Research of Lossy Image Compression Algorithm Based on Fractal Discrete Cosine Transform

Vladislav Butorov
Samara National Research University
Samara, Russia
twoeightnine@list.ru

Marina Chicheva
Image Processing Systems Institute of RAS - Branch of the FSRC "Crystallography and Photonics" RAS;
Samara National Research University
Samara, Russia
mchi@geosamara.ru

***Abstract*—Lossy image compression algorithm based on fractal discrete cosine transform is proposed in this paper. The created algorithm is compared to an algorithm based on two-dimensional discrete cosine transform. It is shown experimentally that the described algorithm brings less distortion concerning block structure in comparison with square blocks of two-dimensional discrete cosine transform. It is remarked that visual quality characteristics of both algorithms vary poorly for several values ranges of entropy of compressed image.**

Keywords*—lossy image compression, discrete cosine transform, canonical number system, fractal DCT

I. Introduction

For compression of images, lossy compression algorithms are widely used, since the losses introduced into the image can be invisible to the eyes and practically do not affect the visual quality. In such algorithms, compression is performed in the frequency domain, to obtain values in which discrete orthogonal transforms (DOTs) are used. Discrete cosine transform (DCT), namely, its two-dimensional variation, is widely used in the field of image processing. Since two-dimensional DCT is defined on a square region, the resulting artifacts in compression have a very noticeable mesh structure. To eliminate this effect, one can use the classical one-dimensional DCT applied to the sweep generated by some canonical number system (CNS) [1], or the fractal DCT (FDCT) defined on the fractal region generated by CNS [2]. In this paper, we study a lossy compression algorithm that uses various variations of the FDCT. The results of the algorithm are compared with a compression based on two-dimensional DCT.

II. The Theoretical Basis

A. Fractal DCT

This section provides brief theoretical information about the CNS in imaginary quadratic fields [3]-[6], k- fundamental domains, and FDCT [2].

Let $Q(\sqrt{d})$ is a quadratic field: $Q(\sqrt{d}) = \{z = a + bd; a, b \in Q\}$, d is an integer, free of squares. Then the field element $z \in Q(\sqrt{d})$ is called a whole algebraic field element if its norm and trace are integers

$$Norm(z) = (a + b\sqrt{d})(a - b\sqrt{d}),$$
$$Tr(z) = (a + b\sqrt{d}) + (a - b\sqrt{d}).$$

The whole algebraic element $\alpha \in Q(\sqrt{d})$ is the basis of the CNS in the ring of integer elements $Q(\sqrt{d})$, if any whole element of this field is uniquely representable in the form of a finite sum

$$z = \sum_{j=0}^{k} z_j\, \alpha^j, z_j \in N = \{0,1,..,|Norm(\alpha)| - 1\} \qquad (1)$$

CNS in the field $Q(\sqrt{d})$ is called a pair $\{\alpha, N\}$, k-fundamental domain G_k is the set of algebraic elements of the field $Q(\sqrt{d})$, created by k-membered sum of a formula (1),

$$G_k = \{\sum_{j=0}^{k-1} z_j\, \alpha^j, z_j \in N\}. \qquad (2)$$

Let $\Lambda\text{COS}_k(m,n) = cos\left(\frac{\pi \text{Im}\left(\bar{\alpha}^{k+1} n(x+\beta)\right)}{Norm(\alpha^k)\text{Im}(\alpha)}\right)$, where parameter β is set for the reason of orthogonality:

$$\sum_{x \in G_k} \Lambda\,\text{COS}_k(p,x) \cdot \Lambda\text{COS}_k(q,x) = 0, p \neq q,$$

For example, for $Norm(\alpha) = 2$ parameter β is calculated as

$$\beta = \frac{\alpha^{k+1} - 2\alpha^k + 1}{2(\alpha - 1)}.$$

Then FDCT over G_k is called a transformation

$$X(m) = \lambda(m) \sum_{n \in G_k} x\,(n) \Lambda\text{X}\text{O}\Sigma_k(m,n),$$

where $m \in D_k$, and $\lambda(m)$ is FDCT.

Reverse FDCT (RFDCT) is called

$$x(n) = \sum_{m \in D_k} \lambda\,(m) X(m) \Lambda\text{X}\text{O}\Sigma_k(m,n),$$

where $n \in G_k$, and $\lambda(m)$ is the normalizing coefficient of FDCT.

