

Trabalho Prático nº 1

Licenciatura em engenharia informática



FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE DE
COIMBRA

Trabalho realizado por:

Diogo José Gaspar Ferreira nº 2022213517

Gonçalo Fernandes Ferreira, nº2022210563

João Guilherme Pilar de Figueiredo nº 2022215732

José Pedro de Pinho Santos, nº 2022227876

Exercício 1.4

Neste exercício, à medida que o fator de qualidade das imagens é reduzido, a percepção da perda de detalhes e degradação da qualidade visual aumenta, os pormenores tornam-se menos claros e dá uma sensação de desfoque. A razão para a perda de qualidade depende apenas do facto que a compressão segundo o codec JPEG é destrutiva, causando uma grande perda de informação relevante da imagem à medida que o fator de qualidade também diminui.

Em suma, derivado do codec JPEG ser destrutivo, se o fator de qualidade for baixo, a taxa de compressão vai ser elevada e consequentemente a degradação na qualidade será notória. Inversamente, se o fator de qualidade for elevado, a taxa de compressão será menor resultando em menos destruição.

Imagem	Qualidade	Taxa de Compressão
Airport	25	$\frac{2\,430\,122\text{ bytes}}{67\,916\text{ bytes}} = 35.78 \Rightarrow 36:1$
Airport	50	$\frac{2\,430\,122\text{ bytes}}{105\,143\text{ bytes}} = 23.11 \Rightarrow 23:1$
Airport	75	$\frac{2\,310\,000\text{ bytes}}{158\,287\text{ bytes}} = 15.35 \Rightarrow 15:1$
Geometric	25	$\frac{3\,003\,122\text{ bytes}}{34\,506\text{ bytes}} = 87.03 \Rightarrow 87:1$
Geometric	50	$\frac{3\,003\,122\text{ bytes}}{46\,387\text{ bytes}} = 64.74 \Rightarrow 65:1$
Geometric	75	$\frac{3\,003\,122\text{ bytes}}{60\,681\text{ bytes}} = 49.49 \Rightarrow 49:1$
Nature	25	$\frac{5\,760\,122\text{ bytes}}{119\,910\text{ bytes}} = 48.04 \Rightarrow 48:1$
Nature	50	$\frac{5\,760\,122\text{ bytes}}{310\,963\text{ bytes}} = 18.52 \Rightarrow 19:1$
Nature	75	$\frac{5\,760\,122\text{ bytes}}{401\,000\text{ bytes}} = 14.00 \Rightarrow 14:1$

Exercício 5.5

Comparando Y em relação a R, G e B

A luminância (Y) representa a intensidade da luz na imagem e é influenciada pelos canais RGB.

O canal G (verde) é o que tem mais influência sobre Y, estando estes canais bastante correlacionados.

O canal B (azul) exerce a menor influência sobre Y, o que significa que áreas dominadas pelo azul aparecem mais escuras na luminância (Y).

O canal R (vermelho) também influencia Y, mas menos do que G.

Comparando Y em relação a Cb e Cr

Cb (Diferença Azul - Luminância): Representa a quantidade de azul na imagem em relação à luminância. Áreas ricas em azul terão valores mais altos de Cb, enquanto regiões com pouca presença de azul terão valores baixos.

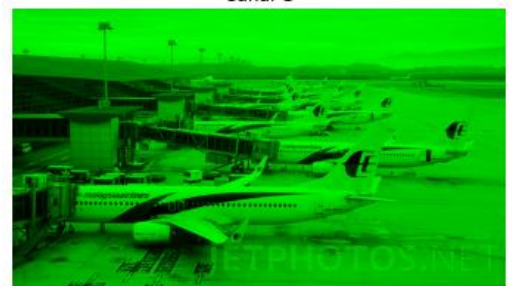
Cr (Diferença Vermelho - Luminância): Mede a quantidade de vermelho na imagem em relação à luminância. Áreas com muito vermelho irão apresentar valores elevados de Cr.

Como Y é uma média ponderada dos três canais RGB, ele não pode ser diretamente comparado a Cb e Cr, pois estes representam a diferença entre a luminância e as cores individuais. No entanto, juntos, esses componentes permitem reconstruir a imagem original separando a informação de brilho das informações cromáticas.

Canal R



Canal G



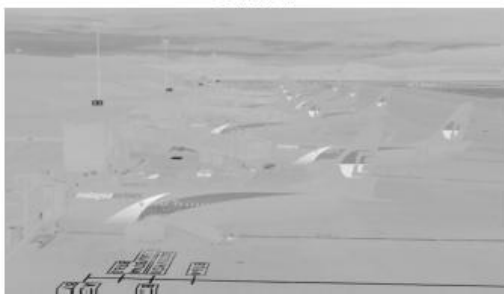
Canal B



Canal Cr



Canal Cb



Canal Y



Exercício 6.5

No codec JPEG, o downsampling aproveita a menor sensibilidade do olho humano às variações de cor, reduzindo os dados dos canais Cb e Cr.

