Journal of Electronic Systems
Print ISSN: 2278 – 652X
Online ISSN: 2278 – 6538
DLINE JOURNALS
JES 2024; 14 (1)
https://doi.org/10.6025/jes/2024/14/1/1-9

Use of Wavelet Spectrum for Digital Image Compression

Teodora G. Sechkova[1], Ivo R. Draganov[2]
[1,2]The Faculty of Telecommunications, 8 Kliment Ohridski Blvd., 1000 Sofia, Bulgaria
{teodora.sechkova@gmail.com}
{idraganov@tu-sofia.bg}

ABSTRACT

In this paper, we present a new method for digital image compression. This method involves dividing the wavelet spectrum into several sub-bands after a certain level of regular or irregular decay. Each sub-band is then decomposed using an inverse pyramid algorithm and a linear orthogonal transform like DCT. Depending on whether you want lossless compression or lossy compression, all or only some spectral coefficients of the inverse pyramid are retained. Entropy coding is then applied. Higher compression ratios are achieved at higher image quality levels than some popular algorithms from the practice.

Received: 18 September 2023
Revised: 12 December 2023
Accepted: 28 December 2023
Copyright: with Author(s)

Keywords: *Image Compression, Wavelet, IPD, DCT*

1. Introduction

During the last three decades a large number of algorithms were designed for digital images compression. Many methods rely on spectral decorrelation combined with entropy coding [1]. Historically one of the most popular approaches that had become a standard and is still in wide use is the JPEG compression algorithm [2]. It can achieve extremely high compression ratios (over 100 times) at relatively high peak signal-to-noise ratios (PSNR, in some cases above 30 dB). Nevertheless some serious disadvantages could be pointed out for it such as the block effect due to the localizing property of the discrete cosine transform (DCT) and the lack of any scaling ability apart from the group change of the quantization matrix coefficients in some more flexible fashion, e.g. some multilevel or multistage processing. Extension like that could provide even more opportunities for scalable progressive transmission of image data over narrowband communication channels such as far distance satellite links for planetary observation.

Another large group of algorithms for image compression is based on the spatial-scale decomposition dividing the information for large and small objects from the image into separate bands. The wavelet transforms [3, 4] are the essence of such approaches and naturally led to the well-known JPEG 2000 standard [5] in which the block effect is absent. Here some more advanced techniques were introduced at the level of information redundancy reduction such as the Embedded

Zero-Tree Wavelet (EZW) coding and the Embedded Block Coding with Optimal Truncation (EBCOT). Another advantage is the ability for scalable transmission (restoration) of the image based on partial information transmitted in terms of bit-planes. Despite all the novelties in this approach at the level of visual objects inside the image some real multi-level scheme is thought to be proper for further enhancement.

Recently a new approach has been proposed for digital images decomposition denoted as inverse pyramid [6]. It introduces decomposing levels for a single image by dividing it in smaller blocks over which linear orthogonal transform of any kind could be applied. Then restoring of the blocks in the next level is done by using only some of the spectrum coefficients while the difference between the original and the approximated blocks is being preserved and passed to the next level. This particular decomposition is thought to be very efficient and flexible for digital images compression as well as in the field of pattern recognition for simplified object representation and speeding up algorithms. It is in the base of the proposed here algorithm along with the advantages given by the wavelet transforms.

In the next section detailed description of the new algorithm is presented, then in part three some experimental data is given and then conclusion is made in part four.

2. Algorithm Description

The input image I is grayscale consisting of N by N pixels and intensity range from 0 to 255. It is transformed by using the Haar wavelet according to:

$$B_{2x2} = TI_{2x2}\, T^T , \tag{1}$$

where I_{2x2} is a block of the image with size 2×2 and T – the Haar transform matrix of the same size given by:

$$T = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \tag{2}$$

Going through all 2×2 blocks of the image and rearranging the resulting spectral coefficients from the resulting spectrum B into 4 sub-bands the complete 1-level Haar decomposition is achieved.

