

1 2 9 0



UNIVERSIDADE
COIMBRA

Compressão de imagem
Multimédia

Licenciatura em Engenharia Informática
2022/2023

17 de março de 2023

Autores

João Moreira

 joaomoreira@student.dei.uc.pt

 2020230563

Rodrigo Figueiredo

 rfigueiredo@student.dei.uc.pt

 2020236687

Tomás Pinto

 tomaspinto@student.dei.uc.pt

 2020224069

Índice

1	Introdução	3
2	Compressão de Imagem Com um Editor Auxiliar	4
2.1	Resultados	4
2.2	Analise de Resultados	4
3	Modelo YCbCr vs Modelo RGB	5
3.1	Resultados	5
4	Sub-Amostragem	6
4.1	Resultados	7
4.2	Análise de resultados	7
5	Transformada de Cosseno Discreto	8
5.1	Aplicação da DCT no Canal Completo	8
5.2	Aplicação da DCT em Blocos 8x8	9
5.3	Aplicação da DCT em Blocos 64x64	10
6	Quantização	11
7	Modelação de Código de Pulso Diferencial	13
8	Imagens descodificadas e erros	15
9	Métricas de distorção	16
10	Conclusão	18

1. Introdução

Compressão de dados de imagem é a técnica de reduzir o tamanho original de uma imagem de forma a manter a sua qualidade.

Ao contrário do que foi desenvolvido em Teoria da Informação (compressão de dados sem perda), neste trabalho pretende-se explorar a compressão com perda de informação. Neste tipo de algoritmo, tentamos reduzir o tamanho original da imagem com a perda de parte da informação inicial.

Sendo assim, o objetivo principal é explorar a informação pouco relevante visualmente para o ser Humano, de forma que ao eliminá-la seja imperceptível ao olho humano. O codec que iremos usar é o JPEG.

2. Compressão de Imagem Com um Editor Auxiliar

Começámos por comprimir as 3 imagens fornecidas em formato bmp para o formato jpeg nas qualidades alta (75%), média (50%) e baixa (25%) utilizando o GIMP. Os resultados obtidos são apresentados de seguida.

2.1. Resultados

Compress	Barn_Mountains	Logo	Peppers
Original	348	411	576
High	33	9.4	30.5
Medium	21.5	7.6	20.3
Low	13.5	6.2	13.3

Tabela 1: Tamanho final das imagens em Kb.

Compress	Barn_Mountains	Logo	Peppers
High	10:1	44:1	19:1
Medium	16:1	54:1	28:1
Low	26:1	66:1	43:1

Tabela 2: Racio das compressões por imagem.

2.2. Analise de Resultados

Como era esperado, os resultados mostram uma diminuição de tamanho dos ficheiros quando as imagens originais são comprimidas. Podemos verificar que quanto menor é a qualidade menor é o tamanho do ficheiro também.

Algo interessante de observar é que conseguimos comprimir mais de dez vezes o ficheiro original e manter uma qualidade alta.

Ao olhar para as imagens de **qualidade alta** é notório alguma perda de qualidade. As principais zonas onde se nota maior perda de qualidade é nas transições abruptas de cor. Por exemplo, na imagem *logo.bmp* observa-se grandes perdas junto às transições de cores, não se notando perdas nas partes de cores constantes.

Este comportamento é característico do JPEG, visto que ele explora as limitações do olho Humano e acaba por perder alguma sensibilidade nas zonas de transições abruptas.

Sendo assim, concluímos que o JPEG tem uma maior compressão para imagens com transições abruptas de cor (o que gera mais ruído/distorção) e trás mais qualidade para imagens com transições suaves de cor.

3. Modelo YCbCr vs Modelo RGB

Foi feita a conversão das imagens do modelo RGB inicial para o modelo YCbCr.

O modelo RGB possui redundância da luminância nos 3 canais. Esta conversão é importante pois este modelo permite-nos separar as componentes da luminância (canal Y) e da crominância (canais Cb e Cr), que se encontram fundidas no modelo RGB.

Como a visão humana é mais sensível à luminância, o JPEG pode tirar proveito disso para comprimir mais nas componentes Cb e Cr que irão influenciar muito menos a qualidade da imagem final a olho nu comparativamente à componente Y.

3.1. Resultados



Figura 1: Canal R



Figura 2: Canal G



Figura 3: Canal B

Podemos reparar que nas imagens acima, apesar dos três canais R, G e B conterem luminância, o nível de detalhe não é igual em todos. Isto é devido ao facto do olho humano ser mais sensível às tonalidades verdes, de seguida vermelhas e por último azuis, coincidente com a ordem decrescente de detalhe das três imagens.



Figura 4: Canal Y



Figura 5: Canal CB

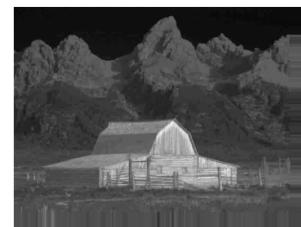


Figura 6: Canal CR

Como é possível ver pelas imagens, o canal Y (canal da luminância) apresenta mais detalhe do que os canais Cb e Cr (canais da crominância). Isto deve-se não à perda de qualidade, mas sim ao facto de o olho humano ter maior sensibilidade à luminância do que a crominância.

