File Format Organization for Effective Still Image Transfer with IDP

Roumen K. Kountchev[1] and Vladimir T. Todorov[2]

Abstract - **In the paper is presented one new image compression method – Inverse Difference Pyramidal decomposition, aimed at image transfer with increasing resolution. Specific feature of the method is that the pyramidal decomposition is performed in the frequency domain, and every successive layer consists of larger number of coefficients, i.e. the pyramid, constituted of these coefficients, is inverse. The higher pyramid levels ensure better quality for the restored image. Basic advantage of pyramidal image decomposition from the known methods is that the received initial information for the lowest pyramid levels is used to upgrade the whole image and no part of the information is send twice, transferring the image data layer by layer.**

Keywords - **Image compression, Inverse difference pyramid, Progressive image transmission, Layered image transfer**

I. INTRODUCTION

The most up-to-date applications of image processing are aimed at such areas as e-commerce, digital libraries and mobile video communications and require the creation of specific structures of the compressed data, which to facilitate the process of data transmission and receiving. The best image compression methods have to ensure random codestream access and processing, progressive transmission by pixel accuracy and resolution, to possess open architecture and to permit content-based description. Some of the most effective methods for still image compression are based on different kinds of pyramidal decompositions. Two basic approaches for pyramid building are known: orthogonal and non-orthogonal [1, 2, 3, 4]. These two groups of pyramidal decompositions have some common deficiencies: they require multiple decimations and interpolations together with low frequency or subband filtration, which cause specific image distortions.

The specific features of the Inverse Difference Pyramid (IDP) method [5, 6], which result from the image decomposition in the frequency domain, suit all these requirements and together with this, avoid the mentioned deficiencies to a great extend. In the paper is represented one application of the IDP method, which permits flexible image transfer layer by layer, gradually increasing its resolution and quality.

II. BASIC PRINCIPLES OF THE IDP METHOD

The essence of the IDP decomposition is easily explained for 8-bit greyscale image. As it was already mentioned, the processing is performed in the frequency domain. For this reason at the beginning of the processing the whole image is processed with two-dimensional (2D) orthogonal transform using only limited number of coefficients. The values of the coefficients, obtained in result of the transform, constitute the first pyramid level. Then, using the values of these coefficients, the image is restored with inverse orthogonal transform and subtracted pixel by pixel from the original one. The difference image, which is of same size as the original one, is divided in 4 sub-images and each sub-image is processed with the 2D orthogonal transform again; the values of the coefficients constitute the second pyramid level. After that the processing continues in similar way as in the already described preceding level. In this way all pyramid levels, consisting of coefficients only, are calculated. The set of coefficients of the orthogonal transform, chosen for every pyramid level, can be different. The image decomposition is stopped when the required image quality (for the restored image) is reached - usually earlier than the last possible pyramid level. The coefficients got in result of the orthogonal transform from all pyramid levels are quantizated, sorted in accordance with their frequency, and scanned sequentially. In order to achieve higher compression ratio the data is processed, applying special entropy coding.

The same method is used for processing of 24-bit color bmp images too. In case when color images are compressed, three similar pyramids are built - one for each color component. For RGB images, preferred standard is usually 4:4:4. For images in format YC_rC_b is selected one of standards 4:2:2, 4:2:0, or 4:1:1. The chrominance pyramids have the same number of levels as the luminance pyramid, but their resolution differs in accordance with the selected standard. Higher compression ratio for color images is obtained using the techniques already explained for greyscale images: limitation of the coefficients and pyramids levels number, quantization with increasing steps, and adaptive entropy coding. For standards 4:2:2, 4:2:0, or 4:1:1 every pyramid could have different settings, but the standard 4:4:4 requires the settings for the three pyramids to be the same. The algorithm offers highest compression ratio with retained high image quality for standard 4:2:0.

All parameters used for the image compression are described in a special header, sent together with the compressed data. The header contains information about the number of pyramid levels (image layers), the start level, and the selected orthogonal transform used for the image processing, the number of coefficients and their corresponding spatial frequencies, the quantization parameters and the kind of entropy and run-length coding used in every level.

