DPCM with an adaptive extrapolator for image compression

M.V. Gashnikov[1]

[1]*Samara National Research University, 34 Moskovskoe Shosse, 443086, Samara, Russia*

Abstract

The method of image compression based on differential pulse-code modulation (DPCM) with an adaptive extrapolator is investigated. This extrapolator automatically adapts to local features of contours (boundaries) in the image. The issue of the negative effect of quantization on the result of optimizing the adaptive extrapolator is investigated. It is experimentally proved that, despite this effect, the adaptive extrapolator has the advantage over prototypes. Also, an experimental research of the considered method is carried out as a whole; a comparison is made with the JPEG method with respect to the maximum error in a half-tone Waterloo set of natural images.

Keywords: image compression; DPCM; extrapolation; quantization; entropy coding; compression ratio, maximum error

1. Introduction

Images have extremely large data sizes. The growth of data storage capacity does not solve this problem. First of all, this applies to multi- and hyperspectral images [1]. Due to the limited available resources, this problem is especially important for on-board image observation systems located on satellite and other aircraft, including atmospheric drones. So we have to use effective, often specialized, methods of image compression.

The number of approaches to image compression is very large [2-6]. Fractal compression methods [15] have the highest compression ratio, but they have not been widely used because of their computational complexity and unnatural distortion of the images. Methods based on wavelet transforms [8] are most preferable in the "efficiency/complexity" coordinates and have the widest possible scope. The method JPEG-2000 [9] is the most common of the wavelet compression methods. Methods based on two-dimensional discrete orthogonal transformations [10] have similar advantages. The method JPEG [11] is the most common of these compression methods. The JPEG method loses [12] a wavelet method in terms of efficiency, but JPEG significantly exceeds JPEG-2000 prevalence due to a lot of software and hardware in use.

However, all transformation-based compression methods require a lot of computing resources. Therefore, these methods are difficult to use in real-time systems, including on-board systems. In addition, such systems usually impose increased requirements for the control of the compressed data quality. But this is very problematic for the mentioned compression methods, because of the difficulties of controlling the compression error in the space of transformation coefficients.

Thus, in the situation of strictly quality control of compressed data and the restrictions on available resources, we need compression methods that do not use any spectral spaces. So these methods have to produce all the processing in the source brightness space. This allows providing low computational complexity and providing quality control of compressed information.

Such methods include differential compression methods [2-3], which decorrelate the signal by using the difference representation of this signal. This class of methods includes hierarchical methods of compression [13-14], which have a number of important advantages when used in on-ground image processing complexes. However, with on-board processing the hierarchical methods have no special advantages, but they have a high structural complexity. Therefore, in the opinion of the author, difference methods based on differential pulse code modulation (DPCM) [2-3] are the most preferable for on-board systems. In addition, the scope of DPCM method is not limited to real-time systems. For example, DPCM is included in other compression methods, such as JPEG. Thus, the task of research and increasing the effectiveness of compression methods based on DPCM is still relevant.

DPCM compression is based on extrapolation (prediction) of image pixels and coding of the difference between the original and extrapolated values of pixels. In [15], a compression method based on DPCM with an adaptive extrapolator is proposed, which is optimized on the basis of the minimum absolute value of extrapolation error. However, the paper [15] does not close a number of important issues related to the investigation of this compression method.

First of all, the effect of quantization on the optimizing the parameters of the adaptive extrapolator is not taken into account. In addition, the results of the extrapolator research are not given, and the conclusion about its effectiveness is made on the basis of research of the compression method as a whole. The present work aims to close these gaps in the research of this compression method, to draw substantiated conclusions about its effectiveness and to develop recommendations for its use.

2. Image compression based on differential pulse code modulation

Here is a brief simplified description of the image compression method based on DPCM. Let the non-negative integer pixels $x(m,n)$ of the original digital image be processed line by line. Let's designate $\bar{x}(m,n)$ the pixels of the decompressed (restored after compression) image. We calculate these values already at the stage of compression when processing the corresponding pixels for organizing the feedback. For each pixel $x(m,n)$, we calculate its extrapolated value $\hat{x}(m,n)$ using an extrapolator $P(...)$ based on the restored values $\bar{x}(m,n)$ of the already processed pixels:

$$\hat{x}(m,n) = P\left\{ \overline{x}(i,j) : i < m \ \ or \ \ i = m \ \ and \ \ j < n \right\}.$$ (1)