Taxa de compressão

Ao usar o downsampling 4:2:2 verificou-se uma redução do tamanho horizontal dos canais Cb e Cr para metade do seu tamanho original. Por outro lado, ao usar downsampling 4:2:0 verificou-se que os canais Cb e Cr sofreram uma redução de tamanho horizontal e vertical de 50%.

Perda de Qualidade (Destrutividade)

Através da compressão 4:2:2 é perceptível uma redução menos acentuada da qualidade da imagem em relação à original, indicando uma baixa destrutividade. Contrariamente, quando aplicada uma compressão 4:2:0 a destrutividade é maior resultando numa perda de qualidade mais significativa.

Y (Downsampling (linear) with [4, 2, 0])



Y (Downsampling (linear) with [4, 2, 2])



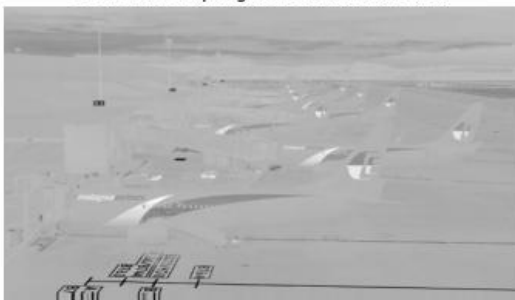
Cr (Downsampling (linear) with [4, 2, 0])



Cr (Downsampling (linear) with [4, 2, 2])



Cb (Downsampling (linear) with [4, 2, 0])



Cb (Downsampling (linear) with [4, 2, 2])



Exercício 7.4

Neste exercício é solicitada a criação de algum método para o cálculo da DCT (Transformada do Cosseno Discreta). Para calcular a DCT de um canal completo foi criada a função “*get_dct()*” enquanto que para os blocos de 8x8 e 64x64 foi criada outra função, “*dct_by_block()*”.

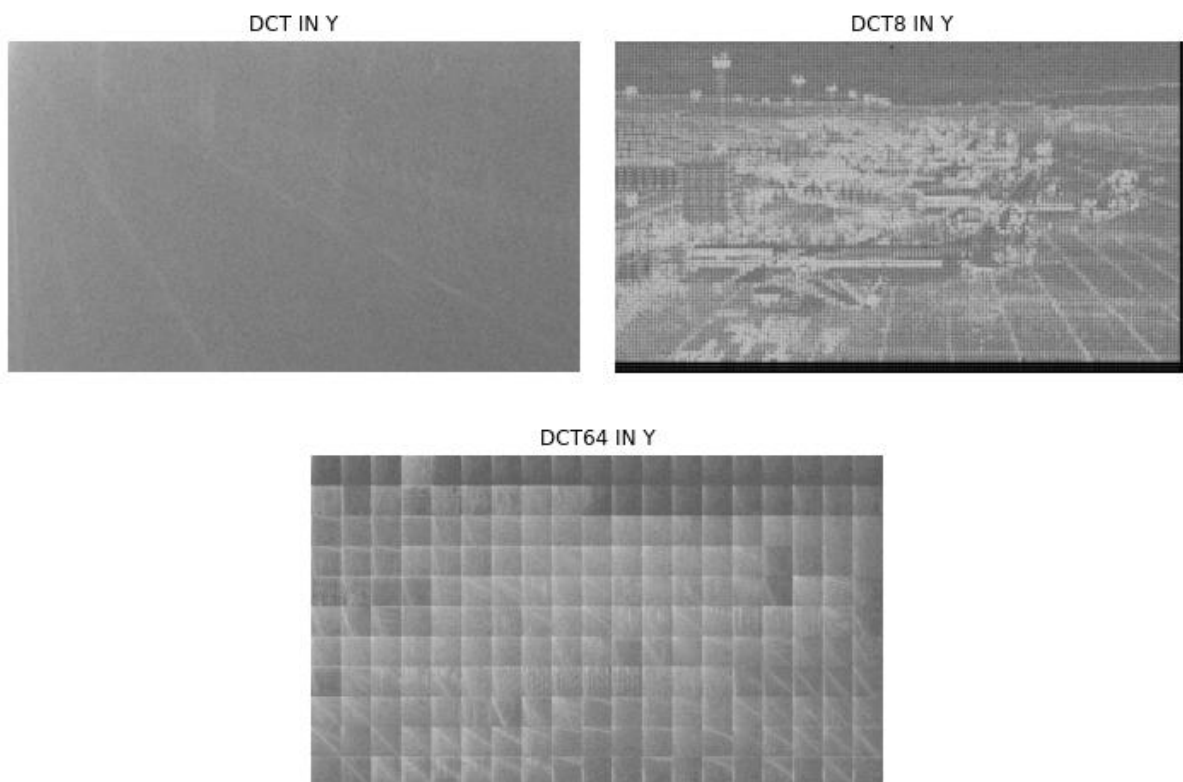
Conclusão sobre o Potencial de Compressão

DCT de um canal completo (ex 7.1) → Potencial de compressão elevado, descarta os pequenos detalhes e explora a correlação de toda a imagem. Não é perceptível a existência de artefactos na imagem.

