



FCTUC FACULDADE DE CIÊNCIAS
E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

Compressão de Imagem

MULTIMÉDIA

Daniela Lopes, 2018278081
Samuel Machado, 2020219391
Ricardo Guegan, 2020211358

Ano letivo 2022/23

Índice

Questão 1.4	2
Questão 5.4.....	3
Questão 7.....	6
Questão 7.1.3	7
Questão 7.2.3	7
Questão 7.3.2	8
Questão 8.3.....	9
Questão 8.4.....	10
Questão 9.3.....	11

QUESTÃO 1.4

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

	Original	Alta (75)	Média (50)	Baixa (25)
Barn_mountains	349 kB	34 kB 10:1	22 kB 16:1	14 kB 25:1
Logo	412 kB	10 kB 41:1	8 kB 51:1	7 kB 58:1
Peppers	577 kB	31 kB 19:1	21 kB 27:1	14 kB 41:1

Perda de informação perceptível - de carácter subjetivo

	Alta (75)	Média (50)	Baixa (25)
Barn_mountains	p++	p++	p+
Logo	p++	p	p-
Peppers	p++	p+	p

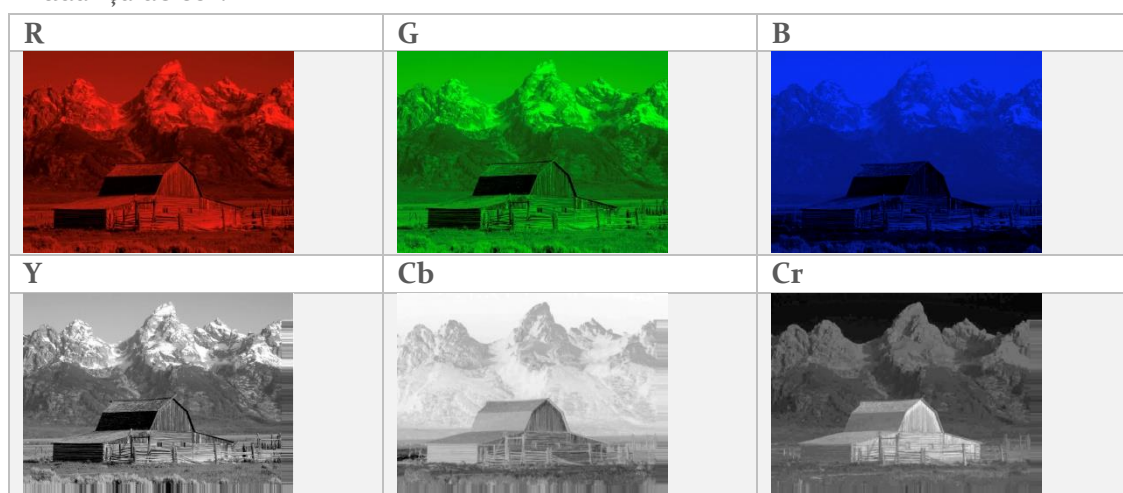
Compressão de alta qualidade

Ao comprimir as imagens com alta qualidade, a única em que foi perceptível perda de qualidade foi a Logo, sendo que acreditamos que isto se deve ao facto de não se tratar de uma imagem foto-realista e conter transições abruptas de cor, pois esta é a dificuldade do JPEG.

Compressão de média e baixa qualidade

Na imagem *logo*, utilizando qualidades de compressão média e baixa, essa perda de qualidade foi gradualmente intensificada.

Nas imagens Barn_mountain e Peppers, podemos verificar também alguma perda de qualidade quando esta foi predefinida para média. No entanto, observámos uma descida abrupta da mesma quando comprimimos as imagens em qualidade baixa, isto é, revelou-se bastante perceptível na sua visualização, principalmente nas zonas de mudança de cor.