Then, an extrapolated value $\hat{x}(m,n)$ is subtracted from the original pixel value $x(m,n)$ to calculate the difference signal $f(m,n)$. Then, the difference signal $f(m,n)$ is quantized $Q(f)$, the result of this quantization is a quantized difference signal $\overline{f}(m,n)$. This signal is encoded and transmitted through a communication channel or sent to an archive file. The quantized signal $\overline{f}(m,n)$ is immediately used to calculate the corresponding recovered pixel value $\overline{x}(m,n)$:

$$f(m,n) = x(m,n) - \hat{x}(m,n), \quad \overline{f}(m,n) = Q(f(m,n)), \quad \overline{x}(m,n) = \overline{f}(m,n) + \hat{x}(m,n).$$ (2)

The restored pixel value $\overline{x}(m,n)$ is used to extrapolate (1) the next pixel.

3. Extrapolation for image compression based on DPCM

The requirement of low computational complexity makes it necessary to apply in DIKM only the simplest [16] extrapolators of the form:

$$\hat{x}^{(0)}(m,n) = \overline{x}(m-1,n),$$ (3)

$$\hat{x}^{(1)}(m,n) = \frac{1}{2}\left(\overline{x}(m-1,n) + \overline{x}(m,n-1) \right),$$ (4)

$$\hat{x}^{(2)}(m,n) = \overline{x}(m,n-1).$$ (5)

These linear extrapolators work worst on contours (object boundaries). As an answer to this problem, extrapolators that are invariant to contours are considered, for example Greham's extrapolator [3], which is invariant to vertical and horizontal contours:

$$\hat{x}^{G}(m,n) = \begin{cases} \hat{x}^{(0)}(m,n), & if \ \lambda_m(m,n) < \lambda_n(m,n); \\ \hat{x}^{(2)}(m,n), & if \ \lambda_m(m,n) \geq \lambda_n(m,n), \end{cases}$$ (6)

где

$$\lambda_m(m,n) = \left| \overline{x}(m,n-1) - \overline{x}(m-1,n-1) \right|, \quad \lambda_n(m,n) = \left| \overline{x}(m-1,n) - \overline{x}(m-1,n-1) \right|.$$ (7)

The smaller difference from the differences (7) gives the direction of the contour in a small neighborhood of the current image pixel. The relation (6) provides extrapolation "along" this direction. Such an extrapolator is more accurate on the contours, but on flat sections it loses to the extrapolator (4), which is more stable to noise due to averaging.

The advantages of the "contour" extrapolator (6) and the "averaging" extrapolator (4) are combined by an adaptive extrapolator:

$$\hat{x}^{A}(m,n) = \begin{cases} \hat{x}^{(0)}(m,n), & if \ \lambda(m,n) < \lambda^{(-)}; \\ \hat{x}^{(1)}(m,n), & if \ \lambda^{(-)} \leq \lambda(m,n) \leq \lambda^{(+)}; \\ \hat{x}^{(2)}(m,n), & if \ \lambda^{(+)} < \lambda(m,n), \end{cases}$$ (8)

where $\lambda(m,n)$ is a «contour direction feature»:

$$\lambda(m,n) = \lambda_m(m,n) - \lambda_n(m,n) ,$$ (9)

$\lambda^{(-)}, \lambda^{(+)}$ are parameters of the adaptive extrapolator, which are selected in the ranges:

$$-x_{max} \leq \lambda^{(-)} \leq 0 \leq \lambda^{(+)} \leq x_{max} ,$$ (10)

where x_{max} is the maximum brightness in the image. If the feature (9) is close to zero (the current pixel is in a flat image area), then the averaging extrapolator (4) is used, but if the feature (9) has a large value (positive or negative), then

extrapolation (6) "along the contour" occurs. The extrapolator (8) adapts to each particular image: if the contours in the image are small, then the parameters $\lambda^{(-)}$, $\lambda^{(+)}$ are set far from zero, but if there are many vertical and/or horizontal contours, the corresponding parameter must have a large absolute value.

The parameters $\lambda^{(+)}$, $\lambda^{(-)}$ are automatically calculated before the actual DPCM processing for each particular image. A special optimization procedure for the adaptive extrapolator is used for this (see below). Then the parameters $\lambda^{(+)}$, $\lambda^{(-)}$ are placed in the archive, because they are also necessary for decompression.