The normalizing coefficient of FDCT and RFDCT is equal and is calculated using the following equation:

$$\lambda(m) = \begin{cases} \sqrt{\frac{1}{Norm(\alpha^k)}}, 2m \equiv 0(mod\alpha^k) \\ \sqrt{\frac{2}{Norm(\alpha^k)}}, 2m \neq 0(mod\alpha^k) \end{cases}.$$

The field D_k is found algorithmically for the reasons of orthogonality of the basis functions:

$$\sum \Lambda\,\text{COS}_k(p,x) \cdot \Lambda\text{COS}_k(q,x) = 0;$$
$$x \in G_k; p, q \in D_k; p \neq q$$
$$\sum \Lambda\,\text{COS}_k(p,x) \cdot \Lambda\text{COS}_k(q,x) \neq 0;$$
$$x \in G_k; p, q \in D_k; p = q$$

The algorithm for calculating this region is described in [2].

B. Two-dimensional DCT

Two-dimensional DCT is called the transformation

$$X(m_1, m_2) = \sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1} x\,(n_1, n_2) cos\left(\frac{\pi k_1(n_1+0.5)}{N_1}\right) cos\left(\frac{\pi k_2(n_2+0.5)}{N_2}\right),$$

where $x(n_1, n_2)$ is a source signal (block of image brightness), N_i is the size of the i-th side of the block, and $X(m_1, m_2)$ is the resulting spectrum of the source signal.

Then the inverse two-dimensional DCT is called the transformation

$$x(n_1, n_2) = \sum_{m_1=0}^{N_1-1}\sum_{m_2=0}^{N_2-1} \lambda_1\,(m_1)\lambda_2(m_2)X(m_1, m_2) \times cos\left(\frac{\pi k_1(m_1+0.5)}{N_1}\right) cos\left(\frac{\pi k_2(m_2+0.5)}{N_2}\right),$$

where $\lambda_i(m)$ is a normalizing coefficient calculated as

$$\lambda_i(m) = \begin{cases} \frac{1}{\sqrt{N_i}}, m = 0 \\ \sqrt{\frac{2}{N_i}}, m \neq 0 \end{cases}.$$

C. Description of the Compression Algorithm

The studied compression algorithm consists of the following steps:

- splitting the image into blocks;
- calculating the DOT for each of the blocks;
- quantization of the obtained frequency domain (lossy compression);
- packing of quantized spectral components for subsequent lossless compression.

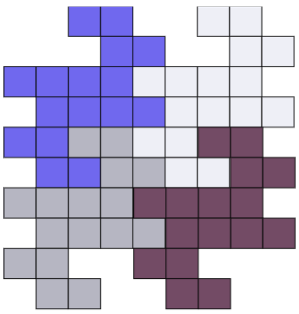


Fig. 1. An example of dividing an image into blocks when using FDCT for *α=-1+i, k=4.*

In the case of FDCP, the partition is performed in accordance with the k- fundamental fractal region (2). This area describes the shape of the block, and the block offsets over the entire image are calculated based on the size of the block (Fig.1). When using two-dimensional DCT, square blocks are used. In cases where the blocks go beyond the image border, the missing values are supplemented by brightness values from the nearest pixel.

Lossy compression is performed by quantizing the spectral components of each block in accordance with the quantization vector (or matrix in the two-dimensional case). Quantization vectors are calculated for each algorithm based on the standard square deviation (SSD) of the corresponding spectral components according to the formula

$$q_i = \left\lfloor \frac{\frac{\sigma_{max}}{\sigma_i}+10}{3} \right\rfloor \cdot Q,$$

where q_i is i-th component of the quantization vector (or matrix), σ_i is i-th component of the mean squared error vector, σ_{max} is the maximum value of the standard deviation for all components, Q is the algorithm parameter, which is the image compression ratio setting. The meaning of this formula is to give more quantization levels to a component with a larger standard deviation. For example, for FDCT $\alpha = -1 + i, k = 3$, the quantization vector at $Q = 1$ is equal to $(3, 4, 5, 4, 6, 7, 7, 6)$.

The quantized values of the spectral components are recorded sequentially, 2 bytes were allocated to each component.