QUESTÃO 5.4

Na comparação entre as diferentes componentes de uma imagem YCbCr com as de uma imagem RGB, a componente na qual a imagem original é mais perceptível é a componente Y, que contém a informação da luminância dos diferentes componentes RGB. Em seguida a nível de perceptibilidade, temos as componentes G, R, B, por esta ordem, pois estas contém informação de luminância e crominância e finalmente temos as componentes Cb e Cr que apenas contém informação de crominância. Estes resultados vão ao encontro com a teórica, pois o olho humano é mais sensível à luz e brilho (luminância) do que às cores (crominância). Também está de acordo com os objetivos do JPEG, pois vai permitir resultar num maior nível de destruição nestas componentes.

	Tamanho original	4:2:2	4:2:0
Barn_mountains	349 kB	232,6 kB	174,5 kB
Logo	412 kB	247,6 kB	206 kB
Peppers	577 kB	384,6 kB	288,5 kB
Compressão	1:1	3:2	2:1

Na fase de subamostragem do codec JPEG, aproveitamos o facto de as componentes Cb e Cr serem menos importantes para a nossa perceção da imagem final (conforme mencionado na alínea anterior). Assim, nesta etapa aplicamos compressão destrutiva à informação destas duas componentes, sendo que a implementação tem alguma liberdade na forma com que reescrevemos a imagem vítima de downsampling. As opções mais comuns são o corte apenas das linhas ou de linhas e colunas que não correspondem com os índices (opção mais simples, mas menos precisa) e o uso de interpolações de diferentes graus, no nosso caso optamos por uma interpolação linear (opção mais precisa, porém ligeiramente mais complexa).

Os fatores de downsampling são expressos por meio de um *ratio*, sendo que o primeiro número refere-se ao número de pixels na horizontal da componente luma, a segunda linha é a proporção dos pixels da luma para os de Cb e finalmente a última linha é a proporção de pixels luma para os de Cr; porém, se este último for o, a proporção dos pixels luma para os de Cb aplica-se também aos de Cr e aplica-se não só na horizontal como também na vertical.

Com as seguintes imagens, é possível verificar que a fase da subamostragem não teve qualquer impacto visual notório, mesmo tendo utilizado interpolação linear.



Imagem original



Imagem reconstruída após subamostragem de ratio 4:2:2



Imagem reconstruída após subamostragem de ratio 4:2:0

Caso fosse importante preservar mais a qualidade de qualquer tipo de imagem, acreditamos que poderia ser melhor optar por outra forma de interpolação, visto que, por exemplo, na imagem *logo.bmp*, nas zonas de alterações abruptas de cor, o efeito desta interpolação foi perceptível. Esta perda seria possivelmente minimizada embora consideremos ser incontornável perante imagens não foto-realistas.

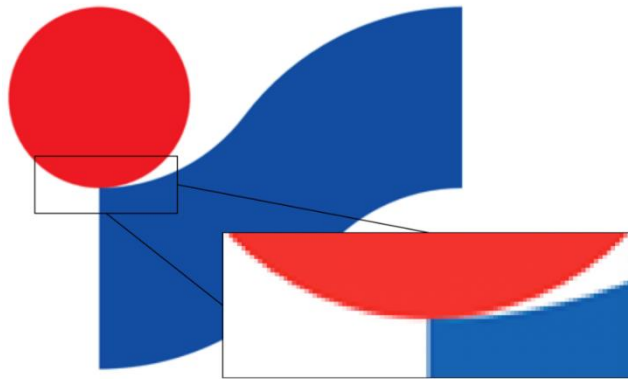


Imagem original

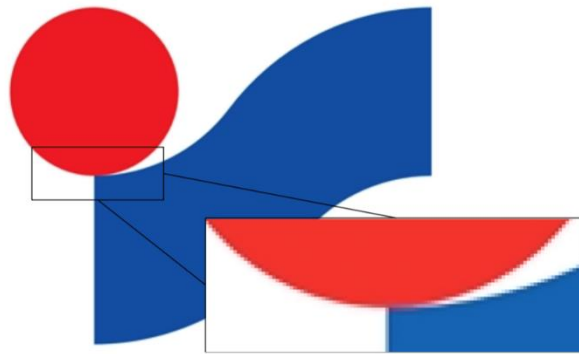


Imagem reconstruída após subamostragem de ratio 4:2:0

QUESTÃO 7

O objetivo da DCT é tirar proveito do facto de que a visão humana tem mais dificuldade a distinguir a intensidade de transições abruptas do que de transições suaves. A DCT permite então codificar a imagem no domínio da frequência de modo que as baixas frequências representam transições suaves e as altas frequências transições abruptas, esta é a vantagem sobre a DFT. Isto permite uma separação das componentes que representam transições abruptas das que representam transições suaves.



