coeficientes. Esta etapa envolve calcular a diferença entre os valores dos coeficientes DC sucessivos de cada bloco e, em seguida, codificar essa diferença num número menor de bits. Esta técnica mostrou-se eficaz quando os valores dos coeficientes DC dos blocos são muito semelhantes, resultando em diferenças mínimas. Tal é comum em áreas de transição suave na imagem, como o céu, onde os tons são consistentes entre blocos adjacentes.

- Qualidade 100 usando a imagem airport

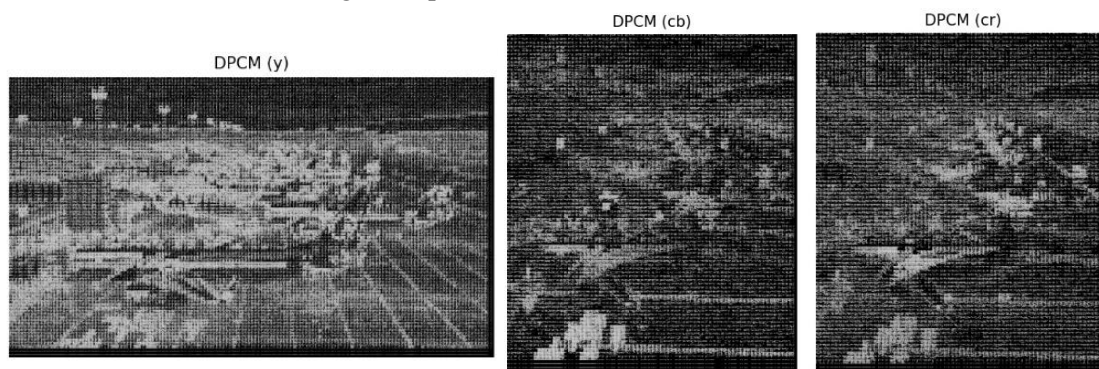


Figura 13. DPCM Qualidade 100

- Qualidade 100 usando a imagem nature

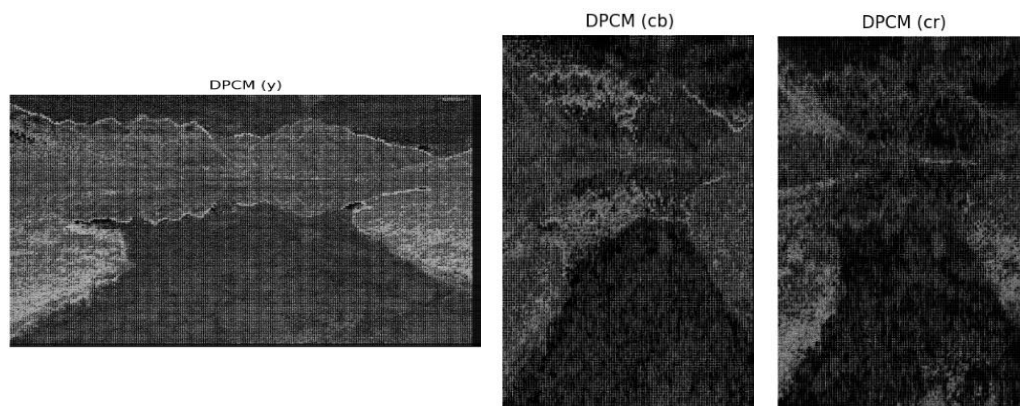


Figura 14. DPCM Qualidade 100

Os canais de luminância (Y), crominância (Cb) e (Cr) obtidos nesta etapa são muito semelhantes aos obtidos durante a quantização, exceto pelo valor dos coeficientes DC, que é modificado pela DPCM, como se pode verificar nas figuras 13 e 14.

9. Métricas de distorção (MSE, RMSE, SNR e PSNR) com os diferentes fatores de qualidade

Nas tabelas abaixo encontram-se os valores das várias métricas de distorção (MSE, RMSE, SNR e PSNR) para cada uma das imagens, e com os seus fatores de qualidade correspondentes. Isso permite comparar a imagem original com a sua reconstrução e avaliar o nível de distorção introduzido durante o processo de compressão.

	<i>Fator de Qualidade</i>				
<i>M.D</i>	10	25	50	75	100
MSE	565.5731	287.9622	164.3970	87.29509	11.2375
RMSE	23.7818	16.9695	12.8217	9.34318	3.3522
SNR	22.1260	25.0576	27.49196	30.24101	39.1442
PSNR	20.6059	23.5375	25.97186	28.7209	37.6241

Tabela 3. Tabela de métricas de Distorção da imagem Airport

	<i>Fator de Qualidade</i>				
<i>M.D</i>	10	25	50	75	100
MSE	221.5934	105.6505	66.2076	40.43424	19.07515
RMSE	14.8860	10.27864	8.136805	6.358793	4.367511
SNR	26.67488	29.8917	31.92137	34.0629	37.3258
PSNR	24.675235	27.8921	29.92173	32.0633	35.3261

Tabela 4. Tabela de métricas de Distorção da imagem Geometric

	<i>Fator de Qualidade</i>				
<i>M.D</i>	10	25	50	75	100
MSE	561.6345	379.19161	35.2533	27.26393	1.26374
RMSE	23.69883	19.472843	5.937449	5.221487	1.124164
SNR	19.478309	21.184260	31.50085	32.61696	45.9563
PSNR	20.636266	22.3422	32.65880	33.77492	47.1142

Tabela 5. Tabela de métricas de Distorção da imagem *Nature*

Fazendo uma análise as imagens por fator de qualidade, percebe-se que quanto menor o fator de qualidade, maior a diferença entre a imagem original e a reconstruída, conforme indicado pelos valores de MSE e RMSE. E isto como exemplo pode se verificar pelo MSE na qualidade 10, obtidos nas imagens *Nature* e *Airport* que ultrapassam os 560 devido sua riqueza a nível de cores, tabelas 5 e 3, enquanto a *Geometric*, apresenta quase que a metade disso, possivelmente pelo facto de ter longas cadeias de pixels adjacentes com a mesma cor tabela 4.

