



UNIVERSIDADE D  
COIMBRA

# Projeto de Multimédia

## Compressão de imagem

**Trabalho realizado por:**

- António Correia (2019216646)
- João Silva (2019216783)
- Pedro Martins (2019216826)

## Índice:

Introdução .....	3
Compressão inicial das imagens (Ex 1.4).....	4
Conversão para modelo de cor YCbCr (Ex 5.4) .....	6
Subamostragem (Ex 6.3) .....	7
Transformada de Cosseno Discreta (DCT) (Ex 7.1.3/7.2.3/7.3.2) .....	10
Quantização (Ex 8.3/8.4) .....	12
Codificação DPCM dos coeficientes DC (Ex 9.3) .....	14
Conclusões.....	15

## Introdução:

Ao longo deste projeto serão abordados pontos e fundamentos cruciais para a compressão de imagem, recorrendo ao codec JPEG para tal. Este método de compressão é classificado como destrutivo (ou lossy), uma vez que aquando da compressão de uma imagem o produto final contém perda de qualidade, sendo que esta redução pode ser limitada ao custo de uma menor taxa de compressão. Todo o código desenvolvido no âmbito deste trabalho será escrito na linguagem Python com o auxílio de várias bibliotecas, tais como: **matplotlib**, **numpy**. O objetivo final é implementar todos os métodos necessários para realizar a compressão e descompressão de uma imagem, seguindo os conceitos do JPEG.

Para tal, o projeto encontra-se separado em várias etapas, das quais se destacam:

- Compressão usando JPEG (Apenas para testar eficiência do algoritmo).
- Tratamento dos canais RGB da imagem.
- Padding da imagem (garantir que a mesma é 16x16 ou múltiplo).
- Conversão para o modelo de cor YCbCr.
- Sub-amostragem dos canais de cor.
- Transformada de Coseno Discreta
- Quantização
- Codificação DPCM dos coeficientes DC.
- Codificação e decodificação end-to-end.

## Compressão das imagens (Ex 1.4):

Primeiramente procedeu-se à compressão de um conjunto de imagens com características diferentes entre si, com recurso ao algoritmo de compressão destrutiva **JPEG** com diversos parâmetros de qualidade, entre os quais:

- Qualidade Alta: Valor de compressão 75%
- Qualidade Média: Valor de compressão 50%
- Qualidade Baixa: Valor de compressão 25%

Na tabela abaixo são apresentados os tamanhos iniciais e finais de cada uma das imagens para os vários tipos de qualidade acima referidos:

	Original	Alta	Média	Baixa
<b>Barn_mountains.bmp</b>	349 KB	34 KB	22 KB	14 KB
<b>Logo.bmp</b>	412 KB	10 KB	8 KB	7 KB
<b>Peppers.bmp</b>	577 KB	31 KB	21 KB	14 KB

**Tabela 1** – Tamanho das imagens pré e pós compressão.

De forma a ser possível interpretar melhor os resultados acima obtidos, iremos usar a taxa de compressão para cada tipo de qualidade diferente. Esta taxa assenta num valor percentual que representa o quão comprimido um ficheiro ficou, com base no seu tamanho inicial. Assim é mais fácil estabelecer um método de comparação entre as imagens e os tipos de qualidade. Para proceder ao seu cálculo, recorre-se à seguinte fórmula:

$$TaxaDeCompressão(\%) = \left(1 - \frac{TamanhoComprimido}{TamanhoOriginal}\right) \times 100$$

	Alta	Média	Baixa	Média
<b>Barn_mountains.bmp</b>	90.26%	93.69%	95.99%	93.13%
<b>Logo.bmp</b>	97.58%	98.06%	98.31%	97.98%
<b>Peppers.bmp</b>	94.63%	96.36%	97.57%	96.19%
<b>Média</b>	94.16%	96.04%	97.29%	

**Tabela 2** – Taxas de compressão e respetivas médias para as imagens

Após realizar a compressão para JPEG das imagens a analisar e calcular a respetiva taxa de compressão, é possível concluir dois pontos:

- De facto, confirma-se que, quanto menor for a qualidade da imagem final pós-compressão, menor vai ser o tamanho do ficheiro final, o que leva a uma maior taxa de compressão. Este facto pode ser avaliado através do maior valor percentual relativo à compressão de qualidade baixa (Ver Tabela 2).
- A imagem “Logo.bmp” apresenta a maior taxa de compressão, funcionando assim melhor aquando do processo de compressão. Isto deve-se ao facto de que esta imagem é relativamente simples, sem grande variância de cores contudo, o produto final vai estar mais danificado nos contornos relativos à mudança abrupta de cor. Ficheiros com maior complexidade e diferentes cores (Barn\_mountains.bmp) não apresentam estes danos numa escala tão elevada, o que também origina uma menor taxa de compressão.