Peppers

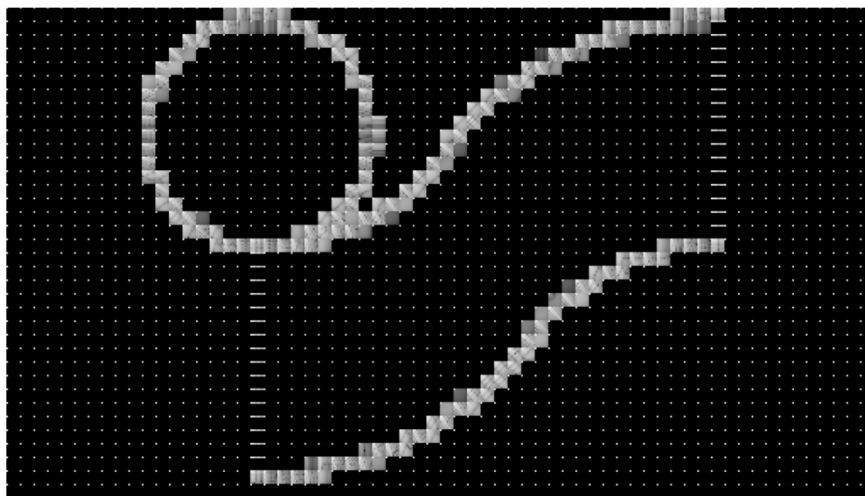
Questão 7.1.3

Como podemos ver nas imagens acima, a aplicação da DCT à imagem completa, “peppers”, devolve coeficientes AC com elevada energia, valores muito dispersos, logo com nível elevado de entropia o que dificulta a compressão de métodos entrópicos, não só isso como ao aplicar a quantização à imagem neste estado a perda de informação seria demasiado elevada.

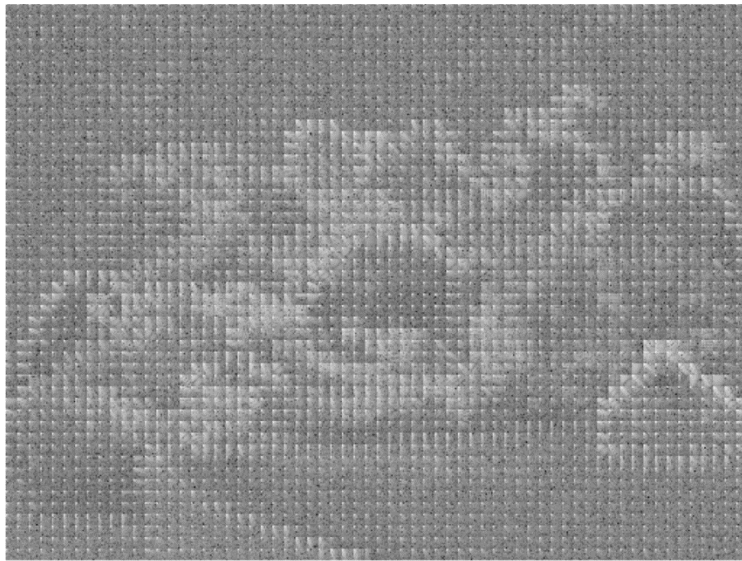
Questão 7.2.3

Vamos agora analisar os resultados da aplicação da DCT em blocos 8x8. Podemos verificar que, de forma geral, em relação à aplicação da imagem na íntegra, a imagem ficou mais uniforme, pois estamos a tirar partido da redundância espacial da imagem, o que faz com que em cada bloco os valores dos coeficientes AC, principalmente os das altas frequências sejam mais baixos isto porque dentro de cada bloco é provável que as transições de cor nele sejam suaves. Isto é especialmente notório da imagem “Logo”, onde os blocos que não contêm transições de cor têm os seus coeficientes AC a zero.

