

Trabalho Prático n.º 1

Compressão de Imagem

Multimédia

Licenciatura em Engenharia Informática

2021/2022

Duarte Emanuel Ramos Meneses – 2019216949 – duartemeneses@student.dei.uc.pt

Inês Martins Marçal – 2019215917 – inesmarcal@student.dei.uc.pt

Patrícia Beatriz Silva Costa – 2019213995 – patriciacosta@student.dei.uc.pt

Índice

[Pergunta 1 5](#_Toc97129422)

[Pergunta 5 6](#_Toc97129423)

[Pergunta 6 8](#_Toc97129424)

[Pergunta 7 10](#_Toc97129425)

Pergunta 1

Compressão de imagens bmp no formato jpeg utilizando um editor de imagem (GIMP).

|  |  |  |  |
| --- | --- | --- | --- |
|  | Tamanho em disco | | |
| Qualidade | barn\_mountains.bmp | logo.bmp | peppers.bmp |
| Original | 349 KB | 412 KB | 577 KB |
| Alta (75%) | 34 KB | 10 KB | 31 KB |
| Média (50%) | 22 KB | 8 KB | 21 KB |
| Baixa (25%) | 14 KB | 7 KB | 14 KB |

Como podemos verificar analisando os dados da tabela acima, quando se comprime uma imagem, o tamanho em disco da resultante é inferior ao original (qualquer que seja a qualidade dessa compressão). Fica ainda evidente nos dados acima que quanto mais baixa a qualidade de compressão, menor o espaço de armazenamento necessário.

Estas compressões foram feitas, aproximadamente, nas seguintes proporções:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Proporção de compressão | | |
| Qualidade | barn\_mountains.bmp | logo.bmp | peppers.bmp |
| Alta (75%) | 10:1 | 41:1 | 19:1 |
| Média (50%) | 16:1 | 52:1 | 27:1 |
| Baixa (25%) | 25:1 | 59:1 | 41:1 |

Cada proporção ajuda-nos a entender melhor a quantidade de compressão existente. Por exemplo, se a proporção for de 10:1 significa que 10 pixéis se tornaram em apenas 1. Deste modo, fica evidente que quanto maior for o valor do lado esquerdo da proporção, maior a taxa de compressão.

Quanto à qualidade das imagens resultantes, tal como era esperado, quanto menor a qualidade da compressão, menor a qualidade do resultado final. Quanto pior qualidade de compressão se tem, maior a perda de informação, logo mais degradada a imagem ficará ao olho humano. Por exemplo, qualquer que seja a imagem, ao diminuir a qualidade de compressão, a resultante fica com grão.

Na imagem do logo, qualquer tipo de compressão causa ruído. Isto acontece, pois, o JPEG é bom para imagens com transições suaves. Para os logos, a melhor solução é o PNG.

# Pergunta 5

Como no modelo RGB existe redundância da luminância nos 3 canais (R, G e B), o modelo YCbCr tem como intuito reduzir essa mesma redundância. Deste modo, o YCbCr separa a luminância da componente cromática. Esta última é a menos sensível ao olho humano.

Aplicando um colormap para tons de cinzento a cada canal do modelo YCbCr, obtemos os seguintes resultados para a imagem peppers.bmp:

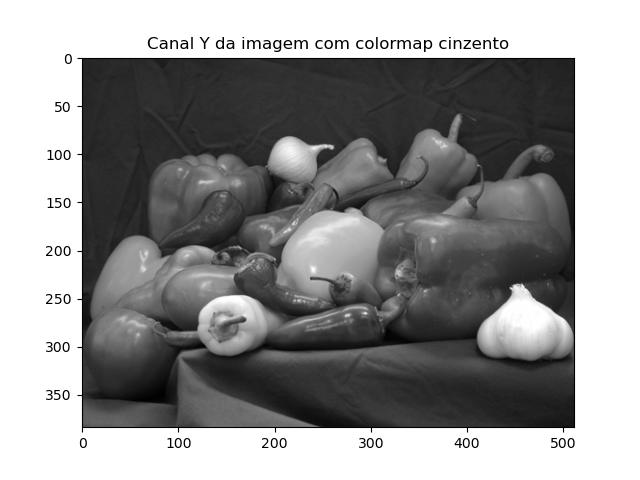


Figura 1 - Canal Y do modelo YcbCr da imagem peppers.bmp utilizando um colormap cinzento

