



FACULDADE DE LETRAS
UNIVERSIDADE D
COIMBRA

Trabalho Prático 1

Compressão de Imagem

Trabalho realizado por:

Carlos Alfredo de Oliveira Matos, nº 2020245868

Francisco Loureiro Cruz Pessoa Macedo, nº 2020223771

Mariana Santos Magueijo, nº 2020246886

Introdução

Realizámos este projeto no âmbito da cadeira de Multimédia. Este visa o entendimento e resolução de questões fundamentais de compressão de imagem, em particular, através do codec JPEG.

Exercício 1

Para este exercício, optámos por criar uma função que nos permitisse realizar a compressão de cada imagem com as qualidades pedidas (alta, média e baixa). Deste modo, a função de compressão recebe o relative path, o nome da imagem e a extensão em que a mesma se encontra.

<i>logo</i>	Qualidade	Tamanho (em bytes)	Taxa de compressão (%)	Resolução
	100	421556	0	500x281
	75	7838	98.14	500x281
	50	6328	98.5	500x281
	25	5188	98.77	500x281

<i>barn_moutains</i>	Qualidade	Tamanho (em bytes)	Taxa de compressão	Resolução
	100	356456	0	400x297
	75	28477	92.01	400x297
	50	18666	94.76	400x297
	25	11976	96.64	400x297

<i>peppers</i>	Qualidade	Tamanho (em bytes)	Taxa de compressão	Resolução
	100	589880	0	512x384
	75	23509	96.01	512x384
	50	15671	97.34	512x384
	25	10553	98.21	512x384

Conclusões

Nas imagens comprimidas com qualidade alta (75) não se nota grande diferença quando comparado com as originais. No entanto, no conjunto de imagens *logo*, são perceptíveis algumas diferenças devido à mudança abrupta de cores.

Da qualidade média (50) para a qualidade alta, já se verificam algumas mudanças nas imagens. Em contraste com a qualidade alta, desta vez não se nota grande diferença na *logo*.

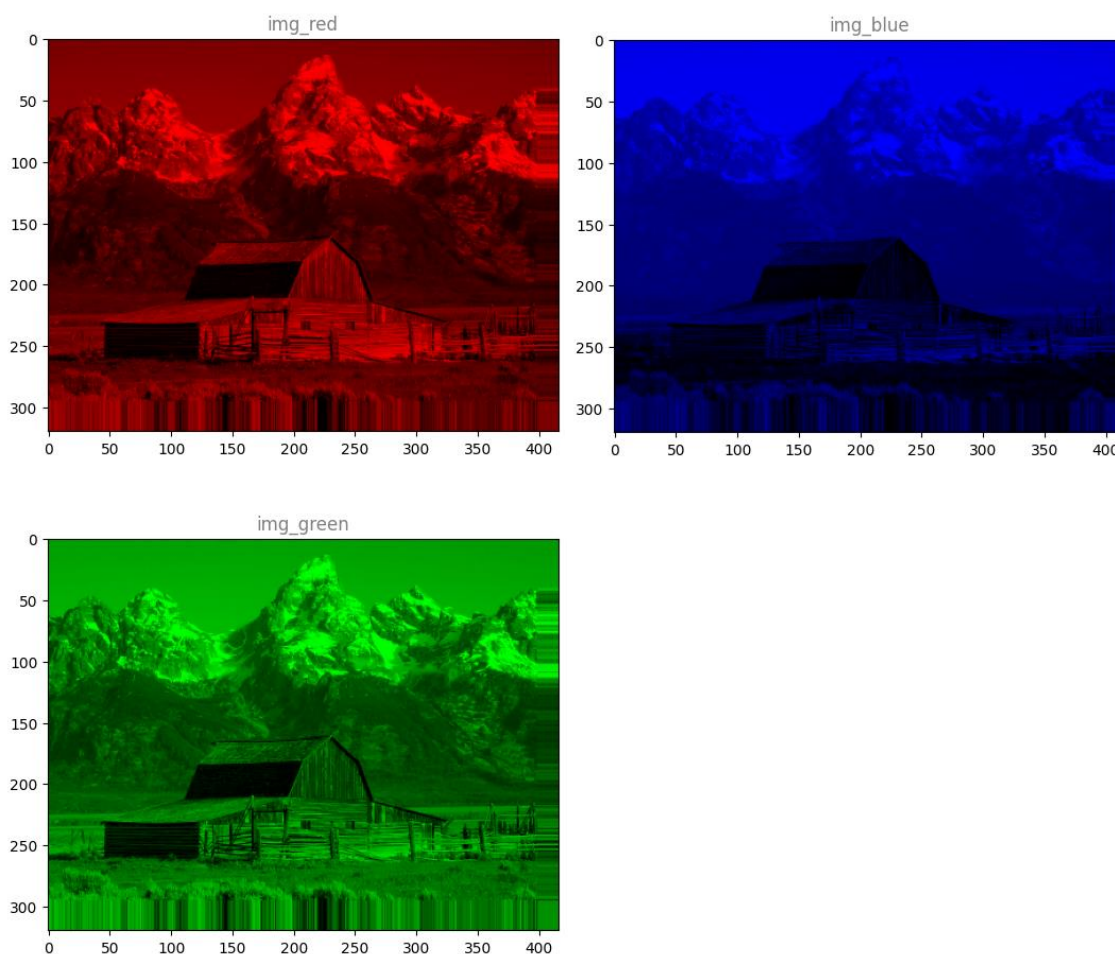
Por fim, nas imagens comprimidas com qualidade baixa (25), podemos identificar algumas imperfeições gráficas nos tons de azul na *barn_mountains* e nos tons de vermelho da *peppers* e algumas incongruências nas transições de cores na *logo*.