É também importante notar que ao olhar para a imagem Peppers 8x8, podemos ver os contornos dos pimentos a claro pois estas são as zonas de transição de cor abrupta, com mais energia.



Bloco 8x8 – Logo

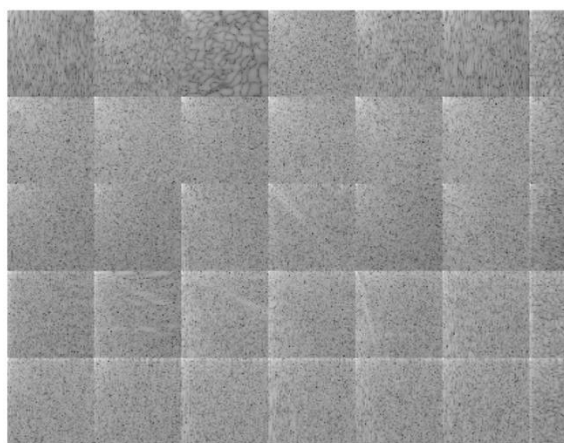


Bloco 8x8 - Peppers

Questão 7.3.2

A aplicação da DCT em blocos 64x64 tem vantagens na medida em que tem uma melhor resolução dos coeficientes das baixas frequências porém tem a desvantagem de não conseguir isolar o contexto espacial como faz a aplicação em blocos 8x8, logo como podemos ver na imagem voltamos a ter valores bastantes claros nas frequências mais elevadas o que quer dizer que temos várias transições abruptas, o que reduz bastante o potencial da quantização, pois não nos vai permitir descartar tão bem a informação das altas frequências. Voltamos também a ter uma range de valores muito mais amplos, o que vai diminuir a performance dos métodos de codificação entrópicos.

Após esta análise, concluímos que para imagens foto realistas como o Peppers e o Barn_mountains os blocos 8x8 são aqueles que têm um melhor compromisso de resolução das diferentes frequências e de isolamento do contexto.



Barn Mountains 64x64

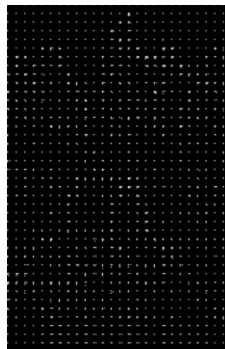
QUESTÃO 8.3

Quantização, esta é mais uma das etapas do JPEG onde a implementação pode variar de acordo com as matrizes de quantização escolhidas para cada componente. Este processo procura, dado um fator de qualidade, maximizar a taxa de compressão enquanto procura minimizar as perdas perceptuais, o que leva a uma maior destruição nas componentes das elevadas frequências, relativamente às baixas. Desta forma é possível optar por diferentes compromissos quanto à relação qualidade/compressão.