Sendo assim, iremos ver ao longo do trabalho como o JPEG usa os canais Cb e Cr a seu favor.

Nota: Para melhor visualização as imagens relativas aos canais Y, Cb e Cr são aplicadas um colormap de tons de cinza.

4. Sub-Amostragem

Na fase da sub-amostragem, *downsampling*, é reduzida a quantidade de dados nos canais relativos à cor (Cb e Cr) com o objetivo de melhorar a compressão, o que resulta numa resolução espacial menor nas imagens, pelo que não há alterações no canal Y.

Como explicado anteriormente, o olho humano é menos sensível à crominância e a redução de informação pode ser feita sem comprometer significativamente a qualidade da imagem final.

Sendo assim, de forma a diminuir a informação podemos remover parte dela:

- Sub-amostragem 4 : 2 : 2 a imagem ficará reduzida horizontalmente, com metade das colunas.
- Sub-amostragem 4 : 2 : 0 a imagem ficará reduzida verticalmente e horizontalmente, com metade das colunas e linhas.

Nesta redução foi aplicada uma interpolação cúbica (opção do grupo de trabalho).

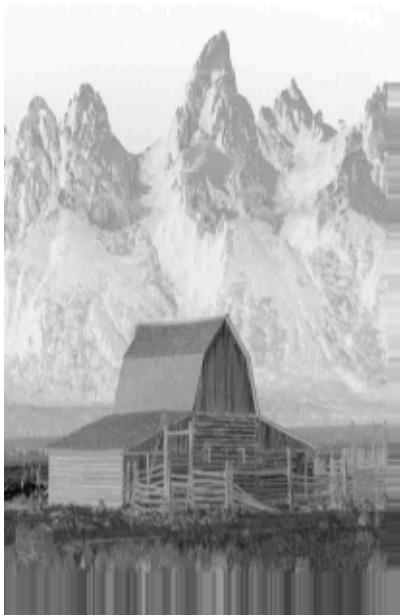


Figura 7: Sub-amostragem canal CB 4:2:2



Figura 8: Sub-amostragem canal CR 4:2:2



Figura 9: Sub-amostragem canal CB 4:2:0

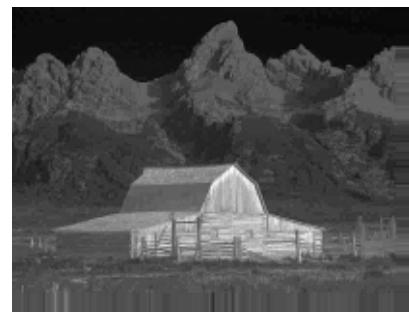


Figura 10: Sub-amostragem canal CR 4:2:0

4.1. Resultados

Imagens	Tamanho
Cb	122
Cr	112
Cb_422	57
Cr_422	54
Cb_420	34
Cr_420	32

Tabela 3: Tamanho final das sub-amostragens em Kb.

Imagens	Taxa de Compressão
Cb_422	2 : 1
Cr_422	2 : 1
Cb_420	4 : 1
Cr_420	4 : 1

Tabela 4: Taxa de compressão das sub-amostragens.

4.2. Análise de resultados

Como era esperado, para fatores 4 : 2 : 2 o tamanho das imagens diminui para metade e fatores 4 : 2 : 0 reduziu para um quarto do tamanho original. É possível observar também que os detalhes mantém-se apesar de a informação ter diminuído.

Apesar de o fator 4 : 2 : 0 apresentar uma maior compressão face o fator 4 : 2 : 2, em contra partida é o que apresenta maior destrutividade.

Para as outras imagens, vai-se observar o mesmo comportamento.

5. Transformada de Cosseno Discreto

A DCT permite-nos converter um sinal num espectro de frequências. Desta forma, em imagens podemos converter o valor de cada píxel para o domínio da frequência, o que nos permite ter uma análise das frequências que estão a ser usadas na imagem e estabelecer uma relação com aquelas que ao olho Humano não é tão sensível.

Visto que a DFT resulta em números imaginários e tem um padrão repetitivo, seria necessário o dobro da memória face à DCT. Sendo assim, é mais vantajoso usar a DCT.

5.1. Aplicação da DCT no Canal Completo

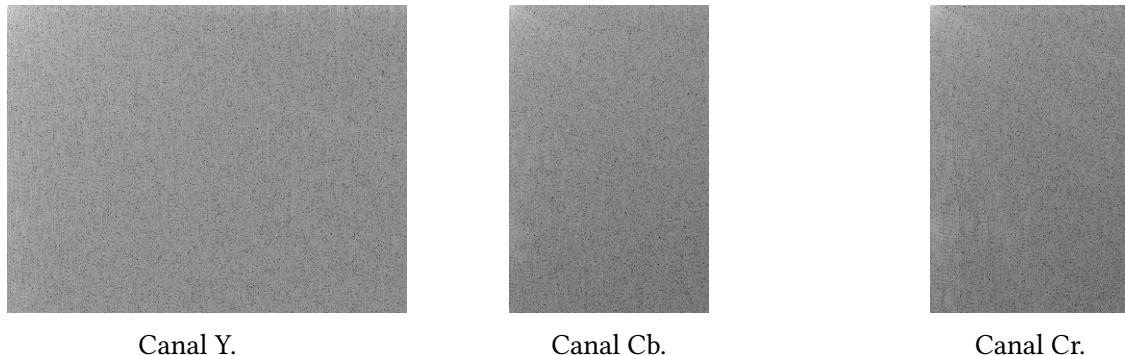


Figura 11: DCT da imagem Barn_Mountains.