Then each sub-band of the wavelet spectrum BR (R = 1÷4) is divided to blocks with dimensions $2^n\times 2^n$ and each of them is presented with Inverse Pyramidal Decomposition (IPD) [6]:

$$[B_R(2^n)] = [\tilde{B}_{0R}(2^n)] + \sum_{p=1}^{2} [\tilde{E}_{p-1,R}(2^n)] + [E_{2,R}(2^n)] , \tag{3}$$

where 3 levels are used and $[E_{2,R}(2^n)]$ is the matrix of the residual from the decomposition. In (3) each matrix is with dimensions $2^n\times 2^n$. The first component $[B_{0R}(2^n)]$ for the level $p = 0$ is a rough approximation of the block $[B_R(2^n)]$. It is obtained by applying inverse 2D-DCT over the transformed block in correspondence with the expression:

$$\tilde{[}B_{0R}(2^n)] = [T_0(2^n)]^{-1} [S_{0R}(2^n)] [T_0(2^n)]^{-1} , \tag{4}$$

where $[T_0(2^n)]^{-1}$ is a matrix with dimensions $2^n\times 2^n$ for the inverse 2D-DCT.

The matrix is the transform block of the cut 2D-DCT over $[B_R(2^n)]$. Here $m_0(u, v)$ are the elements of the binary matrix mask $[M_0(2^n)]$ with the help of which the preserved coefficients are being determined $[S_{0R}(2^n)]$ in accordance to the equation:

$$m_0(u,v) = \begin{cases} 1, \ if \ \ S_{0R}(u,v) \ is \ retained \ coefficient, \\ 0 \quad - \quad otherwise, \end{cases} \quad (5)$$

for $u,v = 0,1,.., 2^n\text{-}1$.

The values of the elements $m_0(u,v)$ *are chosen by the condition the retained coefficients to correspond to those with the highest average energy into the transformed blocks for all the blocks to which the image has been divided. The transformed block from* $[B_R(2^n)]$ *is found by the 2D-DCT:*

$$[S_{0R}(2^n)] = [T_R(2^n)]\,[B_R(2^n)]\,[T_0(2^n)]\,, \quad (6)$$

where $[T_0(2^n)]$ *is a matrix with dimensions* $2^n \times 2^n$ *for level = 0 which is used for implementing the DCT.*

The rest components in decomposition (3) are the approximation matrices for $p = 1, 2$. *Each of them consists of sub-matrices with dimensions* $2^{n-p} \times 2^{n-p}$ *for* $k_p=1,2,.,4p$ *obtained by its quad-tree split. On the other hand each submatrix is calculated by:*

$$[\tilde{E}^{k_p}_{p-1,R}(2^{n-p})] = [T_p(2^{n-p})]^{-1}\,[\tilde{S}^{k_p}_{pR}(2^{n-p})] + [T_p(2^{n-p})]^{-1}\,, \quad (7)$$

for $k_p=1,2,.,4p$. *Here 4p is the number of the branches of the quad-tree in level* p *of the decomposition;* $[T_p(2^{n-p})]^{-1}$– *matrix for inverse 2D-DCT;* $[\tilde{S}^{k_p}_{pR}(2^{n-p})]$- *the transformed block of the cut 2D-DCT of the difference matrix* $[E^{k_p}_{p-1,R}(2^{n-p})]$. *The elements* $\tilde{s}^{k_p}_{pR}(u,v) = m_p(u,v).s^{k_p}_{pR}(u,v)$ *of the matrix* $[\tilde{S}^{k_p}_{pR}(2^{n-p})]$ *depend on the elements* $m_p(u,v)$ *of the binary mask* $[M_p(2^{n-p})]$ for $u,v = 0,1,.., 2^{n-p}\text{-}1$ *according:*

$$m_p(u,v) = \begin{cases} 1, \ if \ \ s^{k_p}_{pR}(u,v) \ - retained \ coefficient, \\ 0 \quad - \quad otherwise, \end{cases} \quad (8)$$

Here $s^{k_p}_{pR}(u,v)$ *are elements of the transformed block* $[s^{k_p}_{pR}(2^{n-p})]$ *which is obtained also by the 2D-DCT:*

$$[\tilde{S}^{k_p}_{pR}(2^{n-p})] = [T_p(2^{n-p})]^{-1}\,[\tilde{E}^{k_p}_{p-1,R}(2^{n-p})] + [T_p(2^{n-p})]\,, \quad (9)$$

where $[T_p(2^{n-p})]$ *is a matrix with dimensions* $2^{n-p} \times 2^{n-p}$ *for level = 0 by which DCT is applied.*