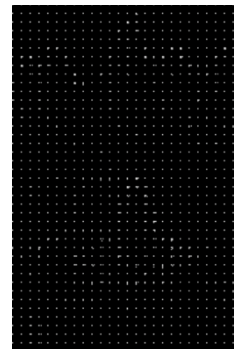
QF = 25



Canal Y

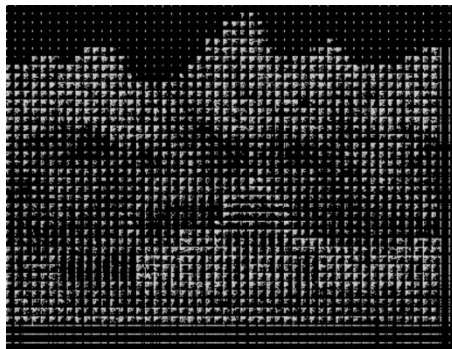


Canal Cb

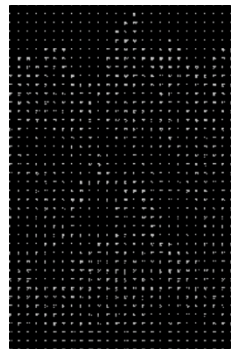


Canal Cr

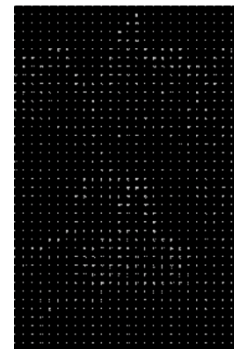
QF = 50



Canal Y

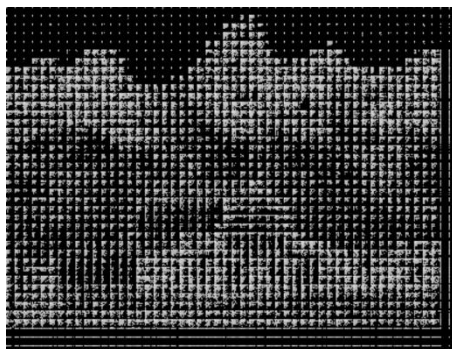


Canal Cb

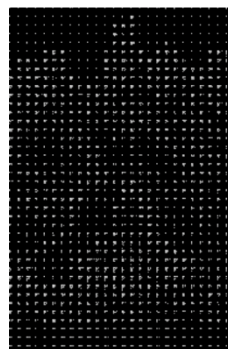


Canal Cr

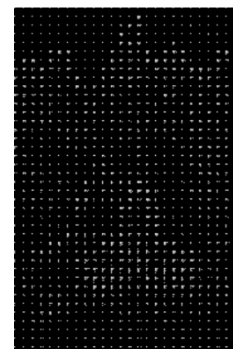
QF = 75



Canal Y



Canal Cb



Canal Cr

Um fator de qualidade mais baixo leva a uma imagem mais escura, quer isto dizer que os valores da matriz da quantização são maiores, ou seja, estaremos a dividir por valores mais elevados e, conseqüentemente, obteremos valores mais próximos de zero. Assim, poderemos utilizar melhor métodos de codificação como o RLE e os métodos de codificação entrópica.

Esta proporção entre o quão escura é a imagem (valores baixos) e o quão mais baixo é o fator de qualidade pode ser comprovada visualmente com as imagens acima.

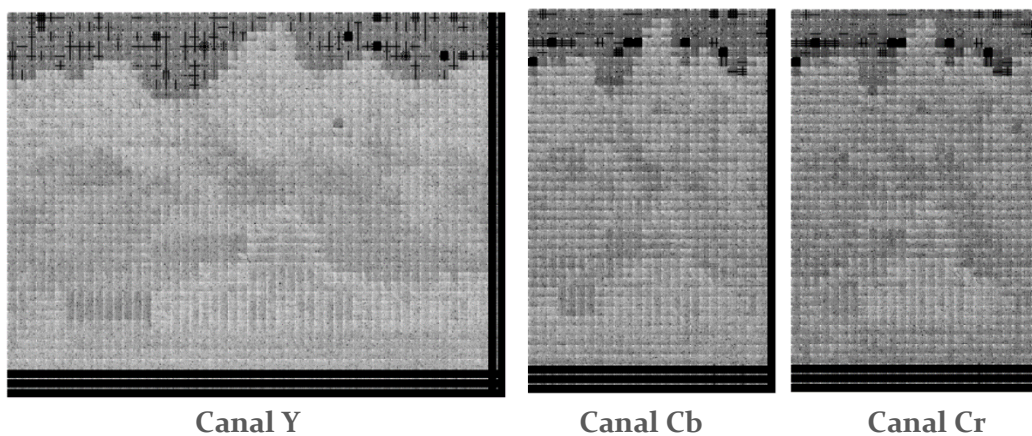
Outra observação pertinente é o quão mais escuras ficam as componentes Cb e Cr em relação à luma, o que é normal pois aplicamos uma matriz de quantização com coeficientes geralmente maiores a essas componentes pois no modelo YCbCr a informação das componentes Cb e Cr é menos importante a nível de percepção da imagem final do que a da luma.