III. Research

A. Description of the experiment

The comparison was carried out on 10 halftone images 512×512 in size from the Waterloo Gray Set. All images were compressed by algorithms using two-dimensional DCT on blocks 4 x 4 and 8 x 8 and FDCT with parameters $\alpha = -1 + i; k = 3,4,5,6$.

As a comparative measure of visual quality, PSNR, or the ratio of peak signal to noise, and MSSIM, or a measure of structural similarity averaged over the image, were chosen.

PSNR is calculated by the formula

$$PSNR(x, y) = 20log_{10}255 - 10log_{10}\frac{1}{N_1N_2}\sum_{i=0}^{N_1-1}\sum_{j=0}^{N_2-1}[x(i,j) - y(i,j)]^2,$$

where *x* and *y* are the compared grayscale images, N_1, N_2 are image width and height respectively; PSNR value is measured in decibels. The higher the PSNR value, the less the image has changed compared to the original.

MSSIM is calculated as the average SSIM for disjoint 8×8 blocks :

$$SSIM(x, y) = \frac{(2\mu_x\mu_y+C_1^2)(2\sigma_{xy}+C_2^2)}{(\mu_x^2+\mu_y^2+C_1^2)(\sigma_x^2+\sigma_y^2+C_2^2)},$$

$$MSSIM = \frac{1}{M}\sum_{i=0}^{M-1} SSIM\,(x, y),$$

where *x* and *y* are grayscale images being compared, *M* is the number of 8×8 blocks, $C_1 = 2.55$, $C_2 = 7.65$. MSSIM values range from -1 to 1, the higher value corresponds to a better visual similarity of two images [7].

To assess the degree of compression, informational entropy was used. Information entropy shows how much information the spectral component carries on average after compression [8], and describes the theoretical limit of sequence compression. Accordingly, the lower the value of entropy, the greater the compression ratio can be achieved by compressing this sequence. Entropy was calculated from a sequence of quantized spectral components by the formula

$$H = -\sum_{i=0}^{65535} p_i\,log_2 p_i,$$

where p_i – is the probability of occurrence of the value of i in the sequence.

B. Results

As a result of the study, it turned out that for most images for equal values of entropy, algorithms based on two-dimensional DCT show the best values of comparative measures of visual quality compared to algorithms based on FDCT (Fig. 2), but the following can be noted: firstly, when the entropy is one and a half bits per sample and higher, the MSSIM value for FDCT-based algorithms differ no more than by 1%, which means that the difference is almost imperceptible.

Secondly, starting from a certain value of entropy, the visual quality of images compressed by the FDCT algorithm is superior to the visual quality of images compressed by the algorithm based on two-dimensional DCT. This can also be seen from the graphs in Fig. 2. Such a property can be useful in image transmission systems for which low PSNR values (about 20 dB) are acceptable.

Moreover, Fig. 3 shows the nature of the distortions introduced by the FDCT fractal blocks. Compared to the square blocks of two-dimensional DCT, the fractal structure is less noticeable, and the boundaries of the objects in the image are sharper, although more noisy.

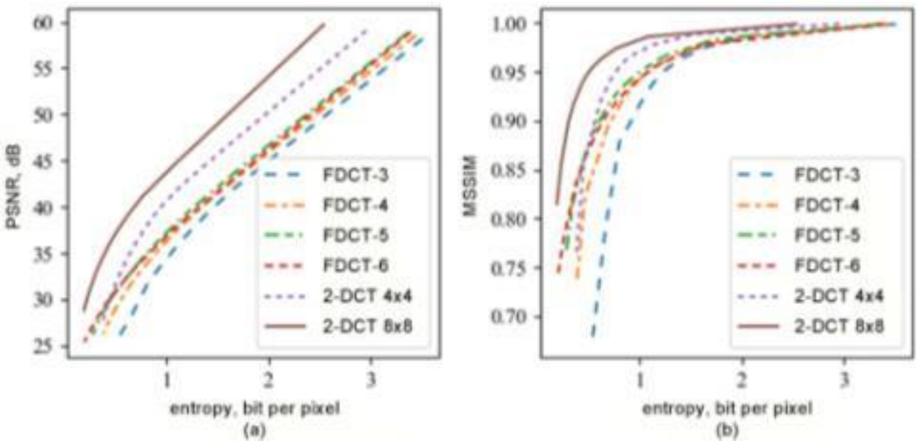


Fig. 2. Dependence of the visual characteristics of the Cameraman image on the information entropy: a) PSNR; b) MSSIM.