Exercício 3

No exercício 3 é onde realizamos as leituras e separações necessárias para as etapas do JPEG.

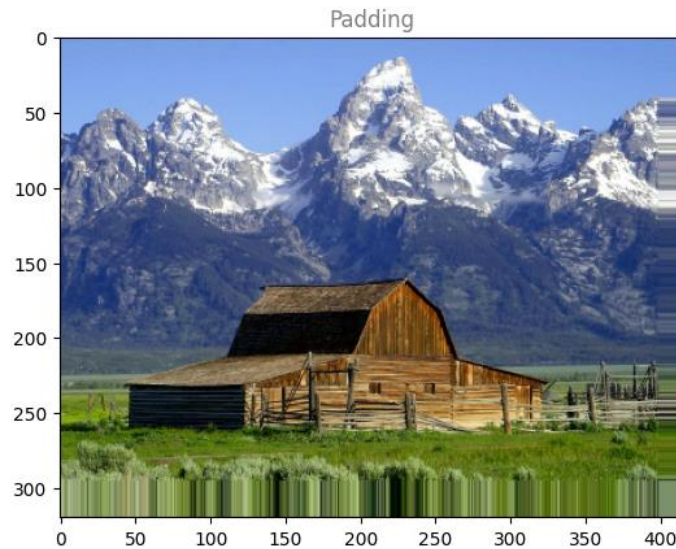
O colormap é essencialmente uma tabela que mapeia valores numéricos para cores. Normalmente, os valores mais baixos são associados a cores mais escuras, enquanto os valores mais altos a cores mais claras. Esta escala de cores é usada para tornar mais fácil a identificação de padrões ou diferenças num conjunto de dados.

NOTA: nas imagens apresentadas abaixo está aplicado o padding referente ao exercício 4



Exercício 4

No exercício 4 fazemos um pré-processamento da imagem através do *padding*. O *padding* consiste na adição de pixels extras em torno da borda de uma imagem para aumentar o seu tamanho ou para alinhá-la num tamanho específico. No nosso caso, queríamos que as linhas e colunas fossem múltiplas de 32. Para isso, implementámos a função “padding” que recebe como argumento a imagem principal e começa por verificar se a sua dimensão é múltipla de 32. Caso não seja, adiciona-lhe as linhas e/ou colunas em falta.