DCT em blocos 8x8 (ex 7.2) → Equilíbrio entre compressão e qualidade (usado em JPEG). Dos 3 tamanhos de blocos, a probabilidade de surgirem artefactos é mais elevada neste, como é visível na imagem abaixo.

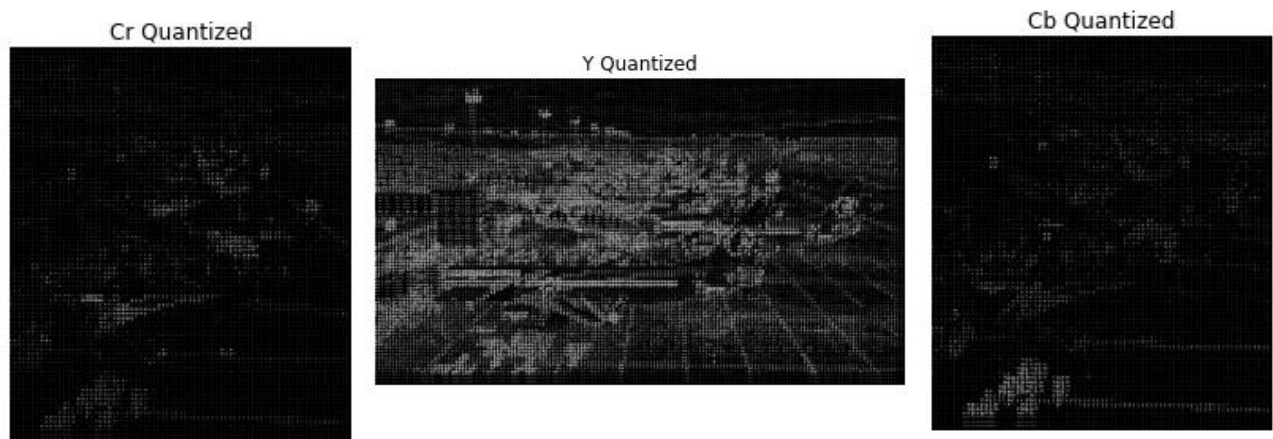
DCT em blocos 64x64 (ex 7.3) → Maior compressão, mas maior perda de detalhes sendo útil para imagens com menos textura.

Concluimos então que quanto maior for o tamanho do bloco, menor será o número de blocos e consequentemente, a visibilidade de artefactos diminui.



Exercício 8.5

Quando usamos valores baixos para o fator de qualidade (por exemplo 10 ou 25), a quantização é mais intensa. Isso significa que eliminamos muitos detalhes, sobretudo nas frequências mais altas, o que resulta numa compressão maior, mas também numa imagem com menos nitidez. Por outro lado, valores mais altos (como 75 ou 100) aplicam uma quantização mais suave, mantendo mais informação e, consequentemente, a qualidade visual, mas gerando ficheiros mais pesados. É, basicamente, um equilíbrio entre reduzir o tamanho do ficheiro e preservar a qualidade.