Por outro lado, quanto maior o fator de qualidade, menor é a distorção na imagem comprimida, conforme evidenciado pelos valores de SNR e PSNR. Isso foi possível observar nos resultados obtidos na imagem *Nature* com qualidade 10, o SNR= 19.478 e PSNR = 20.6362, para um fator de qualidade 100 a mesma imagem tem o SNR=45.956 e PSNR=47.114, tabela 5.

Pode-se observar que em relação a imagem original as imagens *Nature* e *Airport* apresentam diferenças notáveis principalmente na região do céu, e pela imagem *Geometric* também podemos verificar que há diferenças, mas não é tao significativa, observada na transição dos padrões geométricos. Essas diferenças entre as imagens originais e as reconstruídas confirmam a relação entre a degradação da informação e o fator de qualidade, resultante do processo de quantização.



Figura 15. Imagem Airport reconstruída com $Q=10$ e $Q=100$



Figura 16. Imagem Geometric reconstruída com $Q=10$ e $Q=100$



Figura 17. Imagem Nature reconstruída com $Q=10$ e $Q=100$

Em suma, os resultados demonstram que o codec JPEG é eficaz na compressão de imagens, mas a qualidade e a taxa de compressão são influenciadas pelo fator de qualidade escolhido, figuras 15, 16 e 17.

➤ Imagens de Diferenças com $Qf=10$



Figura 18. Imagem Original e de diferenças - Nature



Figura 19. Imagem Original e de diferenças - Airport

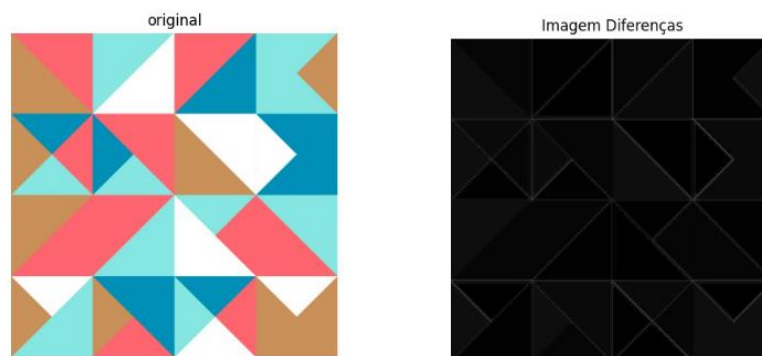


Figura 20. Imagem Original e de diferenças - Geometric

Ficheiros	Fator de Qualidade				
	10	25	50	75	100
<i>airport.bmb</i>	7.429	5.108	3.769	2.701	0.234
<i>geometric.bmp</i>	2.778	1.036	0.836	0.349	0.037
<i>nature.bmp</i>	9.058	7.757	2.150	1.985	0.206

Tabela 6. Tabela de valor medio da diferença absoluta entre o canal y da imagem reconstruída com a original

Ficheiros	Fator de Qualidade				
	10	25	50	75	100
<i>airport.bmb</i>	187.081	97.878	76.202	42.286	16.391
<i>geometric.bmp</i>	79.94	60.504	39.995	27.734	9.269
<i>nature.bmp</i>	134.450	115.928	30.994	28.526	4.695

Tabela 7. Tabela de valor Máximo da diferença absoluta entre o canal y da imagem reconstruída com a original

Foi calculada a diferença entre a imagem final comprimida e a imagem original para obter a imagem de erro, conforme apresentado nas figuras 18,19 e 20. A partir das tabelas 6 e 7, é possível estabelecer uma relação clara entre o fator de qualidade e as diferenças entre as imagens. Como esperado, um fator de qualidade mais alto resulta numa menor diferença entre as imagens, enquanto um fator de qualidade mais baixo resulta numa diferença maior.

Em retrospectiva, podemos confirmar as conclusões tiradas anteriormente sobre a relação entre o fator de qualidade e a compressão, bem como a qualidade das imagens obtidas com o codec JPEG para diferentes fatores de qualidade.

10. Conclusão

Este trabalho prático permitiu-nos analisar aprofundadamente o codec JPEG e sua capacidade de compressão em diferentes tipos de imagens. Observamos que o efeito da compressão pode variar de acordo com as características específicas das imagens. Com base nos resultados obtidos, percebe-se que o JPEG é um algoritmo de compressão bastante eficiente tendo uma grande capacidade de compressão e qualidade perceptual, quando são tidas em conta as características específicas da imagem: em imagens com transições abruptas entre cores, é possível observar um certo ruído de quantização nas áreas afetadas, enquanto imagens com transições suaves podem ser comprimidas com maior eficiência sem comprometer muito a qualidade visual.