QUESTÃO 8.4

Ao fazer a comparação entre as imagens pré e pós Quantização, a primeira característica que notamos é que esta fica mais escura, pois como foi explicado acima estamos a dividir cada bloco pela matriz de quantização que lhe corresponde. A segunda característica que é importante notar é que dentro de cada bloco, após a quantização, quanto mais próximo o coeficiente AC se encontra do coeficiente DC, mais a sua cor é semelhante à original, ou seja, pré quantização. Este resultado é o esperado pois tal como explicamos acima queremos ter uma maior perda de informação nos coeficientes das altas frequências do que nos coeficientes das baixas frequências.

Pequena exceção para a imagem logo na qual os blocos que não contém transições são sempre o, pouco importa o fator de qualidade.

Resultado após DCT

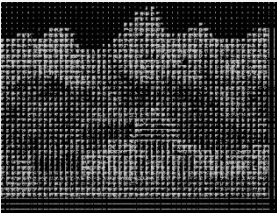
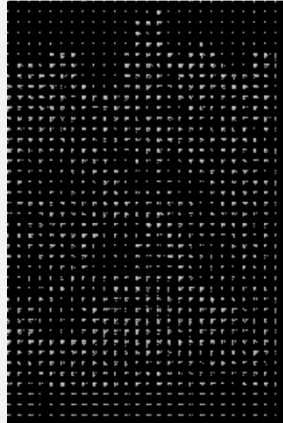
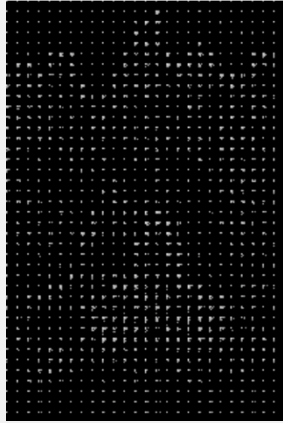
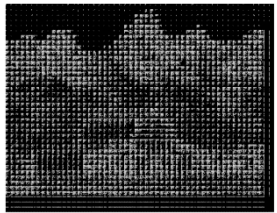
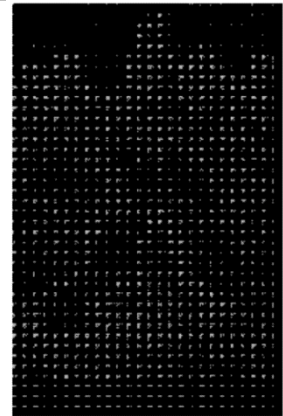
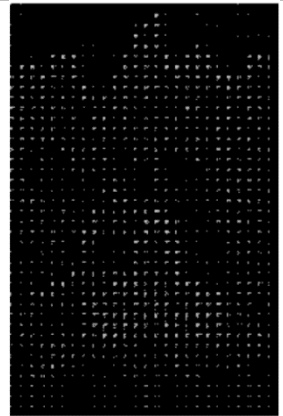


QUESTÃO 9.3



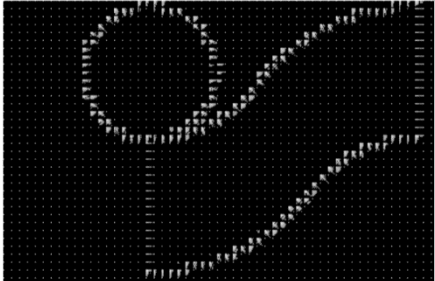
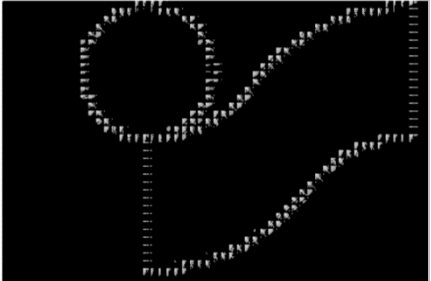
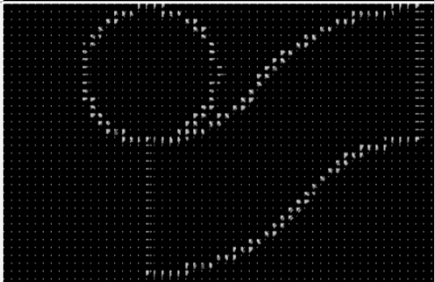
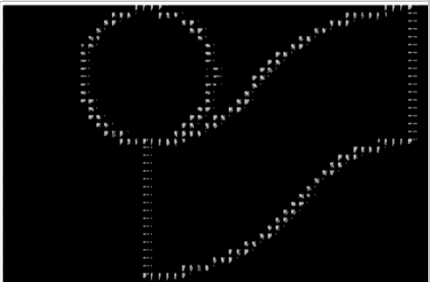
A codificação DPCM dos coeficientes AC tem como objetivo reduzir os valores elevados dos coeficientes AC para assim poder melhor métodos de codificação entrópica. Estes coeficientes serão codificados como uma série de diferenças em relação ao anterior (codificação diferencial), tirando proveito do contexto espacial e conseguindo assim que estes valores fiquem mais próximos dos valores vizinhos. Isto pode ser observado nas seguintes imagens, principalmente na zona do céu da imagem Barn_mountains e em zonas sem transições do Logo.

