

Avaliação do impacto da ordem da DCT em compressão de imagens do tipo JPEG

Otávio F. Jacobi^a, Thiago L. T. Silveira^b, Fábio M. Bayer^c

^a*Apresentador, Acadêmico do Curso de Engenharia Elétrica, Bolsista FIT/UFSM*

^b*Co-autor, Acadêmico do Curso de Ciência da Computação, UFSM*

^c*Orientador, Departamento de Estatística e LACESM, UFSM*

1. Introdução

Dentro das técnicas de processamento de sinais e técnicas de compressão as transformadas discretas desempenham papel fundamental (Britanak et al., 2007). No contexto de compressão de imagens, a transformada discreta do cosseno (DCT) (Rao e Yip, 1990) é uma ferramenta pivotal. Este fato se deve pela conhecida aproximação da DCT para a transformada de Karhunen-Loève (KLT). A KLT tem a distinção de ser ótima em termos de compactação de energia.

A utilização da DCT em processamento de imagens possui uma conveniente formalização matricial (Cintra e Bayer, 2011). A matriz da DCT de ordem k , \mathbf{C}_k , pode ser definida pelos seguintes (i,j) -ésimos elementos

$$c_{i,j} = a_{i-1} \cos \left(\frac{\pi(i-1)(2(j-1)+1)}{2k} \right), \text{ para } i,j = 1, 2, 3, \dots k.$$

Em particular, a DCT de comprimento 8 tem ampla utilização em compressão de imagens, sendo adotada em diversos padrões de compressão, como JPEG (Pennebaker e Mitchell, 1992) e MPEG-1 (Roma e Sousa, 2007). A compressão de imagens digitais visa desconsiderar dados irrelevantes ou correlacionados dentro da imagem. Nestes casos, a perda de qualidade da imagem é inevitável. Porém, ao comprimir uma imagem, é possível armazenar uma menor quantidade de informação mantendo qualidade similar à original.

Contudo, o recente padrão de compressão de vídeo HEVC (Pourazad et al., 2012) usa transformadas de ordens $k = 4, 8, 16$ e 32 . A utilização da DCT de diferentes ordens neste importante método de codificação de vídeo motivou o presente trabalho.

2. Objetivos

Este trabalho tem como objetivo avaliar o desempenho da aplicação da DCT de ordens 4, 8, 16 e 32 na compressão de imagens digitais. As medidas avaliadas são o custo computacional e a qualidade das imagens comprimidas com a DCT sob as diferentes abordagens.

3. Metodologia

A compressão de imagens do tipo JPEG requer que a imagem de entrada seja sub-dividida em blocos de mesma dimensão k . Considerando que N seja a dimensão da imagem (usualmente $N = 512$), ela será sub-dividida em s^2 blocos, em que $s = \frac{N}{k}$ e k é a ordem da matriz DCT. Note que cada bloco \mathbf{B} resultante desta divisão terá, também, dimensão k . Para cada um dos blocos \mathbf{B} , aplicamos a DCT direta:

$$\mathbf{A} = \mathbf{C}_k \times \mathbf{B} \times \mathbf{C}_k^\top,$$

em que \mathbf{C}_k é a matriz da DCT e \mathbf{C}_k^\top é a sua transposta. Pela ortogonalidade da DCT nota-se que $\mathbf{C}^\top = \mathbf{C}^{-1}$.

A DCT tem a propriedade de compactar a energia do sinal nos primeiros coeficientes espectrais, permitindo desconsiderar ou anular os demais coeficientes. Assim, é possível selecionar apenas os r primeiros coeficientes, seguindo a sequência zigzag padrão (ver Figura 1), da matriz \mathbf{A} e ignorar os restantes, atribuindo a eles o valor zero.

Após a etapa de truncamento dos coeficientes da DCT, aplica-se a cada bloco truncado $\hat{\mathbf{A}}$ a DCT bidimensional inversa:

$$\hat{\mathbf{B}} = \mathbf{C}_k^\top \times \hat{\mathbf{A}} \times \mathbf{C}_k.$$

A matriz $\hat{\mathbf{B}}$ é uma aproximação da original \mathbf{B} após ter passado pelo processo de compressão. A imagem comprimida é obtida pela recomposição de todos os s^2 blocos $\hat{\mathbf{B}}$ na imagem de dimensão $N \times N$.

4. Resultados e discussão

Considerando diferentes ordens da DCT na compressão de imagens podemos ter diferentes níveis de qualidade das imagens comprimidas e também diferente custos computacionais dos algoritmos associados. O objetivo aqui é avaliar esses impactos. As implementações para essa aplicação foram feitas utilizando a linguagem R (R Core Team, 2013). A imagem utilizada foi a imagem padrão Lena, 512×512 pixels, em 256 tons de cinzas, obtida de USC-SIPI (2011).

Considerando o algoritmo de Chen et al. (1977), por meio das expressões das complexidades aritméticas apresentadas na Seção 3.2, os resultados de complexidade computacional para cada uma das abordagens é apresentados na Tabela 1. Já os resultados das medidas de qualidade, considerando diferentes níveis de compressão, estão presentes na Figura 2.