Regra geral, quando aplicado o codec JPEG a imagem perde alguma da sua qualidade original (daí ser categorizado como um algoritmo lossy ou destrutivo). Esta perda é assentada à medida que se verifica uma diminuição na qualidade final da imagem (75%→50%→25%), traduzindo-se num ficheiro mais comprimido (com maior taxa de compressão).

## Conversão para modelo de cor YCbCr (Ex 5.4):

Foi realizada a conversão dos canais R, G e B das diferentes imagens presentes no dataset para o modelo YCbCr. Como foi possível observar nos canais anteriormente referidos, o modelo de cor RGB apresenta uma grande redundância de luminância e crominância nos seus 3 canais. Em comparação, o YCbCr descorrelaciona os canais de cor, e agrupa a luminância num só canal (Y), o que origina um aumento da compressibilidade.

Dentro do modelo de cor YCbCr é possível, à semelhança do RGB, separar o mesmo em três canais: Y, Cb e Cr. Como já foi mencionado previamente, o canal Y agrupa toda a luminância, o que vai resultar numa maior presença de informação. Os canais Cb e Cr vão assim possuir menor informação, o que os torna candidatos ideais para as próximas fases de compressão, diminuindo assim a perda de qualidade da imagem original.



Figuras 1 e 2 - Canais Y e Cb, respetivamente



Figura 3 – Canal Cr para a imagem Barn\_mountains.bmp

### Subamostragem (Ex 6.3):

A sub-amostragem representa uma técnica de compressão na qual se pretende baixar a resolução nos canais Cb e Cr (não no Y) e assim reduzir a taxa de amostragem para aumentar a taxa de compressão. Para isso iremos estudar dois tipos de downsampling:

- **4:2:2** - A taxa de amostragem nos canais mencionados (Cb e Cr) é reduzida para metade na direção horizontal.
- **4:2:0** - A taxa de amostragem nos canais mencionados (Cb e Cr) é reduzida para metade em ambas as direções (horizontal e vertical).

Assim, por exemplo, numa imagem com resolução 200x200 quando aplicado o downsample de 4:2:2 terá uma resolução final de 200x100. Com 4:2:0 o resultado será uma imagem com uma resolução de 100x100.

Para a taxa de amostragem 4:2:2 obteram-se os seguintes valores de compressão quando aplicada aos canais Cb e Cr:

	Cb Normal	Cb Downsample	Cr Normal	Cr Downsample
<b>Barn_mountains.bmp</b>	49 KB	30 KB	39 KB	24 KB
<b>Logo.bmp</b>	5 KB	4 KB	6 KB	4 KB
<b>Peppers.bmp</b>	76 KB	46 KB	72 KB	44 KB

**Tabela 3** – Tamanho ocupado pelos canais Cb e Cr sem e com downsample 4:2:2.

Para a taxa de amostragem 4:2:0 obteram-se os seguintes valores de compressão quando aplicada aos canais Cb e Cr:

	Cb Normal	Cb Downsample	Cr Normal	Cr Downsample
<b>Barn_mountains.bmp</b>	49 KB	15 KB	39 KB	12 KB
<b>Logo.bmp</b>	5 KB	2 KB	6 KB	3 KB
<b>Peppers.bmp</b>	76 KB	24 KB	72 KB	22 KB

**Tabela 4** – Tamanho ocupado pelos canais Cb e Cr sem e com downsample 4:2:0.

De seguida, ir-se-á analisar estes valores à luz da taxa de compressão, já previamente enunciada aquando da compressão inicial das imagens, de forma a perceber melhor o quão comprimido é o resultado final com os vários tipos de sub-amostragem estudados no âmbito deste capítulo. Na primeira tabela é representada a taxa de compressão e respetiva média para uma taxa de amostragem 4:2:2:

	Cb Downsample	Cr Downsample
<b>Barn_mountains.bmp</b>	38.78%	38.46%
<b>Logo.bmp</b>	20%	33.33%
<b>Peppers.bmp</b>	39.47%	38.89%
<b>Média</b>	32.75%	36.89%

**Tabela 5** – Taxa de compressão das imagens com downsample 4:2:2.



E agora os valores da taxa de compressão e as respetivas médias para uma taxa de amostragem 4:2:0:

	<b>Cb Downsample</b>	<b>Cr Downsample</b>
<b>Barn_mountains.bmp</b>	69.39%	69.23%
<b>Logo.bmp</b>	60%	50%
<b>Peppers.bmp</b>	68.42%	69.44%
<b>Média</b>	65.94%	62.89%

**Tabela 6** – Taxa de compressão das imagens com downsample 4:2:0.

Após o cálculo de todas as taxas de compressão, é notório que estes valores são bastante superiores na taxa de amostragem 4:2:0. Isto deve-se ao facto de que o produto final é uma imagem com resolução inferior à do 4:2:2, contribuindo assim para um produto final com menor tamanho e mais comprimido.

### **Transformada de Cosseno Discreta (DCT) (Ex 7.1.3/7.2.3/7.3.2):**

Na DCT a maior parte da energia está concentrada nas baixas frequências espaciais. Assim, apenas uma pequena parte das células possui esta energia, o que potencia um aumento da taxa de compressão. Isto é visível quando comparado a outras transformadas como, por exemplo, a Transformada de Fourier Discreta (DFT). Na Figura 4 é notória a aplicação da DCT num channel (Cb com downsample) de uma imagem do dataset. No canto superior esquerdo estão presentes as baixas frequências e, assim, a maior parte da energia.

Este passo (transformação DCT) é essencial para potenciar o aumento da taxa de compressão, visto que, devido ao número de repetições encontradas, é passível de ser aplicado algoritmos de compressão, tais como Run Length Encoding ou Huffman Encoding, o que leva a um ficheiro final mais comprimido.

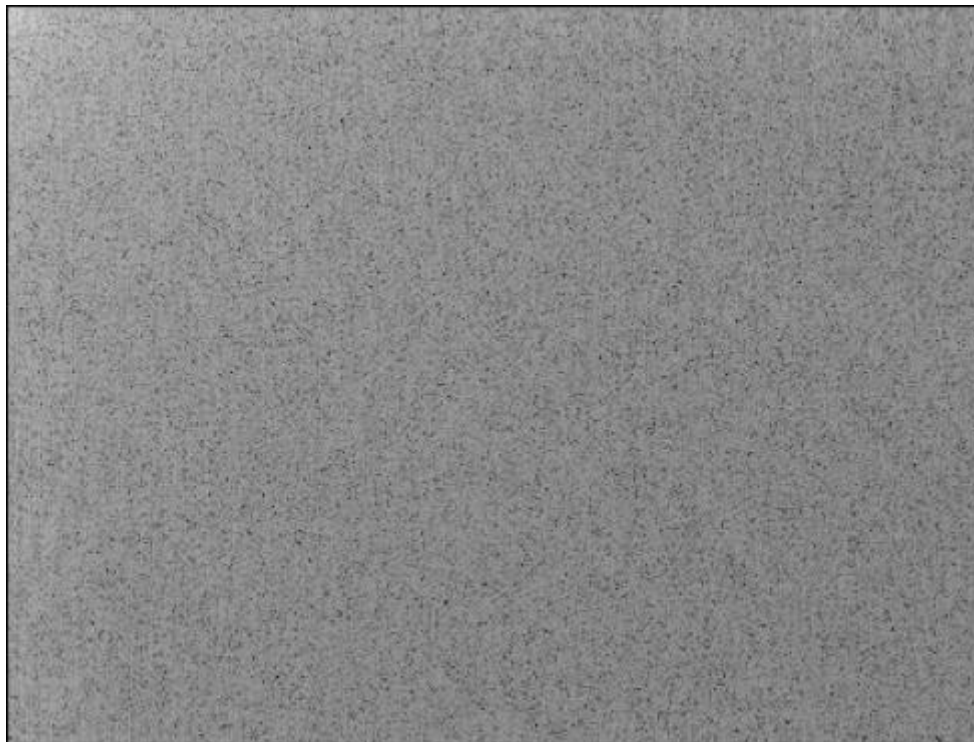
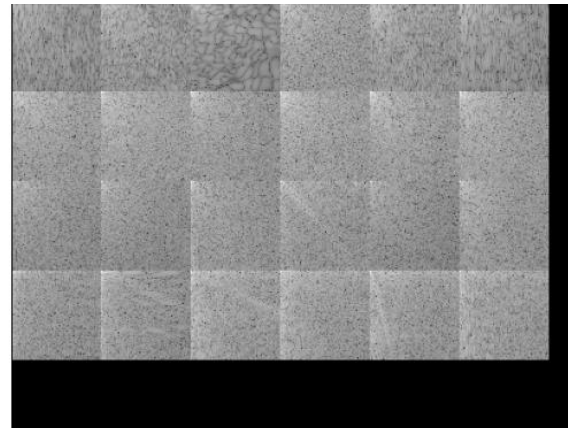
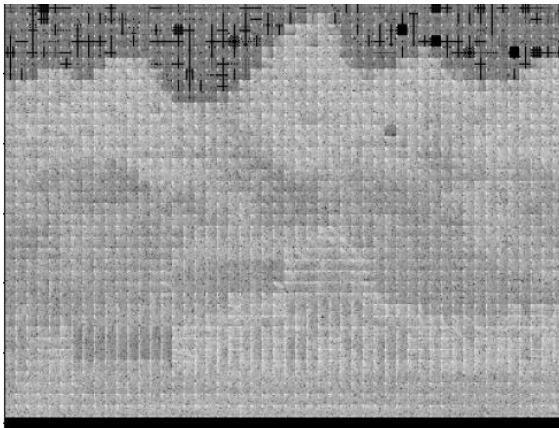


