Context-based method for lossless compression of RGB and multispectral images

A V Borusyak[1], P A Pakhomov[1], D Yu Vasin[1] and V E Turlapov[1]

[1]Lobachevsky State University of Nizhni Novgorod, Prospekt Gagarina 23, Nizhni Novgorod, Russia, 603952

Abstract. We consider the problem of compression of RGB and multispectral images by context-based methods. The algorithm' logic allows for its examination by using the example of full-color images as a particular case of multispectral images. The image-forming channels are divided into two groups: main and additional (detecting) channels. A distinguishing feature of the main channels is a significant correlation between neighbors. A number of variants of prediction from the adjacent channel for the main and additional channels for lossless image compression were considered. In the experiment on a series of images of different contents, the proposed algorithm showed a superior compression ratio in comparison with the popular WinRar, 7z, PNG archivers for all prediction variants. The leader among popular compression methods, JPEG-LS, was surpassed in the record configuration 2b on the image from the Landsat series by 40%. We expect to continue research on a wider sample of images and to use this algorithm to compress multispectral images with a greater number of channels.

1. Introduction

Context-based modeling is an important step in high-performance lossless data compression. Serious advantages offered by high compression degree enable prediction based on the model of matching (coincidence) of contexts. These advantages were successfully demonstrated using the Prediction by Partial Matching (PPM) method published in 1984 [1] to solve the task of text compression. In 2005, the PPM method was significantly improved in [2] by mixing several contexts with weights that change during the method execution (based on machine learning methods). In this method, the model of each context independently estimates the probability and confidence that the next data bit will be 0 or 1. The forecasts are further weighed (the sum of the weights is 1), the weights are corrected by the prediction success criterion. Open source code (www.mattmahoney.net/dc) software (PAQ8) is implemented in the method. This software demonstrated a high rating in several independent tests.

The wide use of this approach to images compression began around the early 2000s, but it was used primarily for binary images containing mostly text and lines [3]. It offered a compression improved by 14% compared with the analog and a 25% better performance. In 2002, a parallel algorithm for this method was developed [4]. In 2001, the context-based method was applied in the development of a new video coding standard for entropy estimation in the coding procedure using binary adaptive arithmetic coding technology, which increased the coding rate by 35% [5].

Of essential importance in the application of the context-based approach is the difference between images and texts consisting in the presence of noise in the images. Therefore, the result obtained in [6] in 2008 was very important for image compression: the prediction method for images with the help of

Prediction by Partial Approximate Matching (PPAM) was presented. Unlike the PPM modeling method that uses exact contexts, PPAM introduces the notion of approximate contexts. Thus, PPAM models the probability of encoding a symbol based on its previous contexts, and contextual occurrences as a result are considered in an approximate manner. The method demonstrated competitive lossless compression and good performance when compressing images that have repeating areas with similar characteristics.

However, the use of the context method for compressing color and multichannel images has remained complicated and ambiguous for a very significant reason: the effective definition and use of contexts for such images is a complex task, since in essence it is the compression of three or more images simultaneously. Nevertheless, in 2011, the publication [7] explored the prospects of using PAQ family methods in combination with machine learning (ML) methods for simple color images and for lossy compression. A number of problems were identified: 1) PAQ can be applied only to one-dimensional sequences, and the expansion for several sequences is not trivial (even in the case of identification of chicken carcass parts); 2) the authors were unable to construct parametric models of typical image contexts, which required for PAQ methods a huge storage capacity. In all four test images used to compare the PAQ-ML method with JPEG and JPEG2000, it was superior to JPEG2000 both in terms of the compression ratio and the quality of the compressed image. The method proposed by the authors showed a significant change in color amounting to the distortion of the palette, while JPEG2000 maintained the ratio of color channels in the local context, was able to locally parameterize the change of this ratio and thus proved to be the winner.