It is possible to represent each group of four neighbouring elements $\tilde{s}^{k_p}_{pR}(u,v)$ *for one and the same* u *and* v *in the way defined by (10) which allows to gain even higher correlation between the spectral coefficients since the last three ones for positions (0,1), (1,0) and (1,1) form differences two by two and these differences often are zero valued because neighboring blocks contain almost identical content.:*

$$\begin{bmatrix} \tilde{d}^{k_p}_{pR}(0,0) \\ \tilde{d}^{k_p}_{pR}(0,1) \\ \tilde{d}^{k_p}_{pR}(1,0) \\ \tilde{d}^{k_p}_{pR}(1,1) \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 4 & 0 & -4 \\ -4 & 0 & 4 & 0 \\ 0 & 0 & -4 & 4 \end{bmatrix} \begin{bmatrix} \tilde{s}^{k_p}_{pR}(0,0) \\ \tilde{s}^{k_p}_{pR}(0,1) \\ \tilde{s}^{k_p}_{pR}(1,0) \\ \tilde{s}^{k_p}_{pR}(1,1) \end{bmatrix} \quad (10)$$

The inverse transform which leads to full restoration of $\tilde{s}^{k\,p}_{pR}$ *(u,v) is given by:*

$$\begin{bmatrix} \tilde{s}^{k\,p}_{pR}(0,0) \\ \tilde{s}^{k\,p}_{pR}(0,1) \\ \tilde{s}^{k\,p}_{pR}(1,0) \\ \tilde{s}^{k\,p}_{pR}(1,1) \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 4 & -1 & -3 & -2 \\ 4 & 3 & 1 & 2 \\ 4 & -1 & 1 & -2 \\ 4 & -1 & 1 & 2 \end{bmatrix}\begin{bmatrix} \tilde{d}^{k\,p}_{pR}(0,0) \\ \tilde{d}^{k\,p}_{pR}(0,1) \\ \tilde{d}^{k\,p}_{pR}(1,0) \\ \tilde{d}^{k\,p}_{pR}(1,1) \end{bmatrix} \quad (11)$$

The difference matrix $[E_{p\text{-}1,\,R}(2^{n\text{-}p})]$ *for level containing the sub-matrices* $[E^{k\,p}_{p-1,R}(2^{n\text{-}p})]$ *is determined by the following equation:*

$$[E_{p\text{-}1,\,R}(2^{n\text{-}p})] = \begin{cases} [B_R(2^n)] - [\tilde{B}_{0R}(2^n)] & -\ for\ p = 1; \\ [E_{p\text{-}2,R}(2^{n\text{-}p})] - [\tilde{E}_{p\text{-}2,R}(2^{n\text{-}p})] & -\ For\ p = 2. \end{cases} \quad (12)$$

Over the coefficients $\tilde{s}^{k\,p}_{pR}$ *(u,v) for all the levels of the pyramid for each sub-band a lossless entropy coding is applied which includes run-length encoding, Huffman coding and arithmetic coding. The resulting values could be stored in a file which volume determines the compressed image size. A decoder for restoring the compressed images with the proposed approach should consist of all the opposite operations to those described by (1)-(12).*

3. Experimental Results

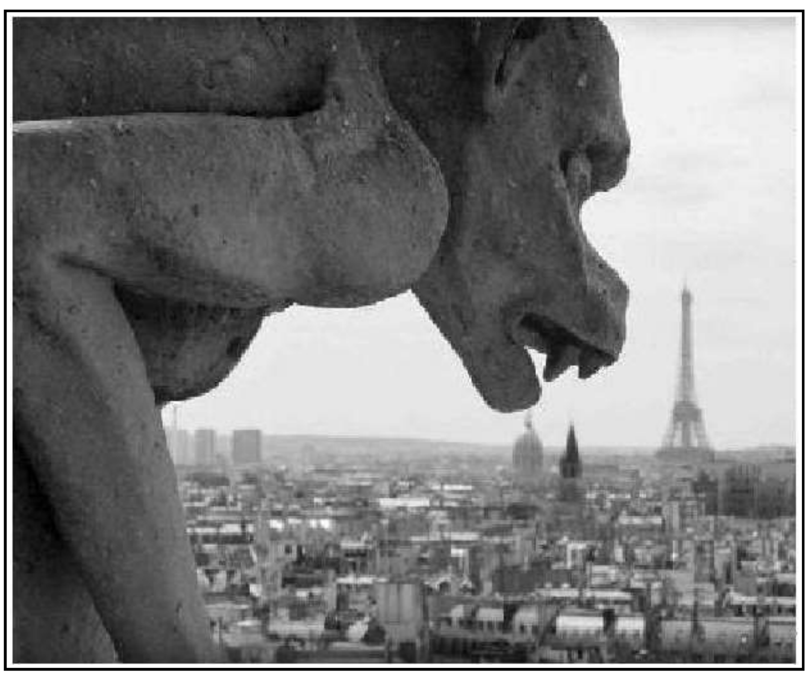
Two test images were used with the proposed approach called Flower and Gargoyl. They are 8 bpp (bits per pixel) grayscale images with size 512×512 pixels shown in Figure 1. Lossy compression was applied over them with 3 levels of the IPD after the Haar transform with 1 level of decomposition into 4 sub-bands –from LL to HH. On each level of the pyramid the 4 lowest frequencies in the DCT spectrum were retained when working with n = 4, 3 and 2 respectively, that is DCT transform matrix of 16×16 for 0th level, 8×8 – for the 1st and 4×4 – for the last one.