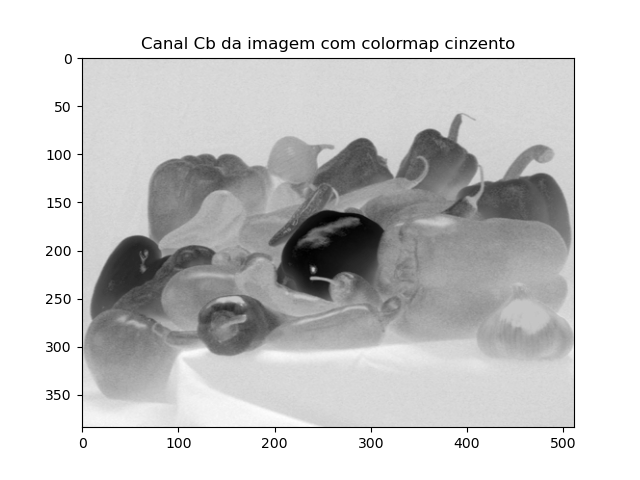
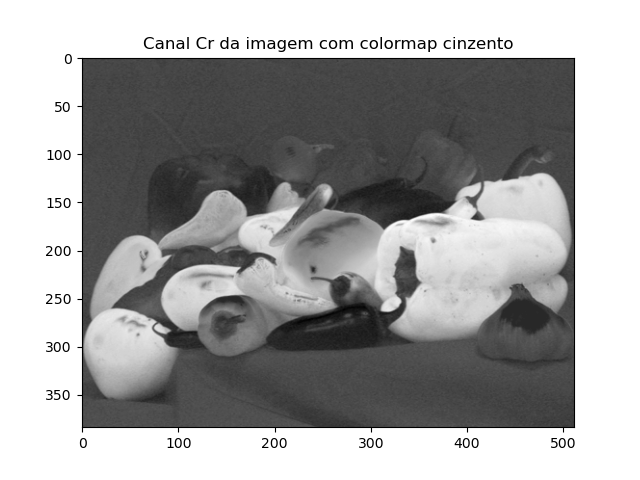
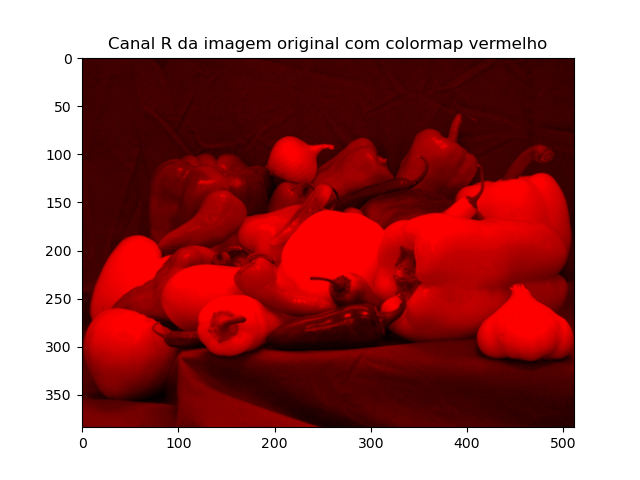


Figura 2 - Canal Cr do modelo YcbCr da imagem peppers.bmp utilizando um colormap cinzento

Figura 3 - Canal Cb do modelo YcbCr da imagem peppers.bmp utilizando um colormap cinzento

Analisando as imagens acima comprovamos o que enunciamos anteriormente. Efetivamente, é no canal Y que conseguimos visualizar a imagem com mais detalhe. Isto acontece uma vez que este canal é o que contem a luminância, enquanto os Cb e Cr contêm a crominância. Como o olho humano é mais sensível à luminância (canal Y) que à crominância (canais Cb e Cr), é a figura 1 que apresenta mais detalhe.

É fácil de perceber, portanto que, sendo os canais Cb e Cr os que menos detalhe apresentam (menos relevância na construção da imagem pois o olho humano não é tão sensível à crominância, propriedade destes canais), são os que servirão de base à compressão. Menor informação nestes canais em pouco influenciará a perceção humana da imagem no seu todo. Já se essa compressão fosse realizada no canal Y, o olho humano detetaria mais facilmente.

