Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, v. 7, n. 1, 2020.

Trabalho apresentado no XXXIX CNMAC, Uberlândia - MG, 2019.

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

A Transformada Discreta do Cosseno para a compressão de imagens JPEG: uma aplicação das bases ortogonais no espaço vetorial $M_{8\times 8}(\mathbb{R})$

Débora Lima Oliveira [1]
Universidade Federal do Sul e Sudeste do Pará
Cecilia Orellana Castro[2]
Universidade Federal do Sul e Sudeste do Pará

1 Introdução

A Transformada Discreta do Cosseno (DCT - Discret Cosine Transform) desenvolvida por Nasir Ahmed [1] em 1974 é uma ferramenta muito utilizada na compressão de imagens JPEG devido à eficiência em seus cálculos. Para utilizar a DCT numa imagem é necessário dividi-la em blocos 8×8, estas matrizes contêm informação vital da imagem e é sobre estes blocos que é aplicada a DCT. Em termos matemáticos, a DCT é um operador linear definido no espaço vetorial das matrizes quadradas de oito linhas e colunas, $M_{8\times 8}(\mathbb{R})$.

Em grande parte da literatura de processamento de imagens encontram-se *fórmulas* como uma definição da DCT sem explicar a álgebra linear envolvida, assim o objetivo deste trabalho é apresentar uma descrição do processo de compressão de imagens JPEG usando a DCT, mostrando principalmente, a matemática por trás do algoritmo mais utilizado na compressão de fotos. [2].

2 A álgebra Linear envolvida na DCT

Dentre as variantes da DCT, a mais utilizada em compressão de imagens JPEG é a DCT II. Baseado em [2], esta variante será explicada neste trabalho. A DCT II particiona a imagem em matrizes de dimensão 8x8. Seja $A = (a(i,j))$ com $i, j = 0, 1, \ldots, 7$ uma destas matrizes resultantes, a matriz A está representada na base canônica do espaço $M_{8\times 8}(\mathbb{R})$,

$$C = \{E_{nm} : E_{n,m}(i,j) = 0 \;\; se(n,m) \neq (i,j) \;\; e \;\; E_{n,m}(i,j) = 1 \;\; se \;\; (n,m) = (i,j)\}. \quad (1)$$

A DCT II considera uma outra base B ortogonal para o espaço $M_{8\times 8}(\mathbb{R})$ de dimensão 64, onde para cada $n, m = 0, 1, \ldots, 7$ tem-se uma matriz V_{nm} sendo parte da base:

[1] debora.oliveira@unifesspa.edu.br
[2] ceciliaoc@unifesspa.edu.br

$$B = \{V_{nm} : V_{nm}(i,j) = \cos(\frac{n\pi}{8}(i+\frac{1}{2}))\cos(\frac{m\pi}{8}(j+\frac{1}{2})) \;\; para \;\; i,j = 0,1,,7\}. \qquad (2)$$

Em termos de álgebra linear, aplicar a DCT II a uma matriz $A \in M_{8\times 8}(\mathbb{R})$ significa fazer uma mudança de coordenadas de A, da base canônica C dada em (1) para a base ortogonal B dada em (2). A vantagem de trabalhar com bases ortogonais deve-se a que o cálculo das coordenadas de um vetor é realizado apenas com produtos escalares entre os vetores da base, fato que simplifica as contas considerando a ortogonanilidade dos mesmos.

Para duas matrizes A_1 e $A_2 \in M_{8\times 8}(\mathbb{R})$ trabalha-se com o produto interno usual: $\langle A_1, A_2\rangle = tr(A_2^T A_1)$ e a norma $\|A_1\| = \sqrt{\langle A_1, A_1\rangle}$, com estas operações prova-se que:

$$\langle V_{nm}, V_{k\ell}\rangle = 0 \; se \; (n,m) \neq (k,\ell) \; e \; \langle V_{nm}, V_{k\ell}\rangle = \|v_n\|^2\|v_m\|^2 \; se \; (n,m) = (k,\ell) \qquad (3)$$

onde $v_n = (\cos(\frac{n\pi}{8})(0+\frac{1}{2})), \cos((\frac{n\pi}{8})(1+\frac{1}{2})), \ldots, \cos((\frac{n\pi}{8})(7+\frac{1}{2}))) \in \mathbb{R}^8$. Além disso, $\|v_n\|^2 = 4$ para $n > 0$ e $\|v_n\|^2 = 8$ para $n = 0$, este fato decorre do Teorema 2 descrito em [2].

Sejam $A(i,j) : i,j = 0,1,\ldots,7$ as coordenadas da matriz A na base B dada em (2), então: $A = \sum_{k=0}^{7}\sum_{\ell=0}^{7} A(k,\ell)V_{k\ell}$, multiplicando escalarmente ambos os lados por V_{ij} esta equação tem-se que $\langle A, V_{ij}\rangle = \sum_{k=0}^{7}\sum_{\ell=0}^{7} A(k,\ell)\langle V_{k\ell}, V_{ij}\rangle = A(i,j)\langle V_{ij}, V_{ij}\rangle$, assim isolando $A(i,j)$, tem-se: $A(i,j) = \frac{\langle A, V_{ij}\rangle}{\|V_{ij}\|^2} = \frac{tr(V_{ij}^T A)}{tr(V_{ij}^T V_{ij})}$ e, realizando os produtos internos obtem-se:

$$A(i,j) = \frac{1}{\|v_i\|^2\|v_j\|^2}\sum_{\ell=0}^{7}\sum_{k=0}^{7} a(k,\ell)\cos(\frac{(2k+1)i\pi}{16})\cos(\frac{(2\ell+1)j\pi}{16}) \qquad (4)$$

onde $a(k,\ell)$ são as coordenadas de A na base canônica C dada em (1), isto é a matriz que contém os píxeles da matriz original. A equação da em (4) é comumente apresentada nas referências bibliográficas sem mencionar sua origem como foi feito neste trabalho.

Observe que ao aplicarmos a DCT II em $a(i,j)$ obtemos $A(i,j)$. A vantagem de trabalhar com $A(i,j)$ é que esta matriz é mais esparsa, isto é, possui mais elementos nulos. Este fato reduz o tamanho da imagem original até a metade segundo os experimentos numéricos feitos pelas autoras usando as funções `dct2` e `idct2` do Octave e a matriz de quantização encontrada em [2]. Pretende-se continuar com o estudo de aspectos teóricos e computacionais com o objetivo de aprimorar o algoritmo implementado em Octave e realizar mais testes numéricos.

Referências

[1] Ahmed, N., Natarajan, T., Rao, K. R. Discrete Cosine Transform. IEEE Transactions on Computers C-23 (1): 90-93. 1974.

[2] M. F. Transformada Discreta de Cosseno: uma aplicação da Álgebra Linear na compressão de imagens do formato JPEG, 2013. http://conteudo.icmc.usp.br/pessoas/frasson/jpeg/jpeg.pdf