Figura 4 – Canal Y<sub>d</sub> da imagem barn\_mountains.bmp após aplicada a DCT



Figuras 5 e 6 - Canal Y\_d da imagem barn\_mountains.bmp para blocos 8x8 e 64x64, respetivamente

Analisando as **Figuras 5 e 6** é possível concluir que aquela onde a imagem é mais perceptível é a 5. Contudo a escolha do tamanho dos blocos é bastante importante, tanto em termos de qualidade da imagem final, como da taxa de compressão que a mesma atinge.

Foi reparado aquando da comparação do produto final de cada imagem para um DCT de 8x8 blocos e 64x64 blocos que no último a qualidade de imagem era inferior e que o tamanho final era superior face aos blocos 8x8. Esta conclusão levou-nos a explorar outras combinações de tamanhos de blocos, notoriamente 1x1, 2x2 e 4x4. Para estes, o tempo de execução é bastante superior (maior quantidade de blocos a aplicar DCT, quantização, etc.) e o resultado é uma imagem com menor qualidade e, sobretudo para os blocos 1x1, uma taxa de compressão quase nula.

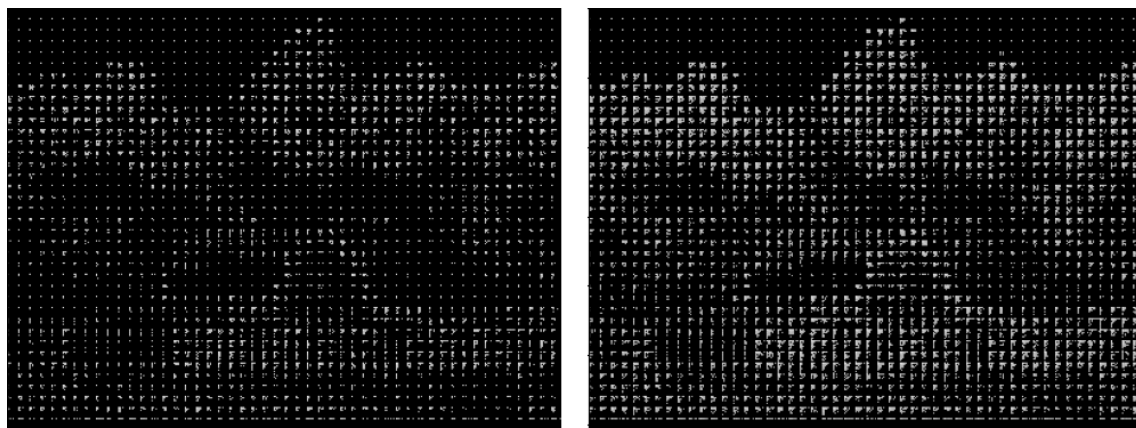
Assim, com base nos testes acima efetuados, concluímos que os blocos 8x8 são uma boa divisão à qual aplicar a DCT e, futuramente, a quantização e codificação DPCM.