Exercício 8.6

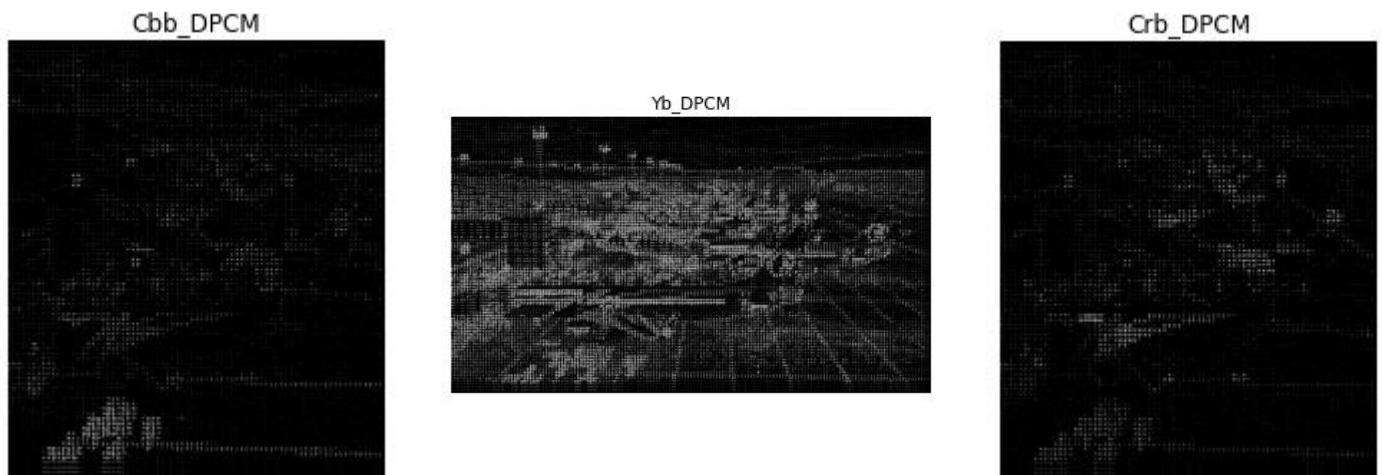
Na alínea 7 usamos a DCT para transformar a imagem para o domínio da frequência, de forma a concentrar a energia nos coeficientes de baixa frequência. Isso, por si só, não perde informação, apenas organiza os dados de modo a facilitar a compressão. Já na alínea 8, a quantização é que efetivamente elimina parte da informação, sobretudo nos coeficientes de alta frequência, onde o olho humano é menos sensível. Assim, enquanto a DCT prepara os dados para serem comprimidos de forma eficiente, a quantização define o quanto vamos sacrificar em termos de qualidade visual para conseguir uma compressão maior.

Exercício 9.5

A DPCM é um método de compressão que diminui a quantidade de dados usados para representar uma imagem ao usar a diferença de amostras consecutivas.

Este método revela-se muito bom para imagens com transições suaves uma vez que a variação entre os coeficientes DC de blocos adjacentes é baixa indicando uma alta correlação. Isto leva a uma gama de valores mais estreita diminuindo a entropia, facilitando a compressão uma vez que os valores mais repetidos necessitam de menos bits para serem codificados.

Observou-se que a matriz continha uma grande quantidade de zeros mostrando que existe uma alta correlação facilitando a compressão entrópica utilizada no codec JPEG.



Exercício 10.5

Análise das Métricas

As métricas abaixo analisadas são referentes à imagem *airport* conforme era pedido no enunciado com um fator de qualidade de 75%.

MSE (Erro Quadrático Médio) = 129,97

Mede a diferença quadrática média entre os pixels da imagem original e da imagem decodificada.

Um MSE acima de 100 sugere uma perda perceptível de informação.

RMSE (Raiz do Erro Quadrático Médio) = 11,40

É a raiz quadrada do MSE.

Valores abaixo de 10 sugerem que muito pouca informação foi perdida. Como o RMSE é 11.40, podemos deduzir que a imagem sofreu alguma degradação.

SNR (Relação Sinal-Ruído) = 28,51 dB

Indica quanta informação original foi preservada em relação ao ruído introduzido pela compressão (um coeficiente de correlação).

Valores acima de 30 dB são considerados bons, então 28,51 dB sugere alguma degradação.

PSNR (Pico da Relação Sinal-Ruído) = 26,99 dB

Compara a intensidade máxima dos pixels com o MSE.

Geralmente, valores acima de 30 dB são bons, enquanto abaixo de 25 dB indicam uma perda perceptível de qualidade.

Com 26,99 dB, há alguma degradação visível, mas a imagem ainda mantém uma qualidade razoável.

AVG_diff (Diferença Média no Canal Y) = 2,62

A diferença média dos valores dos pixels no canal Y (luminância) entre a imagem original e a decodificada é de 2,62.

Pequenas variações são normais, mas valores mais altos podem indicar perda de informação.

Max_diff (Diferença Máxima no Canal Y) = 42,59

Representa a maior diferença de pixel encontrada no canal Y.

Um valor elevado não implica necessariamente um problema generalizado, mas sim que alguns pixels específicos sofreram uma grande perda de informação.

Abaixo segue-se uma tabela com as métricas da imagem *airport* com diferentes fatores de qualidade. Apesar de não ter sido apresentado o mesmo para a imagem *geometric* e *nature* a relação é semelhante.

Fator de Qualidade	25	50	100
MSE	287.9622913580247	164.3970864197531	11.23755061728395
RMSE	16.969451710589375	12.821742721633168	3.3522456081385132
SNR	25.057548187029152	27.49196323004734	39.144187784155775
PSNR	23.53744740217513	25.97186244519332	37.62408699930176
AVG_diff	5.007052245614098	3.6720244246759153	0.20510580381904403
Max_diff	97.9491848494323	75.9967992374397	1.4460648838243912