[1] Roumen Kountchev - Technical University of Sofia,
Department of Radiocommunications
Boul. Kl. Ohridsky 8, Sofia 1000, Bulgaria
e-mail: rkountch@vmei.acad.bg
[2] Vladimir Todorov - T&K Engineering Co.
P.O.Box 12, Sofia 1712, Bulgaria

III. FORMAT OF COMPRESSED IMAGE

The compressed data consists of two basic parts: the image header and the data, obtained after the processing (the compressed image data). If we compress the image and send all the information together, the data structure is relatively simple – it consists only of the header and the compressed data. The header contains the data necessary for the proper decoding. It consists of 3 parts: general header, special data sub-header and sub-header for color data. Our everyday practice requires much more flexible approach in image compression and transfer. Usually such applications as e-commerce, m-trade, B2B, B2C, etc., require fast initial image transfer without many details, which are send to the customer only if required. For such applications the IDP method is very suitable because it permits easy image transfer layer by layer, with increasing resolution and quality. In such cases the data structure is more complicated. The header is divided in as many parts as there are layers, retaining same organization. Here follows the brief description of the basic header data. The used names are created as abbreviations of their functions. This organization is shown in Fig.1

Main Header	Special data sub-header	Sub-header for color data	Additional data	Packet 1 (Lowest level)
Sub-headers for level 2				Packet 2
Sub-headers for level 3				Packet 3

Fig.1. Data structure for multi-layer image transfer.

Kuhdr - general header
Kutype (16 bit integer) - software version;
Rlbytecnt (unsigned 32 bit) - number of bytes after run-length coding;
Bytecnt (unsigned 32 bit) - final number of bytes, after entropy coding;
Prgprm (unsigned, 16 bit) - program parameters, containing all 1-bit parameters (flags): the way of picture scanning, data arrangement, color transform, etc.
Prgprm1 (unsigned, 16 bit) - initial and end level of the truncated pyramid and maximum possible number of pyramid levels for the processed image.

Kusdh - special data sub-header
ILVFL (unsigned, 16 bit) - mask for the coefficients in the initial pyramid level;
MLVFL (unsigned, 16 bit) - mask for the coefficients in the middle pyramid levels;
ELVFL (unsigned, 16 bit) - mask for the coefficients in the highest pyramid level;
Rltype (unsigned char);
Rlcod (unsigned char);
Rlcod1 (unsigned char) – selected kind of RL;
DivCC (char) - division coefficient for the first pyramid level
LVLidp (unsigned char) - defines the selected approximation method for each level (WHT, DCT, etc).
Nn (unsigned, 16 bit) - size of the original image (vertical)
NHn (unsigned, 16 bit) - size of the original image (horizontal direction)

Kusdhc - sub-header for color data
ILVc (unsigned char) - initial pyramid level
ELVc (unsigned char) - end pyramid level
ILVFLc (unsigned 16 bit) - mask for the coefficients in the initial pyramid level;
MLVFLc (unsigned 16 bit) - mask for the coefficients in the middle pyramid levels;
ELVFLc (unsigned 16 bit) - mask for the coefficients in the highest pyramid level;
DivCCc (char) - division coefficient for the first pyramid level;
LVLidpc (unsigned char) - defines the selected approximation method for each level (WHT, DCT, etc).
Nnc (unsigned, 16 bit) - the size of the original image (vertical)
NHnc (unsigned, 16 bit) - the size of the original image (horizontal)
Uvpxlc (unsigned, 16 bit) - used color standard (4:1:1, 4:2:2, etc.)
The **additional data** contains information about other parameters, necessary for definition of the selected kind of lossless data coding.

IV. LAYERED IMAGE TRANSMISSION WITH IDP

The layered image transfer is best illustrated with Fig. 2. In Fig.2a is shown the image, restored using the information from the lowest pyramid layer only. This is the whole picture, processed with low resolution and visualized scaled down (20% of the original image size). If the customer wants to see the whole, full-size picture the remaining information (the higher pyramid levels up to the last one), follow. If only a part of the picture is needed, the customer can select it and after that the corresponding information from the higher pyramid levels only for this part of the image will be transferred and received. In the example are used only three of the pyramid levels. Advantage of the method is that no part of the compressed data is transferred twice. The information from the initial image (the first pyramid level) is used as a basis for the upgraded full-size image (or the selected part of it) and when all the remaining information is transferred to the receiving part, the image is restored. In case when only a part of the image is needed, it can be selected arbitrarily. In these cases the header and the data of the compressed image are rearranged in accordance with the requirements and the corresponding part of the picture is transferred and restored at the receiving side. The method permits processing with different (higher or lower) resolution for selected parts of the image.