### Quantização (Ex 8.3/8.4):

O próximo passo de compressão no JPEG é a quantização. Nesta parte é decidida a qualidade (1-100) de imagem a aplicar aquando da compressão. Abaixo segue-se a fórmula aplicada para calcular o valor quantizado para cada ponto da matriz DCT, à qual se aplica este processo:

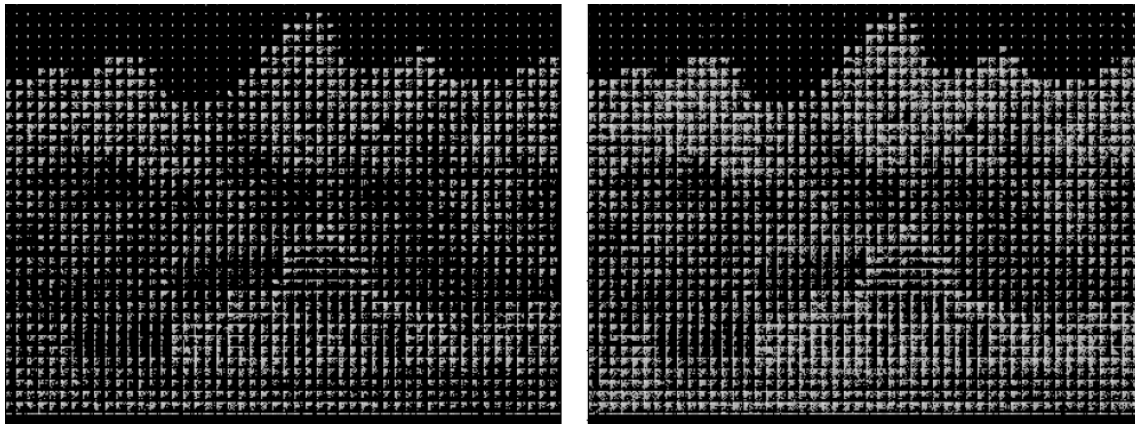
$$\text{Valor quantizado}(i,j) = \frac{DCT(i,j)}{\text{Matriz Quantização}(i,j)}, \text{ arredondado para inteiro}$$

Aqui entra o fator de qualidade, uma vez que este valor vai alterar cada ponto da matriz de quantização. Quanto maior ele for, menor é cada coordenada desta matriz, reduzindo assim a destruição causada por este método. Caso o fator seja pequeno, cada coordenada da matriz terá um valor mais elevado, o que originará uma maior destruição (valores quantizados iguais a 0, ou seja, impossíveis de reconstruir) e, por conseguinte, uma qualidade final de imagem consideravelmente inferior. O que leva à conclusão de que quanto maior for este valor, menor é a taxa de compressão e maior é a qualidade da imagem final.



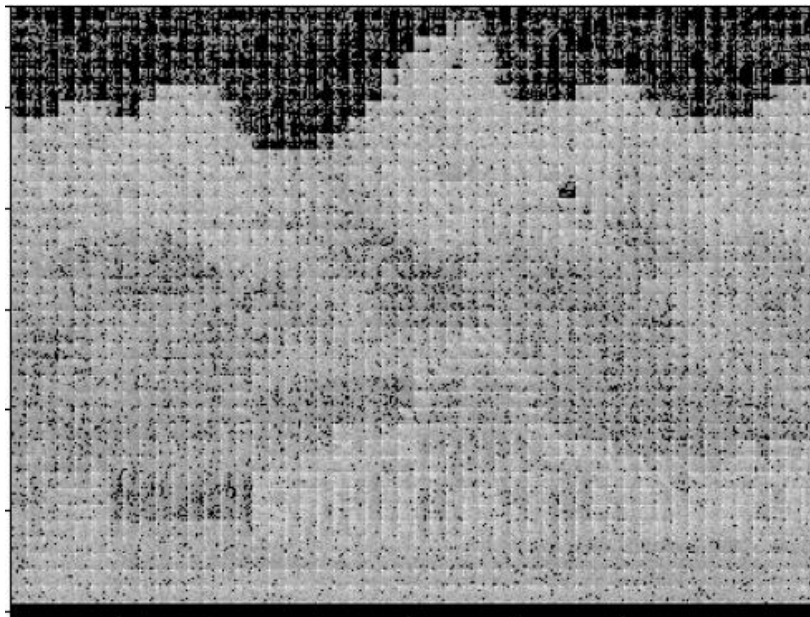
Canal Y\_d da imagem barn\_mountains.bmp após quantização com qualidade 10 e 25, respetivamente