It is also of interest to study the possibilities of using context-based compression methods for color (RGB) and multispectral images of Earth remote sensing (ERS). In the general case, ERS images are multi-channel, i.e. each pixel in the image is specified by the channel value vector. The early compression algorithms included, as a rule, independent operations on individual sample matrices, which were the matrices of the original image channels, or one, two or all the three RGB channels assigned to represent them. Therefore, the publications at that time primarily considered algorithms for processing single-channel (halftone) images, which are basic for implementing all compression methods. More recent publications are related to the compression of hyperspectral images, where the hierarchical compression method for both hyperspectral images (HSI) and for ERS as a whole occupies one of the leading positions [8], [9], [10]. In [9], the following statistical characteristics of the HSI are given:

• the difference between the maximum and minimum brightness gradations reaches thousands and tens of thousands times; such images cannot be converted to "byte images";

• components are very dependent; intercomponent correlation is extremely high (above 0.95 for 85.2% of the pairs of neighboring components);

• most components have high intracomponent correlation (above 0.85 for 87.4% of all components).

In what follows, we will be guided by these considerations.

The hierarchical multiscale representation is based on the results of a number of previous studies. It serves to solve not only the problem of ERS data compression, but also several other problems at the same time, such as the problem of compact storage and high-performance adaptive (in terms of permissible losses and the observer's position) visualization of the terrain surfaces with a controlled value of distortions [11]. In this paper, we use a multi-scale wavelet representation of elevation data and a JPEG2000 encoder to compress 8-bit quantized height differences between their predicted and accurate values. This approach can be used without any significant changes to compress any channel (including the reference channel) of multi- and hyperspectral images.

The high correlation of most of the neighboring HSI channels allows us to apply context-based compression methods at a new level and to use the previous high-correlation channel as the context for the current channel, which has been successfully realized and investigated in publications [9], [10]. The high correlation of the HSI channels has made it possible to bring the level of their lossless compression to the values of the order of 4-5.

Unfortunately, it has not been possible so far to achieve this level for the compression of multispectral data, because the high correlation of neighbors is not a rule for such type of data. When

considering the problem of compressing multispectral data, we will assume the channels of multispectral images to be unequal in terms of their information role in the summary image. One of the channels will be taken as the main (reference) channel, while the others will be used as: 1) special contrast channels for detecting objects of interest; 2) complementary channels, highly correlated with the reference channel (if any). For example, in the RGB image of an oasis in the desert, the yellow sand will be determined by almost identical maps of the red (reference) and green (complementary) channels, and the blue water will be determined by the contrast blue (water-detecting) channel. A similar situation would arise if we were to shift the infrared and ultraviolet (detection) channels to the visible region around the green (reference) channel. Obviously, for evolutionary reasons, the red or green channels are more acceptable to us as support channels. We will take one of them as a reference channel in a "conditionally RGB" image to be compressed, which is quite close to the method of "common reference" spectral components for compressing hyperspectral images [9]. The context-based method for compressing RGB and multispectral images proposed below is the development of an algorithm for adaptive compression of indexed and color images with the use of context modeling [12-14].

2. The algorithm for lossless compression of RGB images

Compression of each pixel is performed channel-by-channel. First, the color component responsible for the red color is compressed, next, the value of the color component of the green color is encoded, and then the value of the blue component is encoded. For each channel, its own context is formed. The structure for context storage is identical to the structure of the algorithm for indexed images [13,14]. For each channel, individual context models of the following orders are used: 6,4,2,1,0. The full-color probability coder (FPC) compression algorithm [14] is as follows: 3 separate keys of the current context and 3 independent forests of AVL-trees are used to store context models (for each of the RGB channels). The following actions are performed in the cycle:

- A consecutive pixel is extracted from the input image file as a current one;
- The maximum-order context (MOC) is formed as the current context: the contexts Cont1, Cont2, Cont3 of the maximum order are formed sequentially for the red, green and blue channels, respectively, as an array of unsigned one-byte integers storing the previous values of the corresponding channel of the current pixel.
- The procedure of channel-by-channel coding of the current pixel value in the current context is performed (for details, see [13]);
- If it is not possible to evaluate and encode the current channel value in the current context, since this value is encountered in the current context for the first time, a lesser order context is formed and this context becomes current, thus a return to point 3 occurs. This continues until the current value of the color component is encoded, which is guaranteed by the fact that occurrence counters for all pixel values in the context of the smallest (zero) order are initially assigned the value of unity. The descent to a lesser order context is realized by applying the exclusion technique, which allows, in case of departure to the contexts of a smaller order m, to exclude from consideration all the values of the pixel occurrence counters that are contained in the context model of the order r, $0<m<r<R$, since none of them is an encoded pixel value.
- The context model (as the MOC) is updated in accordance with the code of the current pixel.
- If not all the pixels are coded, the transition to step 1 occurs, otherwise the encoding is completed.

3. Using the link between channels in context-based image compression

To increase the compression ratio, images were studied to determine the existence and possibility of using the relationship of color channel values.

Experiments have been performed to compare compression of full-color images in three modes:

- Independent channel coding: only the R channel values are used to form the R channel context; only the G channel values, for the context of the G channel; only the B channel values, for the context of the B channel.

- Partial relationship between the channels: for the context of the R channel, only the values of the R channel are used; for the G channel, the values of the R and G channels; for the context of the B channel, the values of all the channels: R, G and B.
- Complete relationship between the channels: for the contexts of each of the channels R, G, B, the values of all the channels R, G, B are used.

In full-color images, the color component values of neighboring pixels often have close or similar values. In most cases, there is a gradual color change from one pixel to another. It is known that for most ERS images, the histograms of the difference between the channel values of adjacent pixels in the line-by-line readout is a normal distribution with zero mean and small sigma. This allows us to use, instead of channel values, the difference between its values for the current and previous pixels. The same increments can be used if necessary to form the contexts: the first value in each line of the context is considered an increment of 0. Since increments can also be negative, it is advisable to apply a linear transformation. As a result, the calculation of increments occurs according to the formula:

$$rn_{i,j}^{q} = (g + (\alpha_{i,j}^{q} - \alpha_{i-1,j}^{q}))\%g,$$

where $rn_{i,j}^{q}$ is the new normalized value of the difference between the current $\alpha_{i,j}^{q}$ and the previous $\alpha_{i-1,j}^{q}$ values of the current color channel, q is the color channel number, g is the number of channel gradations (in full-color images it is 256), % is the modulo operation. The difference $rn_{i,j}^{q}$ is only used as the value of the current element being encoded; for the context elements, the initial values of the color components are used.

The FPC algorithm for full-color image compression is described in detail in [14].

In this algorithm, it is possible to use instead of the channel value the increments between its values for the current and previous pixels Experiments have been performed to compare compression of full-color images in three modes:

a) without using increments between neighboring pixels
b) using increments only for pixel values, and channel eigenvalues for the context.
c) using increments for both the pixel value and the context elements.

4. Results of experiments

To determine the most effective way of context formation, experiments were conducted to compress images of different size and contents, as shown in Tables 1-2. Table 3 contains thumbnail images used for experimental tests.

Table 1 shows the source file sizes and compression ratios for widely used algorithms, such as WinRar, 7z, PNG, JPEG-LS. The JPEG-LS algorithm is chosen for comparison, since it is closer to the proposed algorithm in terms of the compression method used and has a greater compression ratio for most images compared to JPEG2000 in the lossless compression mode. The compression ratio of the entire set of files is taken as the sum of the volume of all source files divided by the sum of the volumes of compressed files. In terms of the compression of the sum of files, the leader is JPEG-LS with the result of 2.18, while for the compression of the landsat.bmp file from the Landsat series, the best performance was shown by the 7z archiver with a result of 2.51 versus 1.98 for JPEG-LS.