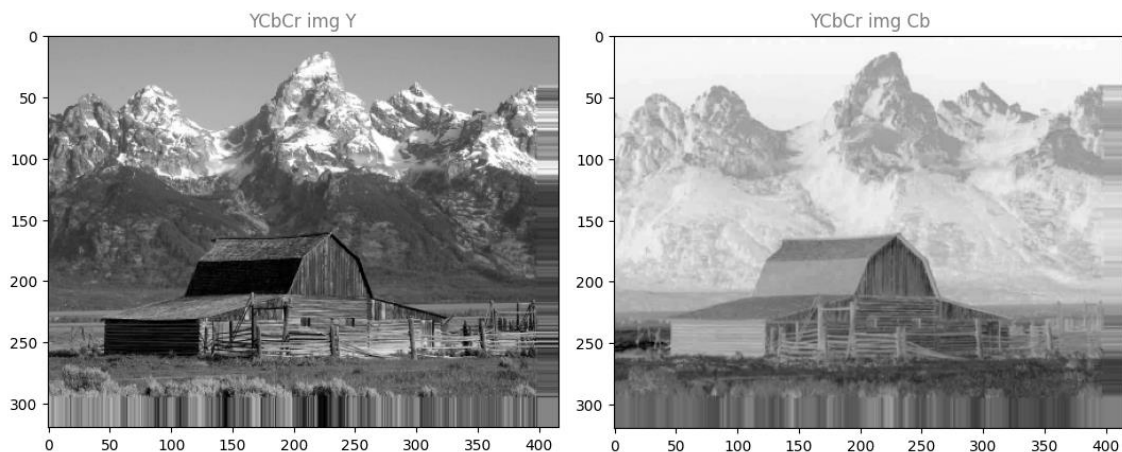


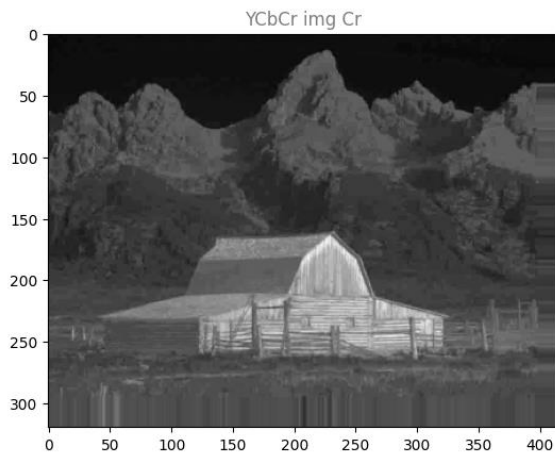
Exercício 5

Nesta fase, fazemos a conversão para o modelo cor YCbCr. Este modelo é usado para separar as informações de cor e brilho para permitir a compressão de forma mais eficiente e reduzir a redundância. A componente Y representa o brilho da imagem, coisa que não existe no modelo RGB, enquanto as componentes Cb e Cr representam as informações de cor azul e vermelho, respectivamente. Este modelo reduz a redundância nos dados.

Conclusões

Sendo que no canal Cb é guardado a informação da cor azul, conseguimos visualizar que esses tons são apresentados em tons mais claros do que os outros, visível, por exemplo, no céu.





Exercício 6

A sub-amostragem dos canais Y, Cb e Cr é uma técnica de processamento de imagem que envolve a redução da resolução dos canais de cor numa imagem para diminuir a quantidade de dados necessários para armazenar ou transmitir uma imagem.

Na sub-amostragem, os canais de cor são amostrados numa taxa inferior à do canal de luminância (Y), que contém a maioria das informações visuais de uma imagem. A taxa de sub-amostragem geralmente é expressa como uma proporção em relação à taxa de amostragem do canal Y.