É de destacar que este método de codificação dos coeficientes, através do cálculo das diferenças em linha, tem dificuldade em comprimir os primeiros coeficientes de cada linha, pois estes têm a sua diferença feita em relação ao último da linha anterior. Consequentemente, acabam por não fazer bem o uso do contexto espacial. Existem alternativas que evitam este problema na codificação, mas não implementadas no JPEG.

Exemplos de componentes Y, Cb e Cr antes e depois da DPCM

Fator de Qualidade 75	Y	Cb	Cr
Sem codificação DPCM			
Com codificação DPCM			

Exemplos de componentes Y antes e depois da DPCM

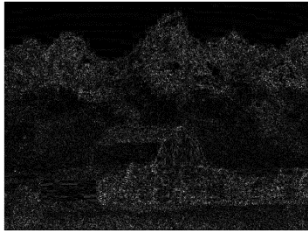
	Componente Y sem DPCM	Componente Y com DPCM
Barn_Mount a ins Fator de Qualidade 75		
Logo Fator de Qualidade 75		
Logo Fator de Qualidade 25		

Métricas de distorção

	10	25	50	75	100
Barn mountains 4:2:2	MSE: 702.603 RMSE: 26.507 SNR: 18.719 PSNR: 19.664 Erro Máx: 119.73 Erro Médio: 9.47	MSE: 396.030 RMSE: 19.901 SNR: 21.209 PSNR: 22.154 Erro Máx: 82.8 Erro Médio: 6.96	MSE: 259.460 RMSE: 16.108 SNR: 23.046 PSNR: 23.990 Erro Máx: 87.4 Erro Médio: 5.55	MSE: 152.144 RMSE: 12.335 SNR: 25.364 PSNR: 26.308 Erro Máx: 44.1 Erro Médio: 4.15	MSE: 12.033 RMSE: 3.468 SNR: 36.3829 PSNR: 37.327 Erro Máx: 3.57 Erro Médio: 0.51
Logo 4:2:2	MSE: 213.536 RMSE: 14.612 SNR: 28.037 PSNR: 24.836 Erro Máx: 102.31 Erro Médio: 5.38	MSE: 65.622 RMSE: 8.100 SNR: 33.162 PSNR: 29.960 Erro Máx: 58.72 Erro Médio: 1.02	MSE: 48.642 RMSE: 6.974 SNR: 34.468 PNR: 31.260 Erro Máx: 46.5 Erro Médio: 1.72	MSE: 23.929 RMSE: 4.891 SNR: 37.542 PSNR: 34.341 Erro Máx: 31.7 Erro Médio: 0.74	MSE: 6.170 RMSE: 2.484 SNR: 43.428 PSNR: 40.227 Erro Máx: 13.09 Erro Médio: 0.41
Peppers 4:2:2	MSE: 279.831 RMSE: 16.728 SNR: 20.456 PSNR: 23.661 Erro Máx: 93.22 Erro Médio: 4.43	MSE: 125.802 RMSE: 11.216 SNR: 23.929 PSNR: 27.133 Erro Máx: 74.5 Erro Médio: 2.69	MSE: 77.055 RMSE: 8.778 SNR: 26.057 PSNR: 29.262 Erro Máx: 32.8 Erro Médio: 2.01	MSE: 50.546 RMSE: 7.109 SNR: 27.889 PSNR: 31.093 Erro Máx: 29.7 Erro Médio: 1.56	MSE: 8.458 RMSE: 2.908 SNR: 35.652 PSNR: 38.857 Erro Máx: 6.82 Erro Médio: 0.52