Como vemos na *Figura 11*, observamos que os coeficientes DC e AC mais altos concentram-se no canto superior esquerdo, ou seja, aquele ponto cintilante é representativo de alta energia para as baixas frequências e os tons mais escuros representam baixa energia nas altas frequências. Logo, significa que estamos perante uma imagem de transições suaves, o que se comprova, pois este espectro pertence à *Figura Barn_Mountains*.

Visto que na gama de altas frequências da *Figura Barn_Mountains* apresenta baixa energia, ela irá comprimir menos do que a *Figura Logo*, que em contra partida apresenta alta energia ao longo do espectro de frequências. Pela mesma razão, a *Figura Barn_Mountains* irá apresentar menos ruído do que a *Figura Logo*.

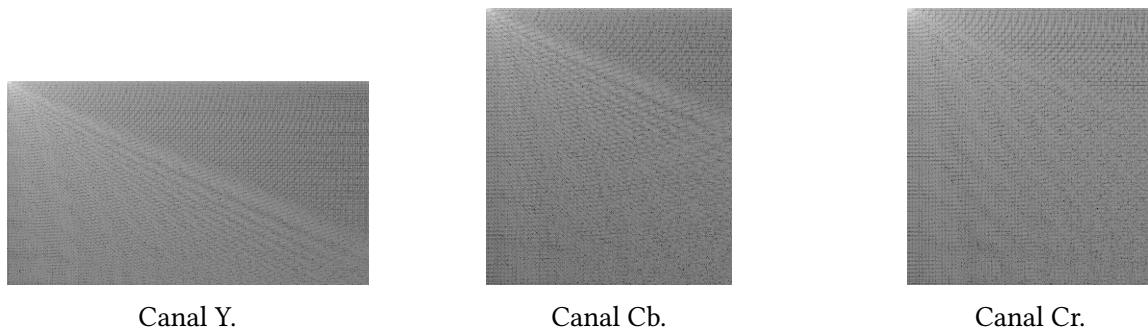


Figura 12: DCT da imagem Logo.

Ao analisar a *Figura 12* vemos que existe uma variação de energia ao longo das altas e baixas frequências, o que significa que esta imagem apresenta simultaneamente transições suaves e abruptas. Neste exemplo, trata-se da *Figura logo*.

Espectros de altas frequências com baixa energia têm tendência a comprimir e destruir mais, em relação a espectros de baixas frequências.

5.2. Aplicação da DCT em Blocos 8x8

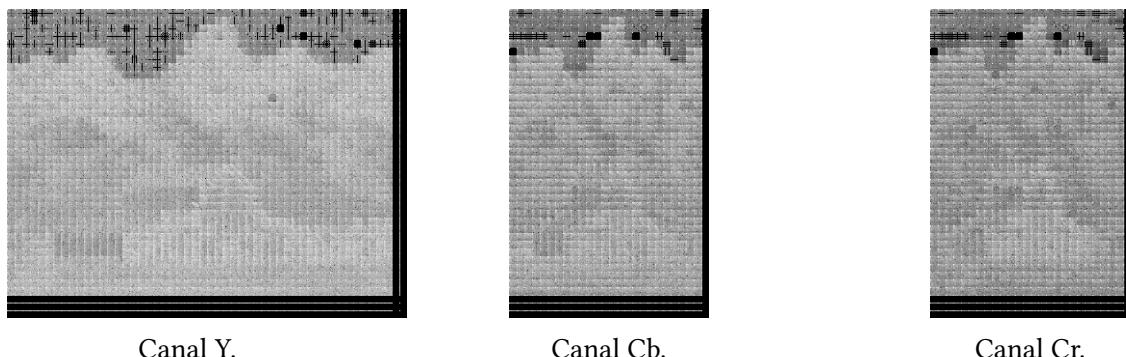


Figura 13: DCT da imagem Barn_Mountains em blocos 8x8.

Como é visível nas imagens acima, quase que é possível reconhecer a imagem inicial. A primeira observação a notar é que o canal Y tem mais detalhes. Um motivo plausível para isso acontecer é o facto de os canais Cb e Cr terem sofrido destruição na sub-amostragem e o canal Y permanecer "*virgem*".

Isto acontece, porque ao limitar o nosso sinal a blocos de 8x8 a probabilidade de ocorrer transições de cores abruptas diminui, no que resulta um sinal de transições de cores suaves.

A aplicação da DCT em blocos 8x8 tem um maior potencial de compressão em relação à imagem inteira pois as altas frequências têm maior probabilidade de terem baixa energia, ou seja, existirá um maior número de coeficientes AC nulos ou próximos de zero.