4. Optimization of the adaptive extrapolator for image compression based on DPCM

Optimization of the adaptive extrapolator (search of parameters $\lambda^{(+)}$, $\lambda^{(-)}$) is based on minimizing the sum of the absolute values of extrapolation errors over the set $\omega = \{(m,n)\}$ of coordinates of all image pixels:

$$\delta\left(\lambda^{(+)},\lambda^{(-)}\right) = \sum_{(m,n)\in\omega} \left|x(m,n)-\hat{x}(m,n)\right| \to \min_{\lambda^{(+)},\lambda^{(-)}}. \tag{11}$$

The error (11) can be divided into three component parts corresponding to different ranges of the "contour direction feature" (9):

$$\delta\left(\lambda^{(+)},\lambda^{(-)}\right) = \delta^{(-)}\left(\lambda^{(-)}\right) + \delta^{(0)} + \delta^{(+)}\left(\lambda^{(+)}\right). \tag{12}$$

where

$$\delta^{(-)}\left(\lambda^{(-)}\right) = \sum_{(m,n)\in\omega^{(-)}} \left|x(m,n)-\hat{x}(m,n)\right|, \quad \delta^{(0)} = \sum_{(m,n)\in\omega^{(0)}} \left|x(m,n)-\hat{x}(m,n)\right|, \quad \delta^{(+)}\left(\lambda^{(+)}\right) = \sum_{(m,n)\in\omega^{(+)}} \left|x(m,n)-\hat{x}(m,n)\right|,$$

$$\omega = \omega^{(-)} \quad \omega^{(0)} \quad \omega^{(+)}, \quad \omega^{(-)} = \left\{(m,n):\lambda(m,n)<0\right\}, \quad \omega^{(0)} = \left\{(m,n):\lambda(m,n)=0\right\}, \quad \omega^{(+)} = \left\{(m,n):\lambda(m,n)>0\right\}.$$

As a result, the two-parameter optimization problem (11) is decomposed into two one-parameter problems that are solved independently of each other:

$$\lambda^{(+)} = \arg\min_{\lambda} \delta^{(+)}(\lambda), \quad \lambda^{(-)} = \arg\min_{\lambda} \delta^{(-)}(\lambda). \tag{13}$$

To solve these problems, a special matrix is filled in the preliminary pass through the image

$$\Delta_{i,\lambda} = \sum_{(m,n)\in\omega(\lambda)} \left|x(m,n)-\hat{x}^{(i)}(m,n)\right|, \quad 0\le i\le 2, \quad -x_{\max}\le\lambda\le x_{\max}. \tag{14}$$

Each element $\Delta_{i,\lambda}$ of this matrix contains the sum of extrapolation errors of extrapolator number i (3-5) for all pixels for which the value of feature (9) is equal λ. For the filled matrix (14), a one-dimensional array of extrapolation error values $\delta^{(+)}$ is filled using a recurrence procedure:

$$\delta^{(+)}\left(x_{\max}\right) = \sum_{\lambda=0}^{x_{\max}} \Delta_{1,x_{\max}}, \quad \delta^{(+)}(\lambda) = \delta^{(+)}(\lambda+1) + \Delta_{2,\lambda} - \Delta_{1,\lambda}, \quad 0\le\lambda<x_{\max}. \tag{15}$$

The computational complexity of this procedure does not depend on the image size. Optimal value $\lambda^{(+)}$ can be found in this array $\delta^{(+)}(\lambda^{(+)})$ by an exhaustive search (the length of this array is only $x_{\max}+1$). Analogously $\lambda^{(-)}$ is calculated.

5. Quantization for image compression based on DPCM

For quantization in DPCM, the Max scale [5-6] is usually used, which provides the minimum relative root-mean-square error

$$\varepsilon_{rel}^2 = \frac{\varepsilon^2}{D_x} = \frac{1}{MND_x} \sum_{(m,n)\in\omega} \left(x(m,n)-\overline{x}(m,n)\right)^2, \tag{16}$$

where $x(m,n)$ is the original image, $\overline{x}(m,n)$ is the restored (decompressed) image, D_x is the image variance, MxN are image sizes.