Nota: os valores com ratio de subamostragem 4:2:0 seguem em anexo

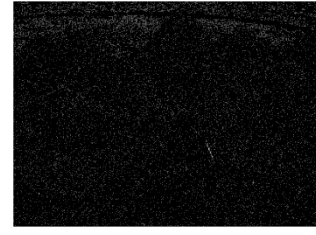
Imagens de diferenças (25-75-100)



QF = 25

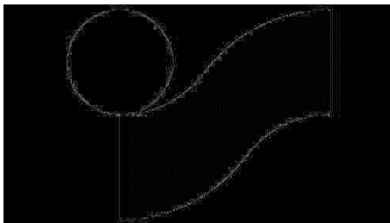


QF = 75

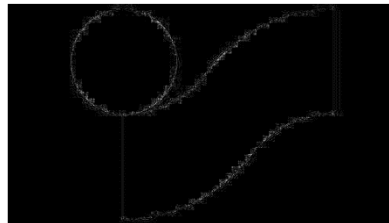


QF = 100

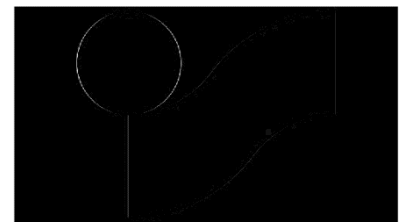
Imagem Barn_mountains



QF = 25

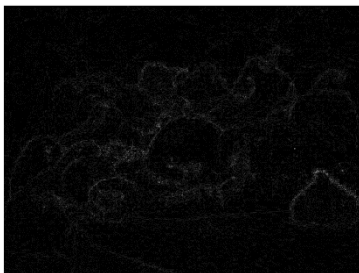


QF = 75

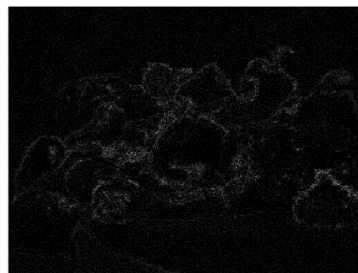


QF = 100

Imagem Logo



QF = 25



QF = 75



QF = 100

Imagem Peppers

Ao observar as imagens de diferenças conseguimos visualizar que, como esperado, a intensidade dos valores das diferenças é inversamente proporcional ao fator de qualidade. É de realçar que facilmente observamos que, quando o fator de qualidade é 100, as diferenças surgem mais dispersas, pois apenas são consequência de *clampings*, arredondamentos e da subamostragem. Ao comparar o fator de qualidade 100 com os outros fatores de qualidade, verificamos que a maior parte do processo destrutivo é feito na etapa de quantização e principalmente nas zonas de transição, daí termos os contornos das diferentes entidades da imagem relativamente bem definidos.

As primeiras duas métricas (MSE e RMSE) devem ser cada vez menores à medida que aumentamos o fator de qualidade, visto que expressam a diferença entre a imagem reconstruída e a nova imagem (ou seja, o erro). O SNR e o PNSR calculam a proporção sinal/ruído, o que leva a que se espere que as imagens comprimidas com um fator de qualidade maior estejam associadas a um valor mais elevado.

Ao observarmos a tabela acima disponibilizada podemos verificar que os valores empíricos foram ao encontro do esperado (mencionado no parágrafo anterior). Pudemos ainda constatar a redução dos valores erro médio e máximo à medida que o fator de qualidade aumenta (salvo exceções ocasionais).