Na comparação acima é notório o ganho de qualidade quando passamos de um fator de 10 para 25, o que se vai correlacionar com uma imagem final com menor taxa de compressão.



Canal Y\_d da imagem barn\_mountains.bmp após quantização com qualidade 50 e 75, respetivamente

De 50 para 75 a qualidade vai cada vez aumentando mais, visto que cada vez é possível observar melhor a imagem esboçada. Contudo o potencial de compressão é cada vez menor.



**Figura 11** - Canal Y\_d da imagem barn\_mountains.bmp após quantização com qualidade 100

Finalmente, ao chegarmos ao fator de qualidade 100, é possível observar ainda melhor a imagem, contudo o potencial de compressão é muito baixo, visto que muito pouco se alterou ao aplicar a quantização quando comparado com a DCT deste canal para os blocos 8x8 (**Figura 5**). Isto deve-se ao facto de que apenas se aplicou compressão entrópica (todos os valores da matriz de quantização são 1) ou seja, arredondou-se cada valor quantizado.

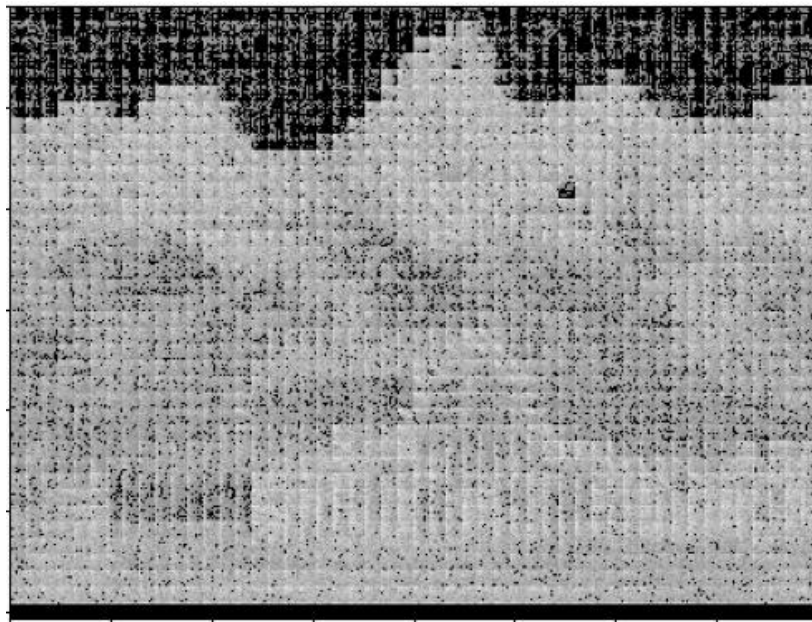


### Codificação DPCM dos coeficientes DC (Ex 9.3):

A codificação DPCM dos coeficientes DC representa a última parte do algoritmo JPEG desenvolvido no âmbito deste projeto e é aplicado à matriz resultante da aplicação da quantização no passo anterior.

Representa uma compressão entrópica, o que leva a imagem obtida a ser bastante semelhante àquela representada na quantização do canal Y com qualidade 100. As únicas diferenças entre estas imagens são os coeficientes DC, o que, mais uma vez, justifica as parecenças entre as mesmas.

Apesar de se encontrar fora do âmbito do objetivo para este trabalho, é importante de referir que visto que se encontram valores bastantes semelhantes após esta etapa, poder-se-iam aplicar algoritmos de compressão lossless (sem perda de qualidade de imagem) tais como Huffman Encoding ou codificação aritmética, de forma a aumentar o potencial de compressão da imagem final.