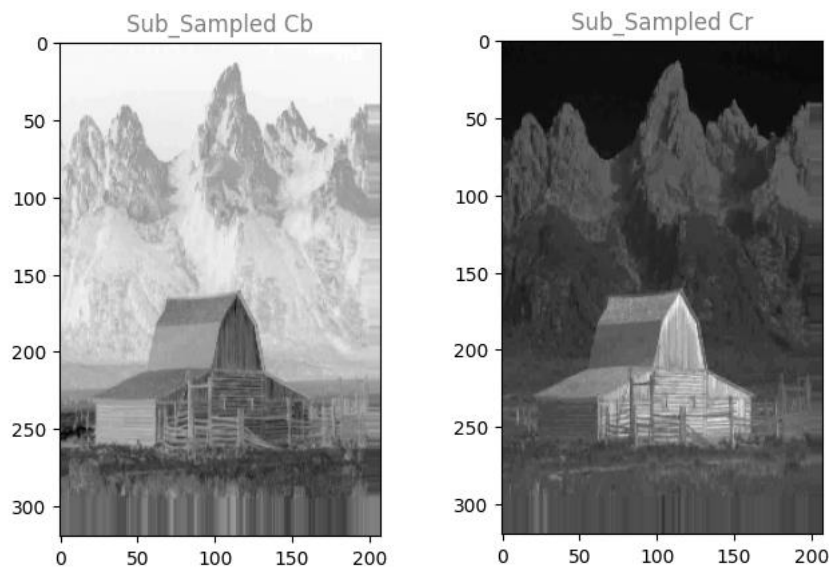
No caso de compressão 4:2:0, o canal Cb e o canal Cr são amostrados numa taxa $1/2$ em relação ao canal Y, ou seja, para cada quatro amostras de Y, existem apenas duas amostras de Cb e duas amostras de Cr, o que faz com que as taxas de amostragem dos canais Cb e Cr sejam reduzidas para metade na horizontal e na vertical. No caso de compressão 4:2:2, significa que, para cada 4 pixels de luminância amostrados, há 2 pixels de crominância amostrados na subamostragem horizontal. Na vertical, a subamostragem é 1:1, o que significa que cada linha de pixels tem a mesma resolução vertical.

Resultados

Esta técnica resulta numa perda de qualidade da imagem, pois as informações de cor são menos precisas e detalhadas do que no caso de não haver sub-amostragem.

Com a compressão 4:2:2, a qualidade da imagem é ligeiramente reduzida em comparação com a imagem original sem compressão.