Tabela 1: Comparação da complexidade aritmética para compressão de imagens com 512×512 pixels.

Tamanho de bloco	Número de blocos	Complexidade de cada bloco			Complexidade total		
		Mult	Adições	Total	Mult	Adições	Total
4×4	16384	64	48	112	1048576	786432	1835008
8×8	4096	416	256	672	1703936	1048576	2752512
16×16	1024	2368	1408	2776	2424832	1441792	3866624
32×32	256	12416	7424	19849	3178496	1900544	5079040

Pode-se verificar que quanto menor o tamanho dos blocos da imagem, menor a complexidade total associada. Entretanto, verifica-se que a qualidade das imagens considerando blocos menores também é menor. Uma análise qualitativa das abordagens considerando diferentes tamanhos de sub-bloco é verificada na Figura 3. Percebe-se que a qualidade é proporcional ao tamanho de bloco k .

5. Conclusões

Foi possível verificar que existe um “trade-off” entre qualidade e complexidade da compressão. Para as mesmas proporções de coeficientes retidos, verifica-se que quanto maior o valor de k melhor a qualidade da imagem comprimida e também maior o custo computacional. Este é um estudo inicial e, com trabalhos futuros, pesquisas sobre aproximações da DCT de complexidade multiplicativa nula e baixa complexidade aditiva serão desenvolvidas.

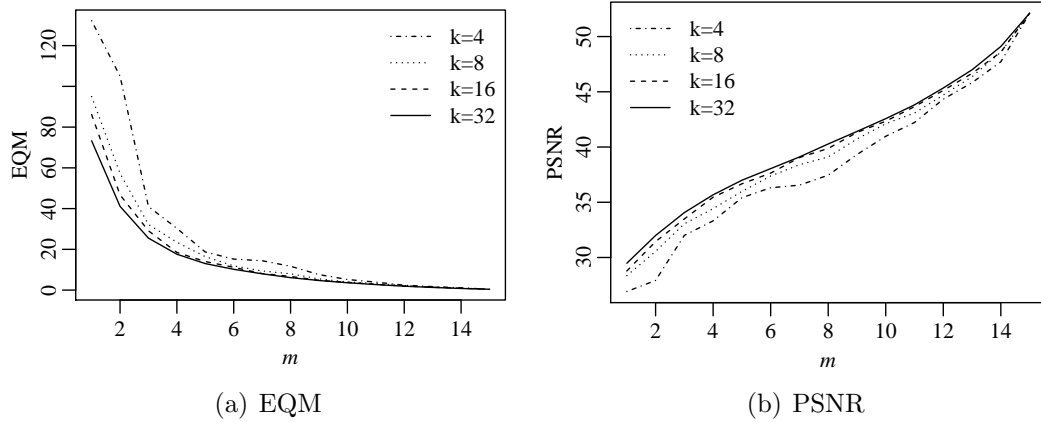


Figura 2: Medidas de qualidade da imagem comprimida sob diferentes tamanhos de bloco considerando proporções de coeficientes retidos iguais a $m/16$.



(a) $k = 4$ e $r = 2$ (b) $k = 8$ e $r = 8$ (c) $k = 16$ e $r = 32$ (d) $k = 32$ e $r = 128$

Figura 3: Imagem Lena comprimida utilizando transformadas com diferentes tamanho de bloco (k) e mesma proporção de coeficientes retidos ($2/16$).

Referências

- Britanak, V., Yip, P., Rao, K. R., 2007. Discrete Cosine and Sine Transforms. Academic Press.
- Chen, W. H., Smith, C., Fralick, S., Sep. 1977. A fast computational algorithm for the Discrete Cosine Transform. IEEE Trans. Commun. 25 (9), 1004–1009.
- Cintra, R. J., Bayer, F. M., Oct. 2011. A DCT approximation for image compression. IEEE Signal Processing Letters 18 (10), 579–582.
- Pennebaker, W. B., Mitchell, J. L., 1992. JPEG Still Image Data Compression Standard. Van Nostrand Reinhold, New York, NY.
- Pourazad, M. T., Doutre, C., Azimi, M., Nasiopoulos, P., Jul. 2012. HEVC: The new gold standard for video compression: How does HEVC compare with H.264/AVC? IEEE Consumer Electronics Magazine 1 (3), 36–46.
- R Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
URL <http://www.R-project.org/>
- Rao, K. R., Yip, P., 1990. Discrete Cosine Transform: Algorithms, Advantages, Applications. Academic Press, San Diego, CA.

Roma, N., Sousa, L., 2007. Efficient hybrid DCT-domain algorithm for video spatial downscaling. EURASIP Journal on Advances in Signal Processing 2007 (2), 30–30.

USC-SIPI, 2011. The USC-SIPI image database. <http://sipi.usc.edu/database/>, University of Southern California, Signal and Image Processing Institute.