Table 2 presents the comparative results separately for cases 1-3 described in Section 3 for the variants a, b and c. As a result of the experiments, it was established that approach No.2b with a partial relationship between RGB channels and using increments for pixel values only, and channel values proper for the context, turned out to be the most effective one. Approach No.3 showed a reduced compression ratio in comparison with approach No.1, Approach No.2 on the average was slightly more successful than No.1. Variant 2b is 3.5% better than 1a and 1b, more than 2% better than JPEG-LS for the whole set of files, and more than 40% better for compression of images from the Landsat series (and also 12% better for this image than 7z). It should be noted that the high compression ratio is achieved due to the greater use of computing resources (270 seconds for the compression of the landsat.bmp file compared to 55 seconds for WinRar, 100 seconds for 7z, 15 seconds for PNG, 5 seconds for JPEG-LS). We have established that the approach using the channel difference as the

values is especially effective in predicting when the correlation between the channels is taken into account.

Table 1. The sizes of test images and compression ratios for commonly used compression algorithms.

File \ algorithm	Size (bytes)	Compression ratio			
		WinRar	7z	PNG	JPEG-LS
30057A.bmp	186786486	1,84	1,72	1,37	2,21
artificial.bmp	18874422	10,64	16,54	11,28	12,07
big_tree.bmp	83101254	1,57	1,35	1,15	1,84
energy_bliss.bmp	3932214	4,50	2,56	1,97	4,72
landsat.bmp	46129734	2,10	2,51	1,93	1,98
The entire set of files	338824110	1,89	1,77	1,43	2,18

Table 2. Compression ratios for the variants 1a-3c of the proposed algorithm.

file \ algorithm	1a	1b	1c	2a	2b	2c	3a	3b	3c
30057A.bmp	2,04	2,04	2,03	2,11	2,16	2,12	1,96	2,10	2,03
artificial.bmp	13,38	13,40	12,65	17,39	19,08	17,65	16,97	19,23	17,18
big_tree.bmp	1,83	1,83	1,75	1,63	1,76	1,70	1,46	1,72	1,66
energy_bliss.bmp	2,44	2,46	2,62	2,85	3,98	4,29	2,74	3,72	4,15
landsat.bmp	2,82	2,81	2,39	2,76	2,82	2,42	2,69	2,71	2,32
The entire set of files	2,16	2,16	2,10	2,14	2,23	2,14	1,97	2,16	2,07

Table 3. Thumbnails of images being compressed.

30057A.bmp (8063x7721)	artificial.bmp (3072x2048)	big_tree.bmp (6088x4550)

energy_bliss.bmp (1280x1024)	landsat.bmp (3792x4055)
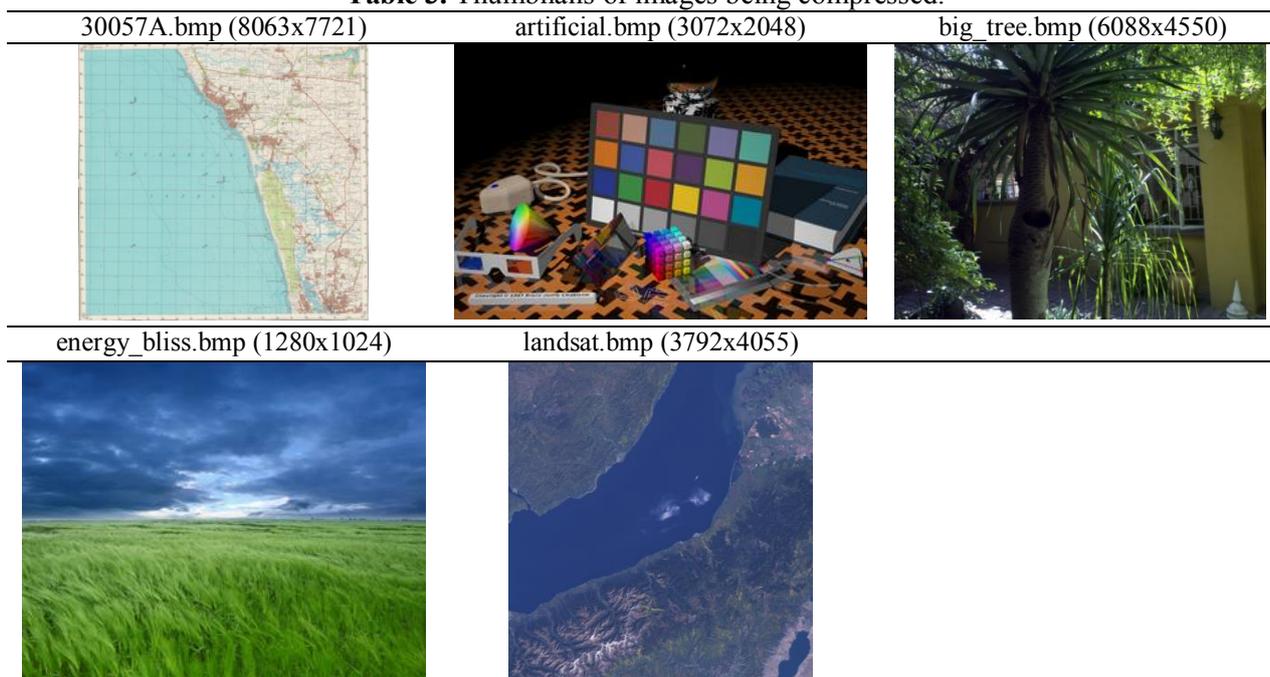
5. A detailed description of the approach used