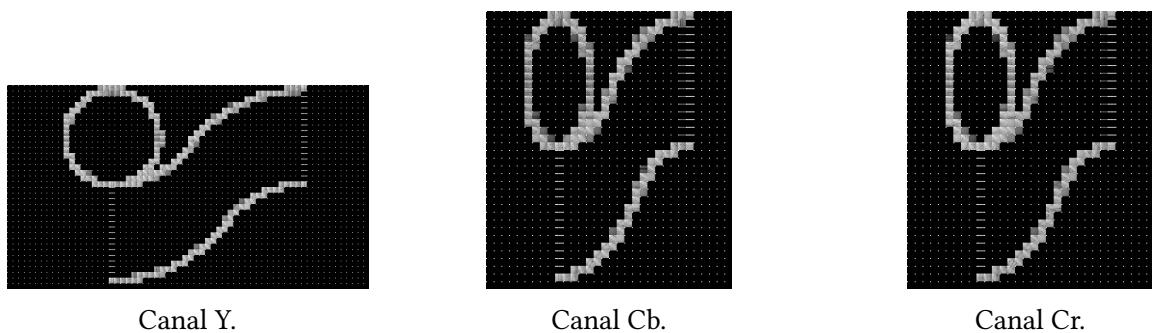
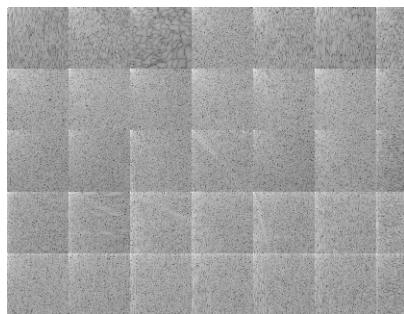


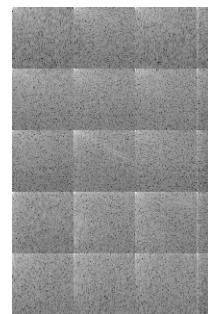
Figura 14: DCT da imagem Barn_Mountains em blocos 8x8.

Nas figuras que antecedem, facilmente verificamos o que foi descrito anteriormente: como existe poucas zonas de transições, a maior parte dos coeficientes AC estão a zero e têm um maior potencial de compressão.

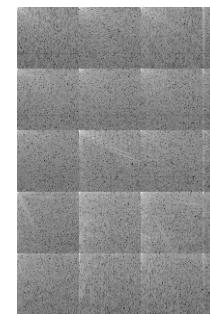
5.3. Aplicação da DCT em Blocos 64x64



DCT do canal Y em blocos
64x64



DCT do canal Cb em blocos
64x64



DCT do canal Cr em blocos
64x64

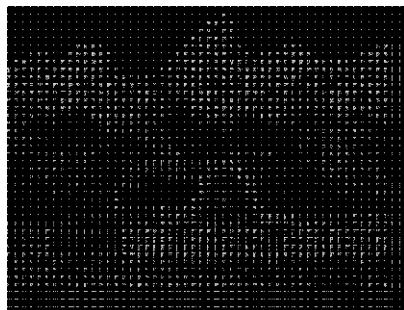
Figura 15: DCT da imagem Barn_Mountains em blocos 64x64.

Como podemos observar aplicar a DCT em blocos de 64x64 deixa o canal ainda muito homogéneo. Com esta estratégia iremos ter melhor compressão do que com o canal inteiro, mas se aplicado com blocos de 8x8 teremos resultados muito melhores.

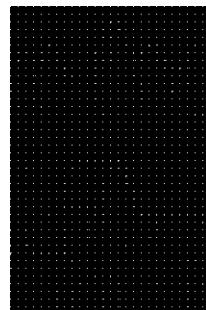
6. Quantização

O olho humano não tem tanta sensibilidade para distinguir componentes de altas frequências (estas frequências correspondem às transições abruptas).

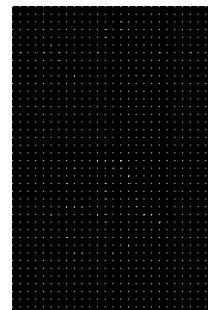
No processo de quantização divide-se as frequências por um valor de quantização já pré-determinado, aliado a um fator de qualidade. Isso permite-nos remover facilmente as altas frequências da imagem (já calculadas pela DCT), pois são menos importantes para o olho humano.



Canal Y.



Canal Cb.

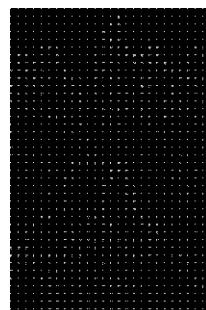


Canal Cr.

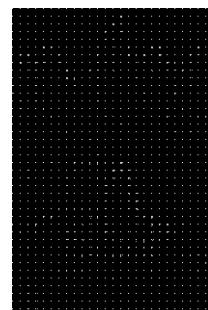
Figura 16: Quantização da imagem Barn_Mountains qualidade 10.



Canal Y.

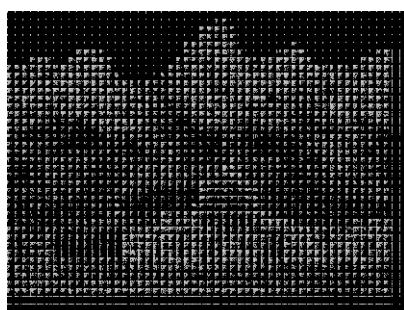


Canal Cb.