**Figura 12** - Canal Y\_d da imagem barn\_mountains.bmp após codificação dos coeficientes DC (qualidade 100)

### Conclusões (Ex 10.2):

Na tabelas abaixo encontram-se representados os valores das métricas de distorção MSE, RMSE, SNR e PSNR para cada imagem presente no dataset, de forma a perceber quais os valores apresentados para os diferentes fatores de qualidade.

	10	25	50	75	100
MSE	745.79	428.12	288.89	179.53	33.81
RMSE	27.31	20.69	17.0	13.4	5.81
SNR	18.46	20.87	22.58	24.64	31.9
PSNR	19.4	21.82	23.52	25.59	32.84

**Tabela 7** – Métricas de distorção para barn\_mountains.bmp e os diversos fatores de qualidade.

	10	25	50	75	100
MSE	349.22	167.62	108.74	74.17	17.42
RMSE	18.69	12.95	10.43	8.61	4.17
SNR	19.49	22.68	24.56	26.22	32.52
PSNR	22.7	25.89	27.77	29.43	35.72

**Tabela 8** – Métricas de distorção para peppers.bmp e os diversos fatores de qualidade.

	10	25	50	75	100
MSE	194.23	85.41	60.76	38.4	14.47
RMSE	13.94	9.24	7.8	6.2	3.8
SNR	28.45	32.02	33.5	35.49	39.73
PSNR	25.25	28.82	30.29	32.29	36.53

**Tabela 9** – Métricas de distorção para logo.bmp e os diversos fatores de qualidade.

Ao analisar as seguintes tabelas é notório que, à medida que o fator de qualidade aumenta, as métricas MSE e RMSE diminuem, o que se encontra dentro do esperado, uma vez que estes valores indicam a diferença média entre os pixels da imagem original e comprimida, logo com qualidade baixa esta diferença é muito maior (qualidade final mais degradada de forma a aumentar a taxa de compressão).

Por outro lado, SNR e PSNR são diretamente proporcionais à qualidade (quanto maior o fator, maior o valor da métrica), o que também se encontra dentro do expectável para esta experiência, visto que para o seu cálculo divide-se pelo valor de MSE, logo quanto maior MSE for, menor as métricas em questão serão.

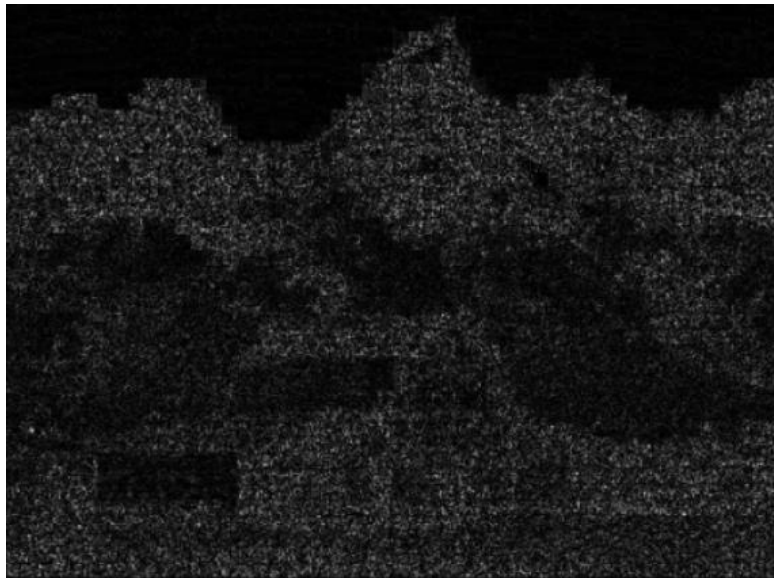
De seguida é possível observar uma comparação para a imagem peppers original e comprimida com fator de qualidade 75:



**Figura 13** – Peppers antes e após aplicação do método de compressão, respetivamente.



Na **Figura 13** é possível notar alguma compressão após aplicação do JPEG com fator de qualidade de 75. Caso este valor diminua, as diferenças serão cada vez mais acentuadas, o que se justifica pelo facto de que os métodos aplicados têm uma componente destrutiva a si associados. Para terminar, segue-se a imagem de erro para barn\_mountains (**Figura 14**), calculada através da subtração da imagem final comprimida e da imagem .bmp inicial:



**Figura 14** – Diferença entre a imagem final e inicial, para o canal Y de barn\_mountains.bmp e fator de qualidade 75.