V. RESULTS

For the evaluation of the quality of the restored image $B'(i, j)$ with size $2^n \times 2^n$ pixels was used the peak signal to noise ratio (PSNR), defined with the equation:

$$PSNR = 10\,\lg_{10}(255^2/\overline{\varepsilon^2})\ \ [dB]\ , \qquad (1)$$

$$\overline{\varepsilon^2} = \frac{1}{4^n}\sum_{i=0}^{2^n-1}\sum_{j=0}^{2^n-1}[B(i,j) - B'(i,j)]^2 \qquad (2)$$

Here $\overline{\varepsilon^2}$ is the mean square error of the restoration, B(i,j) and $B'(i, j)$ are correspondingly pixels from the original and restored images, and 255 - maximum greyscale value for these pixels.
The compression ratio for image coded with IDP, limited up to level p before RLE and EC, is calculated in accordance with the term:

$$C(p) = \frac{4^{n-p}\log_2 m}{K\sum_{k=0}^{p}(A_k + 4^k q_k R_k)} \quad \text{for } p = 0, 1, \ldots, P\text{-}1\ , \qquad (3)$$

where m is the number of the grey levels for single pixel; A_k – the number of bits used for coding the positions of the coefficients selected for the transform of k-th pyramid level; q_k – the number of bits used for the coding of every coefficient from the k-th level, R_k- the number of coefficients of the sub-image in k-th level, and K – number of sub-images in zero level. As an example, for n =9, P=3, m =256, K=1024, q_k=8, R_k=4 and A_k=0 for k=0,1,2 (i.e. for same positions used for the selected spectral coefficients), (Eq.3) we obtain correspondingly C(0)=64, C(1)=12,8 and C(2)=3,04. After RLE and EC was applied over pyramid coefficients, the compression ratio C(p) was increased 4 - 6 times, retaining the same image quality.

In Table 1 are given the results from the compression of the example image: Hungary. The size of the original image Hungary is 1024x768 pixels. Hun1 is the restored image obtained from the restoration of the lowest pyramid level, after compression 248,06, Hun2 – after compression 26,76 and Hun3 – after compression 20,55. These images show the results obtained using pyramid with 1, 2 and 3 levels. Similar results are obtained with picture Ariz. The original image Arizona is 1024x1024 pixels, 24bpp.

The example shows that in most cases the customer will be satisfied with the image quality of the second picture and will not need the last picture with highest quality (Fig.2).
The orthogonal transform used in the lowest pyramid layer was DCT, and for the two following levels – WHT. The selected spectral coefficients are as follows: for level 0 they are $s_0(0,0)$, $s_0(1,0)$ and $s_0(0,1)$; for level 1 - $s_1(0,0)$, $s_1(1,0)$, $s_1(0,1)$ and $s_1(1,1)$; for level 2 - $s_2(0,0)$, $s_2(0,1)$ and $s_2(1,0)$.
Operations quantization and dequantization are performed with correspondence with the equations:

$$s_p^q(u,v)=Q_p\{s_p(u,v)\}=[s_p(u,v)/\Delta_p(u,v)]_{int} \qquad (4)$$

$$s'_p(u,v)=Q_p^{-1}\{s_p^q(u,v)\}=s_p(u,v).\Delta_p(u,v) \qquad (5)$$

The values of the corresponding quantization steps $\Delta(u,v)$ are: $\Delta_0(u,v) = 4$ -for level 0;
$\Delta_1(0,0)= 16$, $\Delta_1(0,1)=\Delta_1(1,0)=32$, $\Delta_1(1,1)=64$ - for level 1;
$\Delta_2(0,0)=16$, $\Delta_2(0,1)=\Delta_2(1,0)=32$, $\Delta_2(1,1)=32$ - for level 2.

Fig.2a. Image "Ariz1"

Fig.2b. Image "Ariz2"

Fig.2c. Image "Ariz3"

Fig.2.Presentation of layered image transfer and decoding.
Fig.2a - restored image from pyramid level 0 (scaled down); Fig.2b – restored from layer 1; Fig.2c – restored from layer 2 (Data from Table 1).

For big values of the compression ratio, the results obtained using IDP are much better than these from JPEG. This is illustrated in Fig.3. The PSNR of the restored image is higher as well (Table 2).