Let us consider in detail variant 2b. The context formation is clearly illustrated in Figure 1. The colors correspond to the channels R - red, G - green, B - blue.

Only the R values from the previously processed pixels are used as the context $Cont_1$ for the red channel value (R). For the context of the green channel $Cont_2$, in addition to the value of the green color component (G) of the previously processed pixels, the R values from the previously processed and current pixels are also used. For the context values of the blue channel $Cont_3$, in addition to the values of the blue color (B) of the previously processed pixels, the values R and G of the previously processed and current pixels are also used.
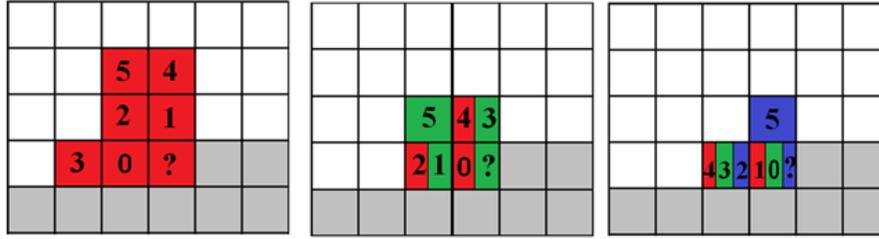
Figure 1. Contexts of the R, G, B channels for the full color coding algorithm (FPC).

To increase the compression ratio, the increment between the corresponding channels for the current and previous pixels is used instead of R, G and B of the current pixel. With this approach, it is possible to use efficiently the correlation between color channels. In the software implementation, each of the three contexts, Cont₁, Cont₂, Cont₃, is a class that includes the active context length and an array of N unsigned 1-byte integers that store the previous values of the corresponding channel for the current pixel, where N is the number of pixels in the context. The contexts Cont₁, Cont₂, Cont₃ are formed according to the formula (1).

$$Cont_\mu(rn_{i,j}^q) = \begin{cases} i_0 = i - t_k + b_k \\ j_0 = j - k \\ t_k = p - k; \\ k = 0,1,...p_\mu; p_{\mu=1} = 2; p_{\mu=2,3} = 1 \\ b_0^{\mu=1} = 0,1,2,...,t_0 - 1; b_1^{\mu=1} = 0,1,2,...2t_1; b_2^{\mu=1} = 0; \\ b_0^{\mu=2} = -1,0,0; b_1^{\mu=2} = 0,0,1; \\ b_0^{\mu=3} = -1,-1,0,0,0; b_1^{\mu=3} = 0; \end{cases} \quad (1)$$

$$i = 0,1,2...m-1; j = 0,1,2...n-1$$
$$q_{\mu=1} = 1; q_{\mu=2} = 1,2,1,2,1,2; q_{\mu=3} = 1,2,3,1,2,3$$
$$\mu = \{1,2,3\}$$

where *m, n are* the width and height of the compressed image, i_0, j_0 are the coordinates of the individual context pixels, $rn_{i,j}^q$ is the current encoded element, p is the height of the context in pixels, q is the channel number (1-red, 2-green, 3-blue), μ is the number of the channel to be coded.