To compare the obtained results in terms of quality and compression levels Lura Wave Smart Compress 3 [7] application was used to compress the test images using the JPEG2000 standard (in JP2 format). The results are given in Table 1. Lossless compression is also possible by preserving the residual after level 2 in the pyramid but which here is omitted.

In the range between 0,5 and 2 bpp for the compression ratio (CR) there is tangible overweight in the PSNR achieved by the proposed approach over the JPEG2000 algorithm from about 4 dB for the higher ratios to around 0,05 dB. Obviously going towards lossless compression both algorithms tend to flatten the results they achieve. And the opposite, with the reduction of the CR the inverse pyramid decomposition over the wavelet spectrum of the image produces higher quality restored images at equal compression levels.

The visual analysis of the restored images from both algorithms (Figure 1) reveals the same tendency – for low CR levels most of the smaller details are preserved in almost identical way for IPD and JPEG2000. At the same time large homogenous areas are clean from any additional artifacts introduced by the compression. But when CR begins to decrease considerably JPEG2000 blurs most of the smaller details in the image mainly because of the quantization of the wavelet spectral coefficients and specially of the low-level HH components. In the new approach where IPD is used no quantization of the wavelet spectrum is applied directly. The second transform in the pyramid (DCT) produces highly efficient decorrelation of these coefficients and even then they are being quantized for the given level retaining the error for the next level. In such a way much *smoother transition is guaranteed from level to level not omitting the details from the image in intolerable degree. Another major advantage is the almost full absence of artifact distortions in the even areas because of the properties of the wavelet spectrum which is used in a first place but not the original image as it is with the classical JPEG approach. Here the block effect is much more suppressed.*

(a)



(b)

(c)

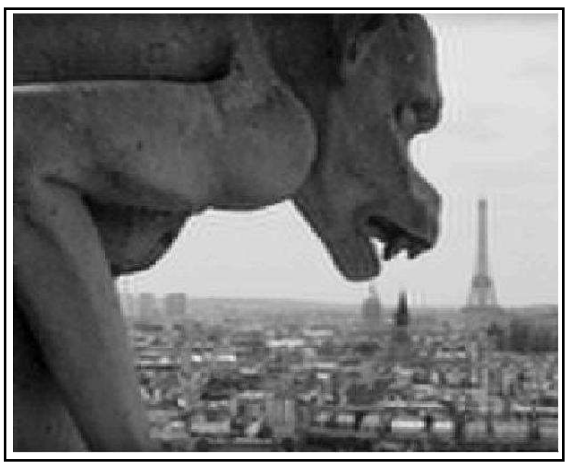
(d)

(e)

(f)

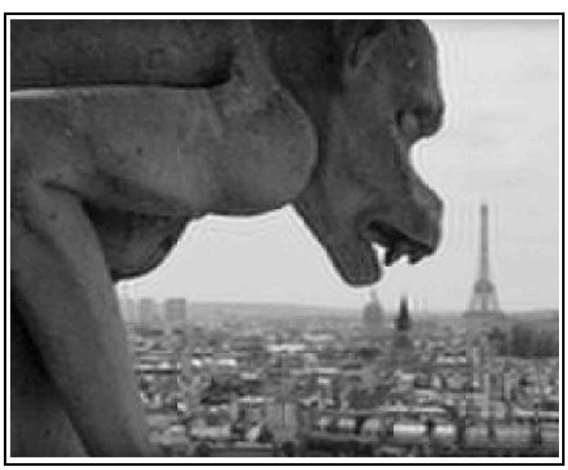
(g)

(h)



(i)