TABLE 1.

RESULTS FROM COMPRESSION WITH IDP

IMAGE	LAYER	COMPRESSION RATIO	PSNR [dB]	FILE [B]
HUN1	0	248,06	25,41	9 511
HUN2	1	64,44	29,28	36 611
HUN3	2	20,55	34,96	114 800
ARIZ1	0	319,38	28,29	7 387
ARIZ2	1	80,31	33,17	29 378
ARIZ3	2	25,16	35,97	93 787

Fig.3a

Fig.3b

Fig.3c

Fig.3d

Fig.3. Comparison of IDP with JPEG compression. Fig.3a - Image Hungary, restored after compression ratio 164,74 with IDP; Fig.3b – Image Hungary after compression ratio 161,87 with JPEG (Microsoft Photo Editor); Fig.3c - Image Myanmar after compression ratio 106,09 with IDP; Fig.3d – Image Myanmar after compression ratio 106,04 with jpeg.

TABLE 2

COMPARISON FOR HIGH COMPRESSION RATIOS

IMAGE	IDP		JPEG	
	COMPR.	PSNR	COMPR	PSNR
HUNGARY	164,74	26,73	161,87	23,95
ARIZONA	314,24	28,38	171,07	24,33
MYANMAR	106,09	19,66	106,04	19,30

The comparison shows the better results obtained with IDP for big compressions. The smaller computational complexity of the method is another advantage which makes its applications welcome.

The software implementation of the IDP Method was performed with fast DCT and WHT [7] in C++. The special program, created on the basis of the described algorithm, is a powerful and flexible tool, permitting all the parameters – the number of pyramid levels, coefficients, quantization steps, etc. - to be set independently. The time necessary for the processing is not noticeable.

VI. CONCLUSION

A new method for layered image compression is presented. The main advantages of the method compared with the standard JPEG [8], are the following:

- It gives better image quality of the restored image for compression with high compression ratios (over 40:1) and similar image quality for lower compression ratios;
- Offers fast initial representation of the image (small size), permitting fast interactive request with higher resolution;
- Ensures easy selection of parts of the image which are represented with higher scale and resolution;
- The method has lower computational complexity than standards JPEG and JPEG 2000 [5, 8, 9].

The future development of the method could be in the following main directions:

- Pyramid adaptation on the basis of image histogram;
- ROI with arbitrary form for higher levels;
- Increasing the effectiveness of processing for color images, using Karhunen-Loeve transform;
- Improving the EC of the coded data with adaptive arithmetic coding;
- Increasing the noise suppression with pre- and post-image processing with homomorphic filtration.

These developments will depend on future applications and requirements.

REFERENCES

[1] B. Aiazzi, S. Baronti, F. Lotti. Embedded Image Coding Based on Laplacian Pyramids with Quantization Feedback. Signal Processing VIII, Trieste, pp. 316-319, 1996.

[2] A. Akansu, R. Haddad. Multiresolutiun Signal Decomposition: Transforms, Subbands and Wavelets, Academic Press, NY, 1992.

[3] O. Egger, P. Fleury, T. Ebrahimi. High-Performance Compression of Visual Information-A Tutorial Review-Part I: Still Pictures, Processing of the IEEE, Vol. 87, No 6, pp. 976-1011, June 1999.

[4] P. Topiwala. Wavelet Image and Video Compression. Kluwer Acad. Publ., NY, 1998.

[5] R. Kountchev, J. Ronsin. Inverse Pyramidal Decomposition with "Oriented" Surfaces. Picture Coding Symposium'99, April 21-23, 1999, Portland, USA, pp.395-398.

[6] R. Kountchev, V. Haese-Coat, J. Ronsin. Inverse Pyramidal Decomposition with multiple DCT. Signal Processing: Image Communication, Vol. 17, February 2002, pp. 201-218.

[7] K. Rao, P. Yip. Discrete Cosine Transform: Algorithms, Advantages, Applications, Academic Press, Toronto, 1990.

[8] G.K.Wallace. The JPEG Still Image Compression Standard. IEEE Trans. Consumer Electronics, Vol. 38, No.1, Feb. 1992.

[9] M. Rabbani, R. Joshi. An Overview of the JPEG 2000 Still Image Compression Standard. Signal Processing: Image Communication, Vol. 17, Jan. 2002, pp. 3-48.