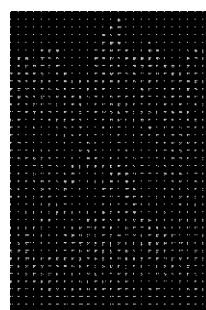


Canal Cr.

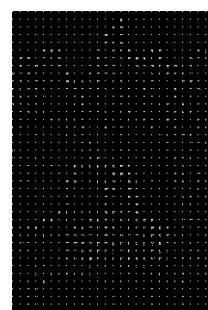
Figura 17: Quantização da imagem Barn_Mountains qualidade 25.



Canal Y.



Canal Cb.



Canal Cr.

Figura 18: Quantização da imagem Barn_Mountains qualidade 50.

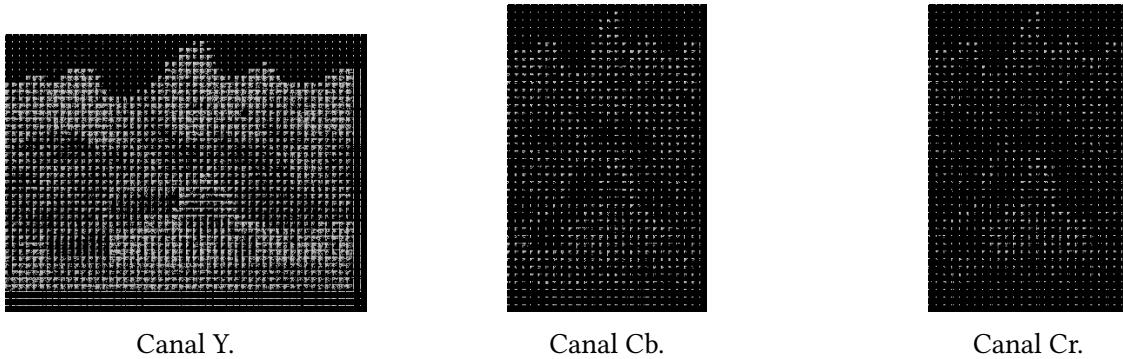


Figura 19: Quantização da imagem Barn_Mountains qualidade 75.

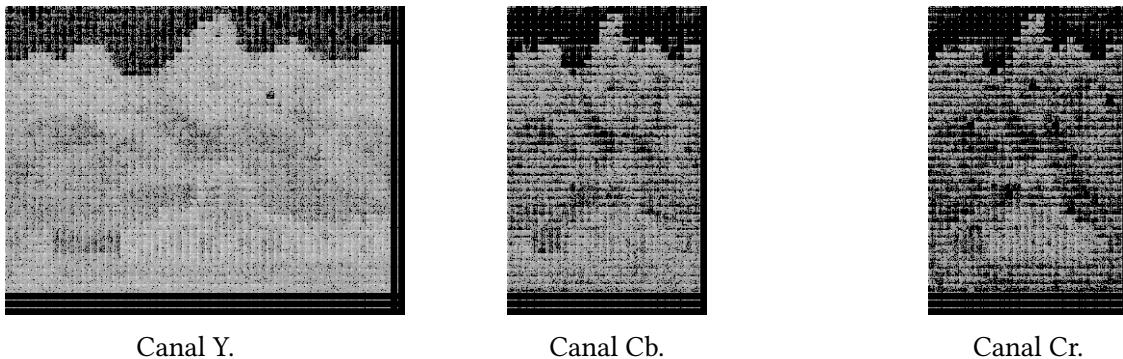


Figura 20: Quantização da imagem Barn_Mountains qualidade 100.

Quando aplicada a quantização, verificamos que a imagem fica mais escura, pois os coeficientes AC das altas frequências foram anulados, pois ao aplicar a matriz de quantização essas frequências serão eliminadas.

À medida que a qualidade diminui, o número de coeficientes AC a zero aumenta, ou seja, um maior número de frequências altas serão eliminadas. Na qualidade 10, já só quase resta os coeficientes DC do espectro (alta energia nas baixas frequências).

Sendo assim, podemos concluir que a compressão aumenta à medida que a qualidade diminui.

Vale alertar para o facto de que quando a qualidade tem um fator de 100, aparenta haver um aumento dos coeficientes AC. Isto acontece devido a questões de arredondamentos na codificação do modelo de cor.

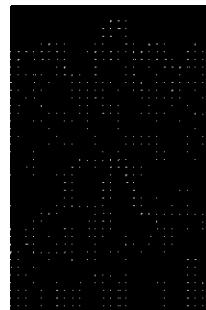
7. Modelação de Código de Pulso Diferencial

De modo a ter uma compressão mais efetiva, aplicamos a DPCM, cujo é um método de codificação preditiva. Este modelo, prevê o valor atual com base no valor anterior.

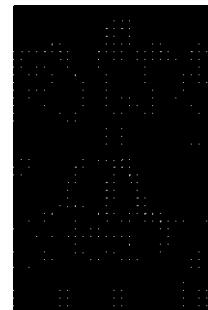
Desta forma, os valores de cada pixel diminuem e podemos reduzir a quantidade de bits necessária para codificar a informação.