A compressão 4:2:0 resulta numa maior redução de dados em comparação com a compressão 4:2:2, o que significa que a qualidade da imagem pode ser mais comprometida.



Exercício 7

Utilizámos a transformada discreta de cosseno para comprimir os canais Y, Cb e Cr em “blocos” de 1x1 (default), 8x8 e 64x64, conhecidos como data sets DCT.

Ao aplicarmos a DCT a um único bloco, em blocos de 8x8 e em blocos de 64x64, concluímos que a imagem é bastante perceptível, ao contrário das outras duas.

Deste modo, conseguimos perceber a razão pela qual o JPEG aplica a DCT em blocos 8x8

Exercício 8

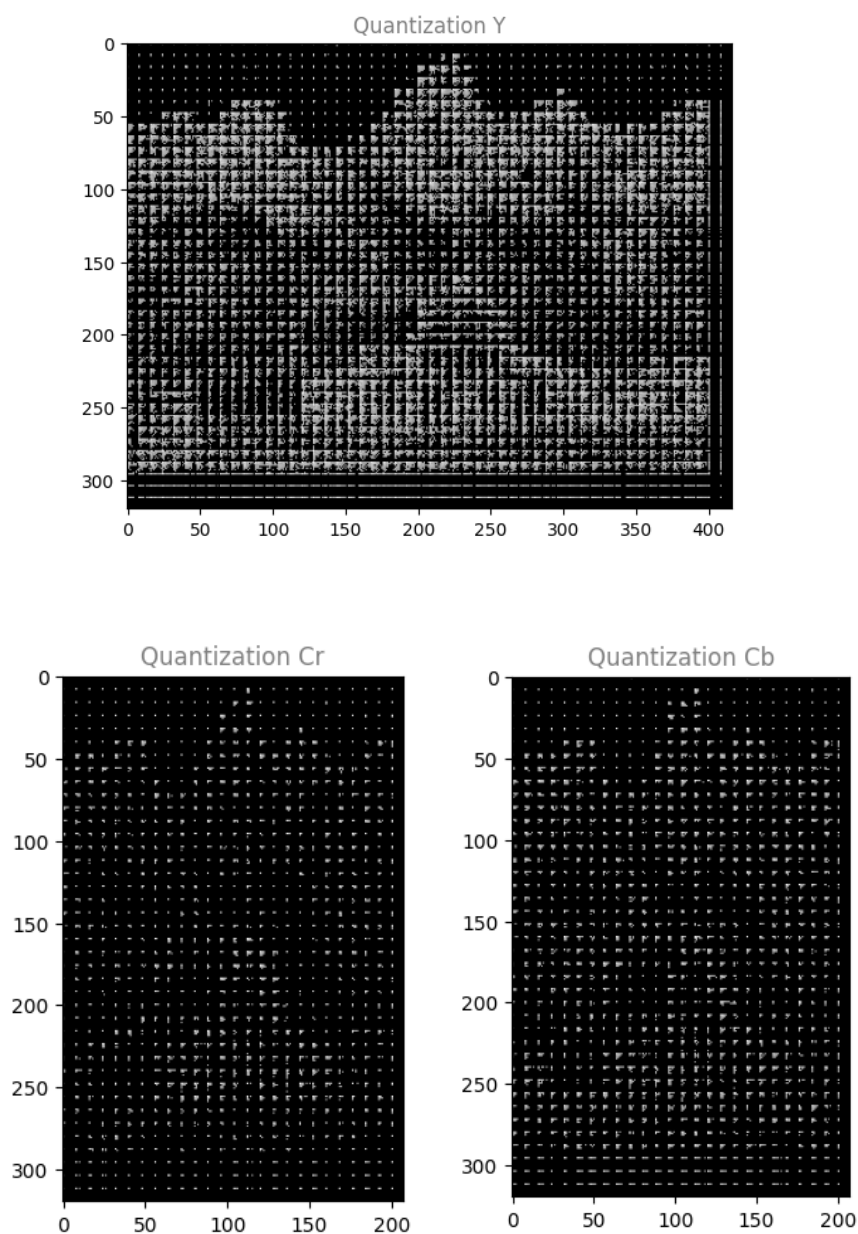
Quantização, em termos matemáticos e de processamento de sinais, é o processo de mapeamento de valores de um dado conjunto de entrada (geralmente, de grandes dimensões), tendo como saída um subconjunto com dimensões substancialmente inferiores.