For unique data, more strong error control [17] is necessary. For example, it is necessary for compression of hyperspectral images [18-20]. In this case, a quantizer with a uniform scale [3] can be used for DPCM. This quantizer provides maximum error control:

$$\varepsilon_{max} = \max_{(m,n)\in\omega}\left|x(m,n)-\overline{x}(m,n)\right| . \tag{17}$$

6. The effect of quantization on the adaptive extrapolator optimization

It should be noted an important nuance that occurs when optimizing an adaptive extrapolator. The optimization procedure described above for the adaptive extrapolator can be performed only in the absence of quantization (when the quantizer is "switched off"). In this case, the original $x(m,n)$ and recovered $\overline{x}(m,n)$ values of the image pixels are equal, the difference $f(m,n)$ and quantized difference $\overline{f}(m,n)$ signals are also equal. The "activation" of the quantizer at the stage of extrapolator optimization would lead to the impossibility of calculating the restored values of the pixels, since they, through the chain of transformations, depend on the parameters $\lambda^{(+)}$, $\lambda^{(-)}$ of the extrapolator, which at this stage are still unknown.

Thus, the general scheme of DPCM compression with an adaptive extrapolator is as follows. First, to optimize the extrapolator, a preliminary pass through the image with the "switched off" quantizer (i.e., zero-error) is performed. In this case, the extrapolator parameters $\lambda^{(+)}$, $\lambda^{(-)}$ are calculated. After that, the DPCM-compression itself is made, in which the quantizer is "switched on" again, and the parameters $\lambda^{(+)}$, $\lambda^{(-)}$ found on the preliminary pass are used for extrapolation.

As a result, the parameters of the extrapolator, optimal by criterion (11) with zero error, are used in compression with nonzero error. In this situation, these parameters are no longer optimal, and the question of how far from the optimum they are, requires additional research, which can only be experimental. Computational experiments were carried out on the so-called set of halftone images "Waterloo" [21], which is traditionally used for the research of compression methods. Typical results are shown in Fig. 1-2.

The investigation was carried out for one-parameter problems (13). The quantities $\delta^{(+)}$, $\delta^{(-)}$, introduced by the relation (12), which are calculated with zero quantization error, will be referred to below as the "estimative total error". The quantities $\delta^{(+)}$, $\delta^{(-)}$, calculated with a non-zero quantization error, will be referred to below as the "true total error". Thus, it is necessary to answer the question of how closely the minimum of the true total error and the minimum of the estimative total error are located.

The dependence of the total error $\delta^{(+)}$ on the extrapolator parameter $\lambda^{(+)}$ is shown in Fig. 1. The value of the second parameter $\lambda^{(-)}$ was fixed (it was chosen in the optimal way). Then the same research was performed for total error $\delta^{(-)}$. The dependence of the total error $\delta^{(-)}$ on the extrapolator parameter $\lambda^{(-)}$ is shown in Fig. 2. Thus, accordingly, the value of the parameter $\lambda^{(+)}$ was fixed.
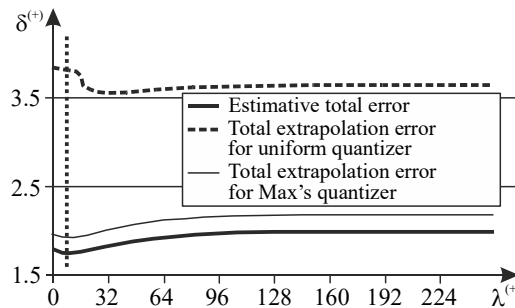

Fig. 1. Dependence of the extrapolation total error $\delta^{(+)}$, corresponding to positive values of the contour direction feature (9), from the extrapolation parameter $\lambda^{(+)}$ for a fixed $\lambda^{(-)} = -19$ (the vertical line shows the position of the minimum of the estimative total error).
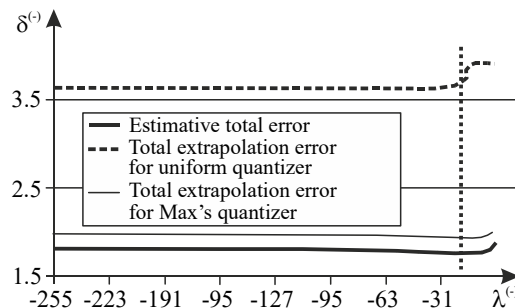

Fig. 2. Dependence of the extrapolation total error $\delta^{(-)}$, corresponding to negative values of the contour direction feature (9), from the extrapolation parameter $\lambda^{(-)}$ for a fixed $\lambda^{(+)} = 7$ (the vertical line shows the position of the minimum of the estimative total error).