Canal Y.



Canal Cb.

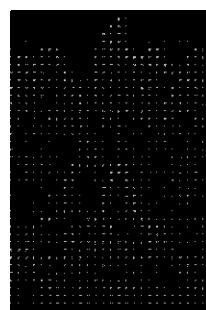


Canal Cr.

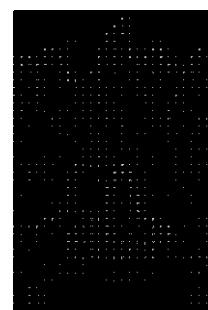
Figura 21: DPCM da imagem Barn_Mountains qualidade 10.



Canal Y.

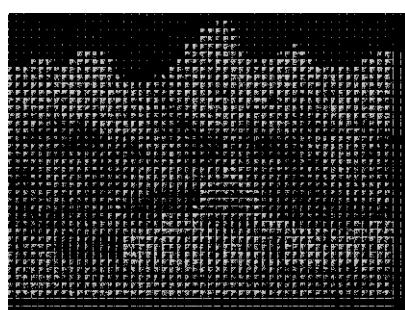


Canal Cb.

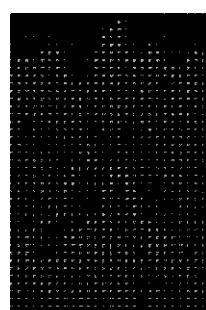


Canal Cr.

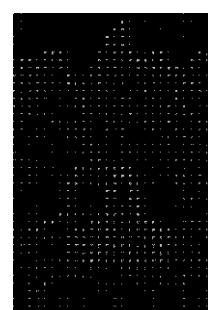
Figura 22: DPCM da imagem Barn_Mountains qualidade 25.



Canal Y.



Canal Cb.



Canal Cr.

Figura 23: DPCM da imagem Barn_Mountains qualidade 50.

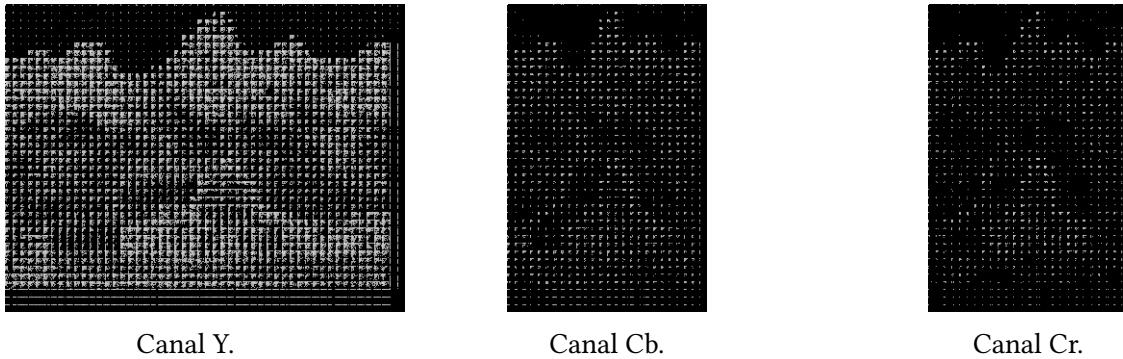


Figura 24: DPCM da imagem Barn_Mountains qualidade 75.

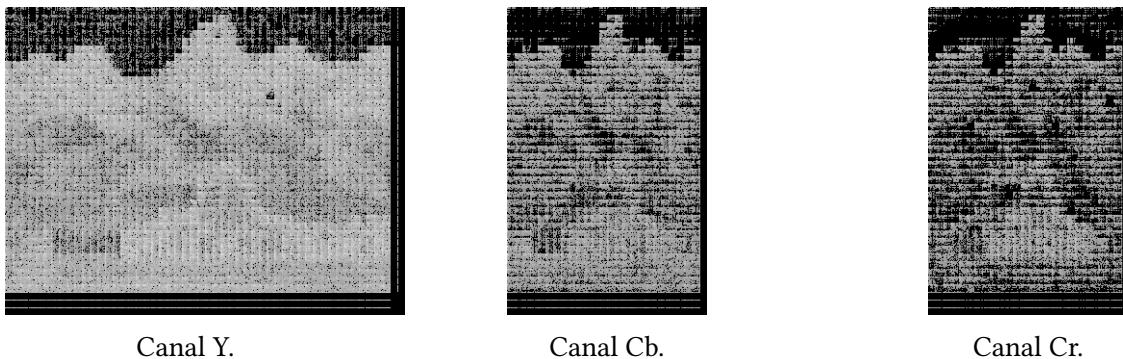


Figura 25: DPCM da imagem Barn_Mountains qualidade 100.

Imagens	Tamanho da DCPM
Y_DCT	244
Cb_DCT	58
Cr_DCT	43

Tabela 5: Tamanho da DCPM em Kb.