6. Conclusion
The problem of compression of RGB and multispectral images by context-based methods has been considered. The logic of the algorithm has been examined using the example of full-color images as a particular case of multispectral images. The channels that form the image are divided into two groups: main and additional (detecting) channels. A distinguishing feature of the main channels is a significant correlation between neighbors. Variants of prediction from the adjacent channel for the main and additional channels for lossless image compression have been considered. In the experiment on a series of images of different contents, the proposed algorithm showed a superior compression ratio in comparison with the popular WinRar, 7z, PNG archivers for all the considered variants of context-based prediction. The leader among the popular archivers, JPEG-LS, was surpassed by our algorithm in the record configuration 2b (and also in the non-record configurations 1a, and 1b) on the image from the Landsat series by 40% with a compression ratio of 2.82 versus 1.99. The best results were demonstrated with the approach, when one channel was used as a reference (master) channel and was compressed independently, and each subsequent channel used the values of pixels from previous channels to form the context. The compression ratio was further increased by using 1-byte increments for channels instead of channel values, while the use of channel values in the context was preserved. We expect to continue research on a wider sample of images and to use this algorithm to compress multispectral images with a greater number of channels.

7. References

[1] Cleary J, and Witten I 1984 Data compression using adaptive coding and partial string matching IEEE *Transactions on Communications* **32(4)** 396-402

[2] Mahoney M 2005 Adaptive weighing of context models for lossless data compression *Florida Tech. Technical Report* **CS-16**

[3] Ageenko E 2000 *Contex-based Compression of Binary Images* (University of Joensuu, Finland) p 120

[4] Ageenko E 2002 Context-based compression of binary images in parallel *Journal Software - Practice & Experience* **32(13)** 1223-1237 DOI 10.1002/spe.480

[5] Marpe D 2001 Video Compression Using Context-Based Adaptive Arithmetic Coding *International Conference on Image Processing* **D-10587** 558-561 DOI: 10.1109/ ICIP.2001.958175

[6] Zhang Y and Adjeroh D A 2008 Prediction by partial approximate matching for lossless image compression *IEEE Trans. Image Process.* **17(6)** 924-935

[7] Knoll B and de Freitas N 2011 *A Machine Learning Perspective on Predictive Coding with PAQ* (University of British Columbia, Vancouver, Canada)

[8] Gashnikov M V, Glumov N I and Sergeev V V 2010 A hierarchical compression method for space images *Automation and Remote Control* **3** 147-161

[9] Gashnikov M V and Glumov N I 2014 Hierarchical compression for hyperspectral image storage *Computer Optics* **38(3)** 482-488

[10] Gashnikov M V and Glumov N I 2016 Onboard processing of hyperspectral data in the remote sensing systems based on hierarchical compression *Computer Optics* **40(4)** 543-551 DOI: 10.18287/2412-6179-2016-40-4-543-551

[11] Yusov E and Turlapov V 2008 JPEG2000-based compressed multiresolution model for real-time large-scale terrain visualization *Int. Conf. on Computer Graphics and Vision, Proceedings* 8 (Access mode: www.graphicon.ru/html/2008/proceedings/English/S8/Paper_1.pdf)

[12] Borusyak A V and Vasin Yu G 2015 Algorithm for compression of indexed images using context-based modelling *Proceedings of the 9th Open German-Russian Conference on Image Recognition and Understanding* 60-62

[13] Borusyak A V and Vasin Yu G 2016 Development of the algorithm for adaptive compression of indexed images using context-based modelling *Pattern Recognition and Image Analysis (Advances in Mathematical Theory and Applications* **26(1)** 4-8

[14] Vasin Yu G and Borusyak A V 2016 Compression of large-format images by means of statistical coding using context-based modeling *Proceedings of International Scientific Conference Situational centers and class 4i information-analytical systems for monitoring and security tasks* 274-278

Acknowledgements