Os resultados da imagem peppers.bmp dividida nos seus canais R, G e B, cada um com o colormap adequado (cada canal com um colormap da cor em questão) são os seguintes:

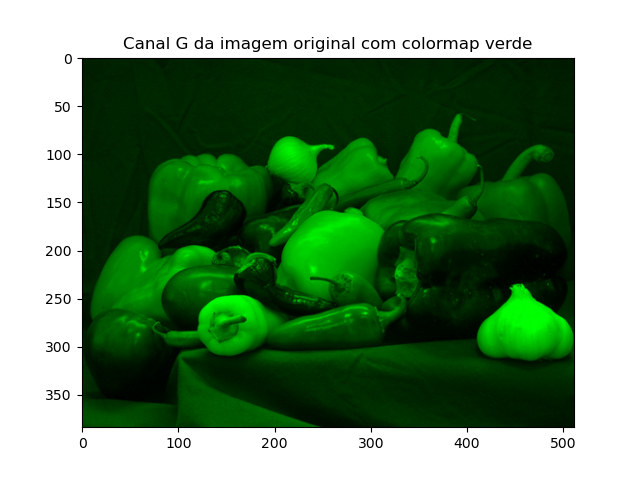


Figura 4 - Canal G do modelo RGB da imagem peppers.bmp utilizando um colormap verde

Figura 5 - Canal R do modelo RGB da imagem peppers.bmp utilizando um colormap vermelho

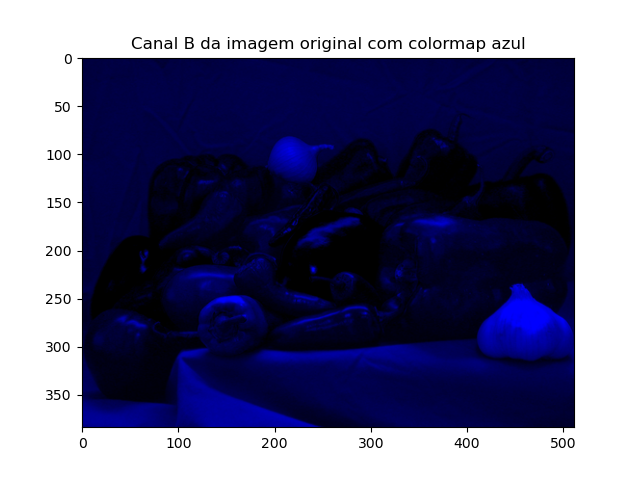


Figura 6 - Canal B do modelo RGB da imagem peppers.bmp utilizando um colormap azul

Comparando agora o canal Y do modelo YCbCr com os canais R, G e B do modelo RGB, fica evidente que existe detalhe em todos. Isto acontece, pois, tal como já referimos em cima, no modelo RGB existe redundância da luminância nos 3 canais. Sendo esta a que está presente no canal Y do YCbCr, é normal que o nível de detalhe seja similar aos canais R, G e B.

Uma questão que nos fez refletir foi o facto de o nível de detalhe entre os canais R, G e B serem ligeiramente diferentes. O canal G é o que apresenta mais detalhe, seguido do R e, por último, o B. Isto acontece, pois, o olho humano é mais sensível às tonalidades verdes, depois às vermelhas e só no fim às azuis. Desta forma, apesar de ambas apresentarem luminância (propriedade a que o olho humano é mais sensível), a cor verde é a que consegue representar mais fielmente os detalhes da imagem original.

# Pergunta 6

* barn\_mountains.bmp

|  |  |  |
| --- | --- | --- |
| Canal | | Tamanho |
| Y | Original | (304, 400) |
| Com downsampling 4:2:0 | (304, 400) |
| Com downsampling 4:2:2 | (304, 200) |
| Cb | Original | (304, 400) |
| Com downsampling 4:2:0 | (152, 200) |
| Com downsampling 4:2:2 | (304, 200) |
| Cr | Original | (304, 400) |
| Com downsampling 4:2:0 | (152, 200) |
| Com downsampling 4:2:2 | (304, 200) |

* logo.bmp

|  |  |  |
| --- | --- | --- |
| Canal | | Tamanho |
| Y | Original | (288, 512) |
| Com downsampling 4:2:0 | (288, 512) |
| Com downsampling 4:2:2 | (288, 512) |
| Cb | Original | (288, 512) |
| Com downsampling 4:2:0 | (144, 256) |
| Com downsampling 4:2:2 | (288, 256) |
| Cr | Original | (288, 512) |
| Com downsampling 4:2:0 | (144, 256) |
| Com downsampling 4:2:2 | (288, 256) |

* peppers.bmp

|  |  |  |
| --- | --- | --- |
| Canal | | Tamanho |
| Y | Original | (384, 512) |
| Com downsampling 4:2:0 | (384, 512) |
| Com downsampling 4:2:2 | (384, 512) |
| Cb | Original | (384, 512) |
| Com downsampling 4:2:0 | (192, 256) |
| Com downsampling 4:2:2 | (384, 256) |
| Cr | Original | (384, 512) |
| Com downsampling 4:2:0 | (192, 256) |
| Com downsampling 4:2:2 | (384, 256) |

Tal como era espectável, em qualquer uma das imagens, o canal Y não sofreu alterações quer no número de colunas quer no número de linhas nas duas variantes de downsampling. Isto acontece, pois, a proporção é de 4:2:2 ou 4:2:0, o que indica que os canais Cb e Cr tem a sua taxa de amostragem reduzida para metade na direção horizontal (no caso de 4:2:2) ou que estes mesmo dois canais têm a sua taxa de amostragem reduzida para metade em ambas as direções (no caso de 4:2:0).