Ao olhar para os resultados da DPCM podemos notar as imagens mais escuras, com menos pontos brancos do que as imagens resultantes da quantização. Isto acontece pelo facto de que na fase de DPCM os coeficientes DC de cada bloco são codificados como a diferença face ao DC do bloco anterior, à excessão do primeiro bloco de todos que se mantém igual. Os coeficientes DC vão ter valores mais baixos e a diferença relativamente aos coeficientes AC diminui, podendo estes ser codificados com menos bits e assim os pontos brancos diminuem.

Em imagens suaves com poucas transições de cor e luz a DPCM obtém melhores resultados uma vez que os coeficientes DC dos vários blocos apresentam valores mais semelhantes.

8. Imagens descodificadas e erros

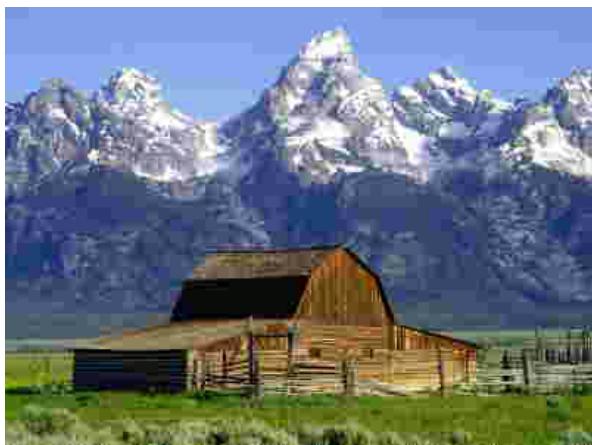


Figura 26: Imagem descodificada com fator de qualidade 10

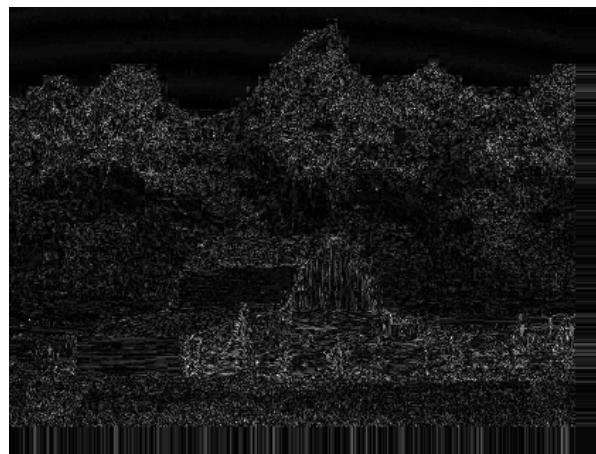


Figura 27: Y error com fator de qualidade 10



Figura 28: Imagem descodificada com fator de qualidade 25

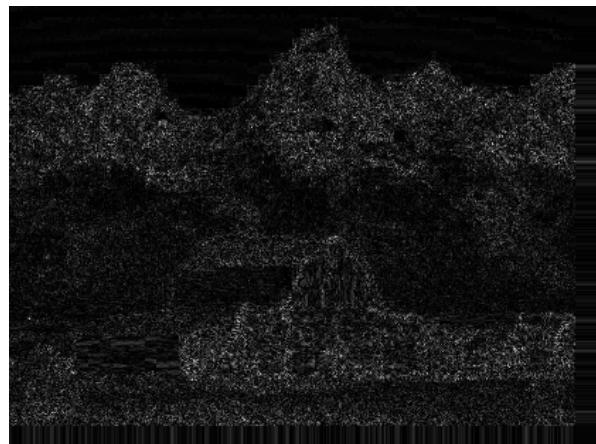


Figura 29: Y error com fator de qualidade 25



Figura 30: Imagem descodificada com fator de qualidade 50

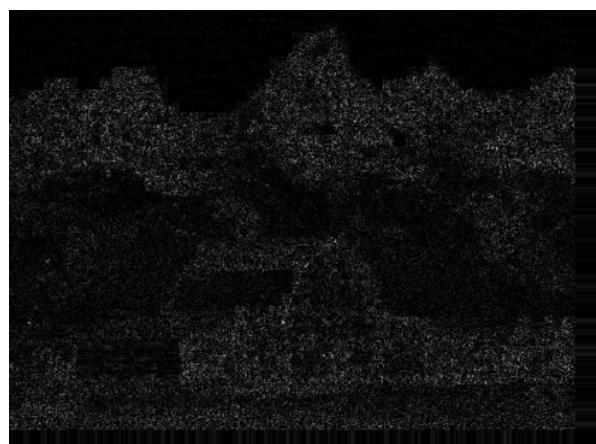


Figura 31: Y error com fator de qualidade 50



Figura 32: Imagem descodificada com fator de qualidade 75

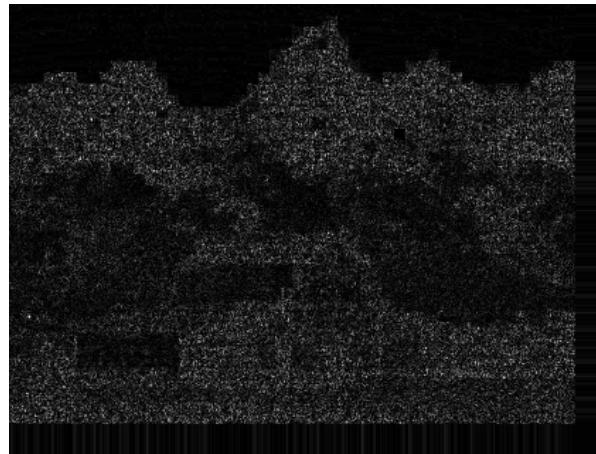


Figura 33: Y error com fator de qualidade 75



Figura 34: Imagem descodificada com fator de qualidade 100

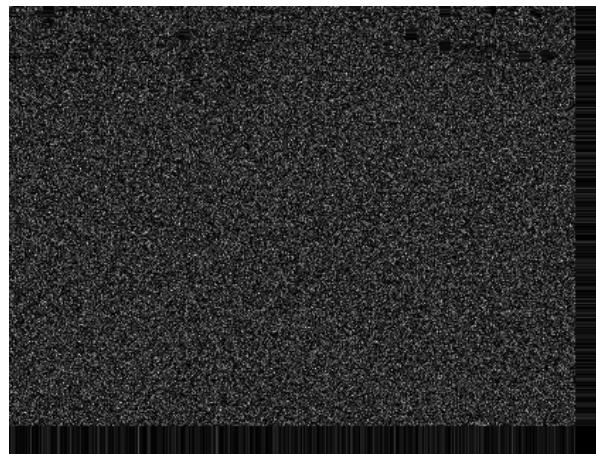


Figura 35: Y error com fator de qualidade 100

Nas figuras acima temos representadas as diferenças entre o canal Y de cada uma das imagens originais e da imagem descodificada respetiva (erro), para cada um dos fatores de qualidade testados. Quanto menor o fator de qualidade mais marcadas¹ estão as imagens (estas marcas são maiores e mais visíveis quanto maior for a compressão ocorrida nessa área), o que era esperado uma vez que quanto menor é o fator, maior é a compressão e perda de informação, resultando num erro maior.

9. Métricas de distorção

Calculamos, para cada imagem com cada fator de qualidade, métricas de distorção que permitem analisar e comparar as diferenças entre a imagem original e as imagens reconstruídas. Fizemos o cálculo do MSE (Mean squared error) que mede a diferença média ao quadrado entre os pixels das duas imagens e o RMSE (Root Mean Square Error) que é a raiz quadrada do anterior. Quanto maiores são estes valores maior a diferença entre as duas imagens, ou seja, maiores as perdas de informação.