These researches allow drawing the following conclusions:

1. Quantization affects the results of optimization of the adaptive extrapolator, since the minimum of the true total error and the minimum of the estimative total error may not be equal
2. When using the Max quantizer, the values found for the parameters of the adaptive extrapolator are closer to the optimum.
3. It is necessary to carry out a research of the extrapolator's efficiency for estimating the effect of a mismatch between the optimums of the true total error and estimative total error.

7. Research of the effectiveness of the adaptive extrapolation algorithm

To evaluate the effectiveness of the adaptive extrapolator for compression, this extrapolator was compared with other extrapolators. The comparison was produced by the entropy H_q of the quantized difference signal. This entropy is a good estimate of compressed data size. The results are shown in Fig. 3-4.
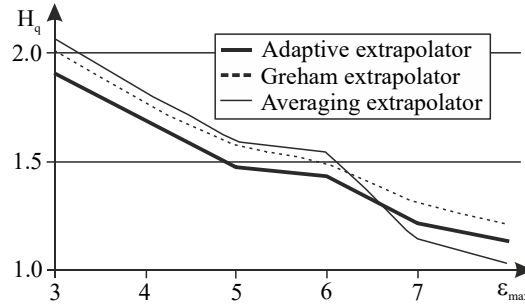
Fig. 3. Dependence of the entropy H_q of the quantized difference signal on the maximum error ε_{max} when using a uniform quantization scale.
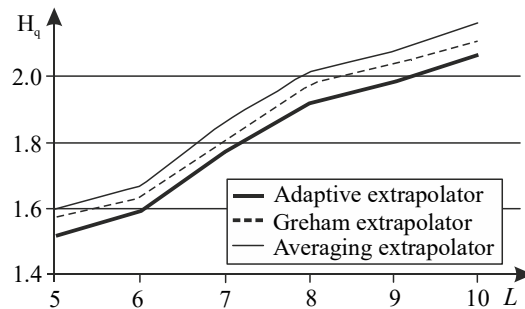
Fig. 4. Dependence of the entropy H_q of the quantized difference signal on the number L of quantized levels when using Max's quantization scale.

Conclusions:
1. The adaptive extrapolator has the advantage over prototypes over the entropy of the quantized signal.
2. Consequently, the negative influence of the quantizer described in the previous section does not have a determining value (at least for small extrapolation errors).
3. With increasing maximum error, the negative influence of the quantizer on the efficiency of the adaptive extrapolator increases. With a maximum error of more than six, adaptive extrapolator loses the advantage.

8. Experimental research of DPCM with an adaptive extrapolator for image compression

To evaluate the efficiency of the image compression method based on DPCM with the adaptive extrapolator, it was compared with method JPEG in the coordinates "error-compression ratio" on the set of images "Waterloo" [21], which is traditionally used for comparison of compression methods. The results obtained, averaged over all images of the set, are shown in Fig. 5. The results of the experiments demonstrate a significant gain (up to two times) of DPCM with an adaptive extrapolator for the JPEG method with respect to the maximum error.
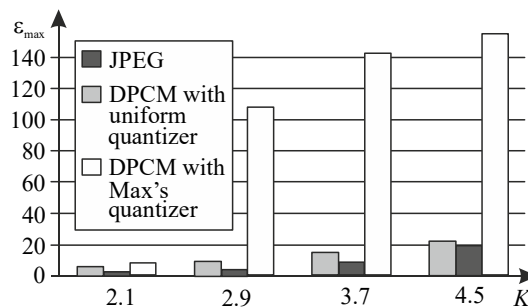
Fig. 5. The dependence of the maximum error ε_{max} on the compression coefficient K_c when comparing DPCM with an adaptive extrapolator versus the JPEG compression method.

9. Conclusion

The significant influence of quantization on the optimization of the parameters of the adaptive extrapolator is described and experimentally confirmed. It has been shown experimentally that, despite this negative effect of quantization, the adaptive extrapolator within the DPCM compression method still has an advantage over prototypes with a small error. Also, computational experiments were carried out to research the efficiency of the DPCM compression method with an adaptive extrapolator and showed its advantage over the JPEG compression method with respect to the maximum error. Based on the obtained results, it can be concluded that the considered method is promising for image compression systems and image transmission systems.