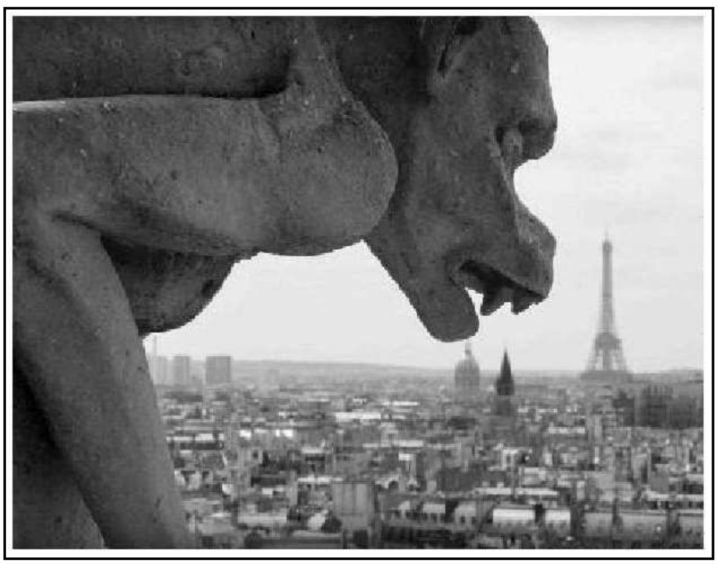
(j)



(k)

(l)

(m)



(n)

Figure 1. Visual test results: a) original Flower image; b) original Gargoyl image; c) IPD 0,5 bpp; d) IPD 1 bpp; e) IPD 2 bpp; f) JP2 0,5 bpp; g) JP2 1 bpp; h) JP2 2 bpp; c) IPD 0,5 bpp; d) IPD 1 bpp; e) IPD 2 bpp; f) JP2 0,5 bpp; g) JP2 1 bpp; h) JP2 2 bpp.

4. Conclusion

In this paper a new approach is presented for highly efficient image compression using wavelet and inverse pyramid decomposition. The quality and the compression level of the images could be smoothly controlled by choosing different levels of both the decompositions and the quantization mask for the pyramid at given level. Progressive image transmission from level to level is possible which proves to be very useful in the case of narrow-band communication channels.

Incorporating the advantages of the wavelet decomposition which minimizes the block effects

CR, bpp	Flower Test Image		Gargoyl Test Image	
	PSNR, dB		PSNR, dB	
	IPD	JPEG2000	IPD	JPEG2000
0,5	34,772	30,876	28,689	27,594
1	35,698	34,952	29,075	28,632
2	36,615	36,560	29,351	28,691

Table 1. Compression Results

due to the absence of localizing properties with the strong decorrelating properties of the orthogonal transforms such as DCT it is possible now to have strongly compressed digital images which when restored have low presence of artifacts. As it is obvious from experimental results higher compression ratios are achieved at equal image quality levels compared to some popular algorithms from the practice such as JPEG2000. This makes the proposed approach a suitable candidate for further improvement and establishing a new codec design for highly efficient image compression.

Acknowledgement

This paper was supported by the National Science Fund of the Bulgarian Ministry of Education, Youth and Science (Contract – DDVU 02/13 – "Public and Private Multimedia Network Throughput Increase by Creating Methods for Assessment, Control and Traffic Optimization").

References

[1] Rao, K. R., Yip, P. (2001). *The Transform and Data Compression Handbook*. CRC Press.

[2] Wallace, G. (1992). The JPEG still picture compression standard. *IEEE Transactions on Consumer Electronics, 38*(1), 18-34.

[3] Walker, J., Nguyen, T. (2000). Wavelet-based image compression. In *Handbook of Transforms and Data Compression* (pp. 267-312). CRC Press.

[4] Jacob, S., Cheeran, A. (2010). Wavelet based image compression. In *Proceedings of the International Conference and Workshop on Emerging Trends in Technology ICWET'10* (p. 999).

[5] Skodras, A., Christopoulos, C., Ebrahimi, T. (2001). The JPEG 2000 still image compression standard. *IEEE Signal Processing Magazine, 18*(5), 36-58.

[6] Kountchev, R., Kountcheva, R. (2008). Image representation with reduced spectrum pyramid. In Tsihrintzis, G., Virvou, M., Howlett, R., Jainm, L. (Eds.), *New Directions in Intelligent Interactive Multimedia* (pp. 275-284). Springer.

[7] Lura Tech. Retrieved from http://www.luratech.com