Obtemos também o SNR (Signal-to-noise ratio) que mede a razão entre a amplitude do sinal

¹Subentendemos *marcadas* como a semelhança que a imagem do erro tem com a imagem original. Uma imagem mais *marcada* significa que se assemelha mais à imagem original.

e a amplitude do ruído na imagem, e o PSNR (Peak signal-to-noise ratio) que é a razão entre a potência máxima possível de um sinal e a potência do ruído presente. Maiores valores destas duas métricas indicam melhor qualidade na imagem (menos ruído), Por fim calculamos a média do Y error representados nas imagens acima.

barn_mountains	10	25	50	75	100
MSE	709.672	399.375	261.932	152.652	8.942
RMSE	26.639	19.984	16.184	12.355	2.990
SNR	18.675	21.172	23.004	25.349	37.671
PSNR	19.620	22.116	23.948	26.293	38.616
Mean	9.452	6.778	5.268	3.844	0.211

Tabela 6: Métricas de distorção da imagem barn-mountains.bmp

logo	10	25	50	75	100
MSE	158.921	66.608	43.168	22.943	4.520
RMSE	12.606	8.161	6.570	4.789	2.126
SNR	29.320	33.096	34.980	37.725	44.780
PSNR	26.118	29.895	31.779	34.524	41.578
Mean	5.211	1.520	1.275	0.445	0.038

Tabela 7: Métricas de distorção da imagem logo.bmp

peppers	10	25	50	75	100
MSE	287.667	129.383	77.917	50.335	4.981
RMSE	16.960	11.374	8.827	7.094	2.231
SNR	20.337	23.807	26.009	27.907	37.952
PSNR	23.541	27.012	29.214	31.112	41.157
Mean	4.487	2.680	1.965	1.483	0.230

Tabela 8: Métricas de distorção da imagem peppers.bmp

Da analise dos resultados obtivemos, para todas as imagens, valores esperados consoante a descrição de cada métrica. Os valores de MSE e RMSE foram decrescendo com o aumento do fator de qualidade, uma vez que as diferenças das imagens descodificadas com as originais vão sendo menores. Já os valores de SNR e PSNR foram aumentando com o aumento do fator de qualidade devido ao facto do ruído progressivamente decrescer. Por fim os valores da média das diferenças entre os canais Y das imagens originais e descodificadas foi decrescendo com o aumento do fator de qualidade, como era também esperado.

10. Conclusão

No final deste trabalho conseguimos adquirir o conhecimento impírico sobre o funcionamento do JPEG:

- Permite fazer compressões nas imagens de alta resolução;
- Imagens de cores sólidas e transições suaves irão ter alta compressão;
- Imagens com transições de cores abruptas podem resultar em perdas de qualidade e detalhe mais acentuadas;