Further research will be aimed at eliminating the need for a preliminary pass through the image when optimizing the adaptive extrapolator, which is necessary to simplify the use of the method of image compression for real-time systems, including on-board systems. The possibility of such an improvement is based on the admissibility of estimating the distribution of extrapolation errors during the actual DPCM processing. The question of the quality of such a "consistently refined" assessment requires additional investigation.

Acknowledgements

References

[1] Chang C. Hyperspectral Data Processing: Algorithm Design and Analysis. Wiley Press, 2013; 1164 p.

[2] Sayood K. Introduction to Data Compression. The Morgan Kaufmann Series in Multimedia Information and Systems, 4ed., 2012; 743 p.

[3] Salomon D. Data Compression. The Complete Reference. Springer-Verlag, 4ed, 2007; 1118 p.

[4] Vatolin D, Smirnov M, Yukin V. Data compression methods. Archive program architecture, image and video compression. Moscow: "Dialog-MIFI" Publisher 2002; 384 p. (in Russian)

[5] Woods E, Gonzalez R. Digital Image Processing. Prentice Hall, 3ed, 2007; 976 p.

[6] Pratt W. Digital image processing. Wiley, 4ed, 2007; 807 p.

[7] Woon WM, Ho ATS, Yu T. Tam SC, Tan SC, Yap LT. Achieving high data compression of self-similar satellite images using fractal. Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS) 2000: 609–611.

[8] Gupta V, Sharma V, Kumar A. Enhanced Image Compression Using Wavelets. International Journal of Research in Engineering and Science (IJRES) 2014; 2(5): 55–62.

[9] Li J. Image Compression: The Mathematics of JPEG-2000. Modern Signal Processing. MSRI Publications 2003; 46: 185–221.

[10] Plonka G, Tasche M. Fast and numerically stable algorithms for discrete cosine transforms. Linear Algebra and its Applications 2005; 394(1): 309–345.

[11] Wallace G. The JPEG Still Picture Compression Standard. Communications of the ACM 1991; 34(4): 30–44.

[12] Ebrahimi F, Chamik M, Winkler S. JPEG vs. JPEG2000: An Objective Comparison of Image Encoding Quality. Proceedings of SPIE Applications of Digital Image Processing XXVII 2004; 5558: 300–308.

[13] Gashnikov M. Interpolation for hyperspectral images compression . CEUR Workshop Proceedings 2016; 1638: 327–333.

[14] Gashnikov M, Glumov NI. Development and Investigation of a Hierarchical Compression Algorithm for Storing Hyperspectral Images. Optical Memory and Neural Networks. Allerton Press 2016; 25(3): 168–179.

[15] Gashnikov M, Mullina SS. Adaptive parameterized predictor for differential image compression. Proceedings of the International Conference "Information Technologies and Nanotechnologies". Samara 2015: 64–67. (in Russian)

[16] Efimov V, Kolesnikov AN. Effectiveness estimation of the hierarchical and line-by-line lossless compression algorithms. Proceedings of the III conference "Pattern recognition and image analisys". Nijni Novgorod 1997; 1: 157–161. (in Russian)

[17] Lin S, Costello D. Error Control Coding: Fundamentals and Applications, second edition. New Jersey: Prentice-Hall, inc. Englewood Cliffs, 2004; 1260 p.

[18] Chang C. Hyperspectral data exploitation: theory and applications. Wiley-Interscience, 2007; 440 p.

[19] Gashnikov MV, Glumov NI. Hierarchical GRID Interpolation under Hyperspectral Images Compression. Optical Memory and Neural Networks (Information Optics). Allerton Press 2014; 23(4): 246–253.

[20] Borengasser M, Hungate W , Watkins R. Hyperspectral Remote Sensing – Principles and Applications. CRC Press, 2004; 128 p.

[21] Waterloo Grey Set. University of Waterloo Fractal coding and analysis group: Mayer Gregory Image Repository. URL: http://links.uwaterloo.ca/ Repository.htm (19.12.2016).

[22] Gashnikov MV, Glumov NI. Onboard processing of hyperspectral data in the remote sensing systems based on hierarchical compression. Computer Optics 2016; 40(4): 543–551. (in Russian)