a b c

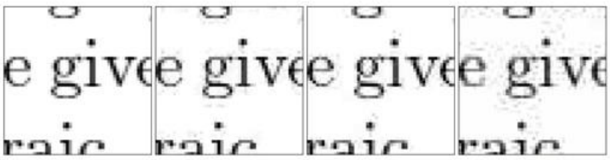
Fig. 3. Fragments of the Cameraman image: a) before compression; b) after compression by an algorithm based on two-dimensional DCT 8×8 (PSNR = 28.7 dB; MSSIM = 0.81); c) after compression by an algorithm based on FDCT $k = 6$ (PSNR = 25.42 dB; MSSIM = 0.74). Both compressed images have an entropy of 0.19 bit/pixel.

Finally, it can be noted that in experiments on images consisting of text, FDCT-based algorithms showed themselves better than algorithms based on two-dimensional DCT, which makes great practical sense when working with scanned documents and books. An example of the operation of algorithms in images containing text is shown in Fig. 4.

IV. Conclusion

In this paper, a lossy image compression algorithm based on a fractal discrete cosine transform was implemented and studied. The implemented algorithm was compared with the algorithm based on two-dimensional DCT. As a result, it turned out that FDCT has a completely different character of distortions introduced into the image during compression: an image compressed by the FDCT algorithm has sharper but more noisy object boundaries compared to two-dimensional DCT; the structure of fractal blocks is less noticeable than the structure of a square block of two-dimensional DCT. Despite the fact that FDCT does not show the best numerical characteristics of visual quality with an equal value of entropy compared to two-dimensional DCT, the actual visual quality differs insignificantly for some values of entropy, which can be used in a number of image processing areas.

Actual problems associated with the FDCT-based compression algorithm are the synthesis of fast FDCT algorithms, the study of FDCT-based algorithms in other k-fundamental areas, as well as the synthesis of the algorithm for reducing the noise introduced by compression when using FDCT.



a b c d

Fig. 4. Image fragments with text: a) before compression; b) after compression by an algorithm based on FDCT $k = 3$ (PSNR = 24.54 dB; MSSIM = 0.92); c) after compression by an algorithm based on two-dimensional DCT 4×4 (PSNR = 28.04 dB; MSSIM = 0.96); c) after image compression by an algorithm based on two-dimensional DCT 8×8 (PSNR = 25.86 dB; MSSIM = 0.89). All compressed images have an entropy of 1.2 bits/pixel.

References

[1] A.M. Belov, “The study of the effectiveness of one-dimensional discrete cosine transforms on the scans of two-dimensional signals generated by canonical number systems,” Computer Optics, vol. 35, no. 4, pp. 519-522, 2011.

[2] M.S. Kasparyan, “Fractal discrete cosine transformations on prefractal areas associated with the fundamental areas of canonical number systems,” Computer Optics, vol. 38, no. 1, pp. 148-153, 2014.

[3] I. Katai and A. Kovacs, “Canonical number system in imaginary quadratic fields,” Acta Mathematica Hungarica, vol. 37, pp. 159-164, 1981.

[4] I. Katai and J. Szabo, “Canonical number systems for complex integers,” Acta Sci. Math. (Szeged), vol. 37, pp. 255-260, 1975.

[5] V.M. Chernov, “Arithmetic methods for the synthesis of fast discrete orthogonal transform algorithms,” M.: Fizmatlit, 2007.

[6] V.M. Chernov, “Exotic" binary number systems for rings of Gauss and Eisenstein integers,” Computer Optics, vol. 42, no. 6, pp. 1068-1073, 2018, DOI: 10.18287/2412-6179-2018-42-6-1068-1073.

[7] Z. Wang, Alan C. Bovik, Hamid R. Sheikh and E.P. Simoncelli, “Image Quality Assessment: From Error Visibility to Structural Similarity,” IEEE Transactions on Image Processing, vol. 13, pp. 600-612, 2004.

[8] V.D. Kolesnik and G.Sh. Poltyrev, “Information theory course,” M.: Nauka, 1982.