Conclusões

Ao diminuirmos o fator de qualidade, os valores da matriz de quantização serão maiores e os resultados da divisão do bloco pela matriz menores.

Concluindo, quanto menor for a precisão da quantização, maior será a taxa de compressão, mas também maior será a perda de qualidade.

Semelhante à DCT, a quantização guarda o valor mais característico do bloco em que é aplicada, que se encontra no canto superior esquerdo de cada bloco. Isto aumenta ainda mais o potencial de compressão do que só aplicando a DCT.

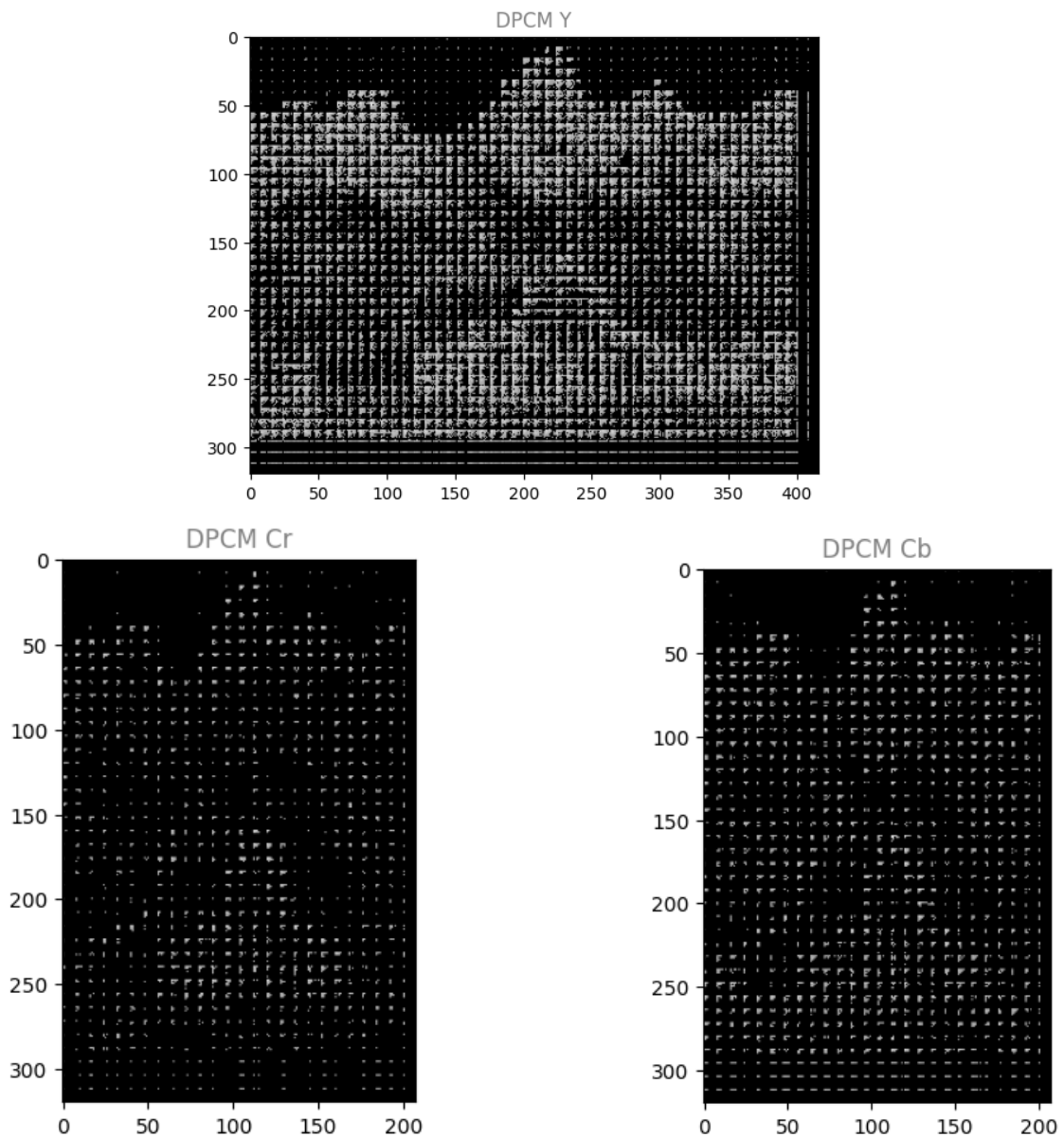


Exercício 9

DPCM é um método de codificação de sinais digitais que explora a redundância estatística nos sinais para obter uma compressão de dados mais eficiente. Os coeficientes DC são os componentes de baixa frequência da DCT de uma imagem que têm informações sobre o brilho médio da imagem.

Conclusões

Analisando os resultados obtidos, foi possível verificar que em zonas da imagem com grandes variâncias de cor, existe uma elevada concentração dos coeficientes nos *plots* devolvidos.



Exercício 10

MSE

É uma métrica usada para avaliar a qualidade da imagem após a codificação e decodificação.

O MSE é calculado como a média do quadrado das diferenças entre os valores de pixel da imagem original e os valores de pixel da imagem codificada e decodificada. Quanto menor o valor do MSE, menor é a diferença entre as duas imagens e melhor é a qualidade da imagem após a codificação e decodificação.

RMSE

O RMSE é calculado como a raiz quadrada da média do quadrado das diferenças entre os valores de pixel da imagem original e os valores de pixel da imagem codificada e decodificada. Como a raiz quadrada é usada.

Assim como o MSE, quanto menor o valor do RMSE, menor é a diferença entre as duas imagens e melhor é a qualidade da imagem após a codificação e decodificação.

SNR

O SNR é definido como a razão entre a potência do sinal (que é a informação útil da imagem) e a potência do ruído (que é qualquer informação indesejada adicionada durante o processo de codificação e decodificação da imagem). Quanto maior a relação sinal-ruído, melhor é a qualidade da imagem.

PSNR

O PSNR é calculado como a relação entre a energia máxima do sinal (ou seja, o valor máximo possível dos valores de pixel na imagem original) e o erro médio quadrático (MSE) entre a imagem original e a imagem codificada e decodificada. Quanto maior o valor do PSNR, maior é a qualidade da imagem.

Conclusão

Através destas métricas é mais fácil e correto fazer uma análise das taxas de compressão e fazer as comparações entre imagens.

	10	25	50	75	100
<i>MSE</i>	712,37344	403,45381	267,08829	160,16191	24,36879
<i>RMSE</i>	26,69032	20,08616	16,34283	12,65551	4,93647
<i>SNR</i>	18,65919	21,12834	22,91973	25,14068	33,31793
<i>PSNR</i>	19,60373	22,07286	23,86425	26,08521	34,26246