Os canais Cb e Cr são os que sofrem compressão uma vez que são os que o olho humano é menos sensível. O Y, sendo o que é mais sensível, não sofre compressão para a imagem não ficar distorcida ao olho humano.

Fica evidente que a proporção 4:4:2, reduzindo apenas a taxa de amostragem na direção horizontal, tem menor destrutividade que a proporção 4:4:0 que reduz a taxa de amostragem em ambas as direções.

# Pergunta 7.1

* barn\_mountains.bmp

|  |  |  |
| --- | --- | --- |
| Canal | Com downsampling 4:2:0 | Com downsampling 4:2:2 |
| Y |  |  |
| Cb |  |  |
| Cr |  |  |

* logo.bmp

|  |  |  |
| --- | --- | --- |
| Canal | Com downsampling 4:2:0 | Com downsampling 4:2:2 |
| Y |  |  |
| Cb |  |  |
| Cr |  |  |

* peppers.bmp

|  |  |  |
| --- | --- | --- |
| Canal | Com downsampling 4:2:0 | Com downsampling 4:2:2 |
| Y |  |  |
| Cb |  |  |
| Cr |  |  |

Como podemos visualizar pelas imagens acima, percebemos que uma concentração da maior parte das frequências da imagem original num número reduzido de pixéis. Isto vai, naturalmente, ajudar à compressão.

É também evidente que as imagens não são distinguíveis pela sua representação após aplicar a DCT.

# Pergunta 7.2

* barn\_mountains.bmp

|  |  |  |
| --- | --- | --- |
| Canal | Com downsampling 4:2:0 | Com downsampling 4:2:2 |
| Y |  |  |
| Cb |  |  |
| Cr |  |  |

* logo.bmp

|  |  |  |
| --- | --- | --- |
| Canal | Com downsampling 4:2:0 | Com downsampling 4:2:2 |
| Y |  |  |
| Cb |  |  |
| Cr |  |  |

* peppers.bmp

|  |  |  |
| --- | --- | --- |
| Canal | Com downsampling 4:2:0 | Com downsampling 4:2:2 |
| Y |  |  |
| Cb |  |  |
| Cr |  |  |

Comparando estes resultados com os anteriores fica claro que realizando a DCT em blocos 8x8 a imagem resultante fica mais semelhante à original, o que permite distingui-las entre si.

Ao dividir a imagem em blocos com o intuito de aplicar a DCT conseguimos filtrar as frequências mais altas da imagem. Isto vai fazer com que a compressão seja mais facilitada uma vez que temos as zonas com mais frequência, podendo descartar as outras.

# Pergunta 7.3

* barn\_mountains.bmp

|  |  |  |
| --- | --- | --- |
| Canal | Com downsampling 4:2:0 | Com downsampling 4:2:2 |
| Y |  |  |
| Cb |  |  |
| Cr |  |  |

* logo.bmp

|  |  |  |
| --- | --- | --- |
| Canal | Com downsampling 4:2:0 | Com downsampling 4:2:2 |
| Y |  |  |
| Cb |  |  |
| Cr |  |  |

* peppers.bmp

|  |  |  |
| --- | --- | --- |
| Canal | Com downsampling 4:2:0 | Com downsampling 4:2:2 |
| Y |  |  |
| Cb |  |  |
| Cr |  |  |

Analisando as imagens anteriores, vemos que aqui volta a acontecer o que aconteceu na pergunta 7.1. As imagens não são distinguíveis pela sua DCT aplicada em blocos 64x64. Embora seja melhor do que aplicar à imagem inteira, uma vez que assim sempre se filtram as frequências mais altas nesses blocos. Na imagem inteira, temos apenas noção de onde se situam essas frequências na imagem e não dá para comprimir tão bem como sabendo em pequenos blocos.

Deste modo, percebemos que quanto menores forem os blocos, melhor será a compressão resultante.