

DOI: <https://doi.org/10.60797/IRJ.2024.146.35>

METHODS AND MEANS OF IMAGE PROCESSING FOR TRANSMISSION IN SPECTRAL REPRESENTATION

Research article

Kostrov B.V.¹, Baranova S.N.^{2,*}, Vyugina A.A.³, Trushina Y.A.⁴, Panina I.S.⁵¹ ORCID : 0009-0005-6557-5192;² ORCID : 0000-0002-2984-4480;³ ORCID : 0000-0002-0087-5988;⁴ ORCID : 0000-0001-6019-225X;⁵ ORCID : 0009-0004-6789-419X;^{1, 2, 3, 4, 5} Ryazan State Radio Engineering University named after V.F. Utkin, Ryazan, Russian Federation

* Corresponding author (baranova.sv.n[at]gmail.com)

Abstract

Images are a large part of all transmitted data. The purpose of the article is to explore the possibilities of using image processing methods and tools to reduce the information load on the channel and preserve the content component. Spatial and spatio-spectral representations of the image and methods of their processing are considered, these are the effectiveness of compression and the possibility of recovery using machine learning algorithms. As a result, it was determined that the combination of the described methods makes it possible to effectively apply image compression at the preprocessing stage without loss of informativeness. The paper summarizes the results of experiments, image parameters during processing and draws appropriate conclusions.

Keywords: image processing, data transfer, machine learning, algorithms of compression, Walsh functions, RLE, MATLAB.

МЕТОДЫ И СРЕДСТВА ОБРАБОТКИ ИЗОБРАЖЕНИЙ ДЛЯ ПЕРЕДАЧИ В СПЕКТРАЛЬНОМ ПРЕДСТАВЛЕНИИ

Научная статья

Костров Б.В.¹, Баранова С.Н.^{2,*}, Вьюгина А.А.³, Трушина Е.А.⁴, Панина И.С.⁵¹ ORCID : 0009-0005-6557-5192;² ORCID : 0000-0002-2984-4480;³ ORCID : 0000-0002-0087-5988;⁴ ORCID : 0000-0001-6019-225X;⁵ ORCID : 0009-0004-6789-419X;^{1, 2, 3, 4, 5} Рязанский государственный радиотехнический университет им. В.Ф. Уткина, Рязань, Российская Федерация

* Корреспондирующий автор (baranova.sv.n[at]gmail.com)

Аннотация

Изображения являются значительной частью всех передаваемых данных. Цель статьи – изучить возможности использования методов и средств обработки изображений для снижения информационной нагрузки на канал и сохранения контентной составляющей. Рассматриваются пространственные и пространственно-спектральные представления изображения и методы их обработки, эффективность сжатия и возможность восстановления с помощью алгоритмов машинного обучения. В результате было установлено, что комбинация описанных методов позволяет эффективно применять сжатие изображений на этапе предварительной обработки без потери информативности. В работе обобщены результаты экспериментов, параметры изображений при обработке и сделаны соответствующие выводы.

Ключевые слова: обработка изображений, передача данных, машинное обучение, алгоритмы сжатия, функции Уолша, RLE, MATLAB.

Introduction

Raster and vector format data play a major role in the modern world, and therefore it is necessary to solve problems related to their transmission over communication channels. The main task in this case is to reduce the amount of transmitted data. At the stage of preparing data for transmission, encoding algorithms can be used to ensure the reliability of transmission, and compression algorithms to reduce the amount of transmitted data [1], [2].

When transferring images, it is especially important to preserve the visual component, since the information content of the image depends on these characteristics. In this case, the amount of lost information depends on the location of the distorted or lost data [1]. One of the actual tasks in the field of image processing is to preserve the content component during transmission over communication channels.

The purpose of the article is to explore the possibilities of using image processing methods and tools to reduce the information load on the channel and preserve the content component.

Image representation options

When processing images, various representations of them are often used, which allows you to use their certain properties. A separation like that can be performed, for example, in the color model: RGB, CMYK, HSB. The article considers image representation models that provide the process of transmitting them over a communication channel:

- 1) Spatial representation;
- 2) Spatio-spectral representation based on Walsh functions;
- 3) Spatio-spectral representation based on pseudorandom sequences.

Spatial representation implies a digital image, which is defined by a discrete array of numbers that make up the $N \times N$ (1) matrix G . For grayscale images, each pixel takes a value from 0 to 255.

$$\mathit{G} = \begin{pmatrix} g_{1,1} & \cdots & g_{1,N} \\ \vdots & \ddots & \vdots \\ g_{N,1} & \cdots & g_{N,N} \end{pmatrix}, \quad (1)$$

where $g_{\{i,j\}}$ – a single pixel of a digital image, (i,j) – the indices of the pixel position in the image matrix, $i = \overline{1, N}; j = \overline{1, N}$.

When translating an image into a spatio-spectral form, the relationship between the values of the image elements is manifested [3]. Let us consider a spatial spectral representation based on Walsh functions.

The image is a two-dimensional discrete function; therefore, it is possible to use orthogonal unitary transformations. Due to a direct unitary transformation of the matrix of the original image, it is possible to obtain a spatio-spectral representation of an image of the same size, the elements of which are calculated using the formula (2) [4].

$$F(u, v) = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} g(i, j) \varphi_{(u,v)}(i, j) \quad (2)$$

where $\varphi_{(u,v)}(i, j)$ – a set of orthonormal functions (the kernel of the direct transformation), u, v – variables of the transformation space.

The original image can be obtained using the inverse transformation (3).

$$g(i, j) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) \varphi_{(i,j)}^{-1}(u, v) \quad (3)$$

where $\varphi_{(i,j)}^{-1}(u, v)$ – the kernel of the inverse transformation.

The Fourier transform and the Walsh transform are often used for the transition to the spatio-spectral domain. The Fourier transform is a common method of analyzing periodic functions and signals, which allows you to represent a function as a sum of sines and cosines. Unlike trigonometric functions, Walsh functions take only two values: +1 and -1.

In this paper, the Walsh functions are used as a non-trigonometric orthogonal basis of the function. The Walsh transform has faster execution compared to the Fourier transform, which makes it more efficient for real-time signal processing, and it is also more resistant to noise and signal distortion [5].

The direct transformation in matrix form is described by the expression (4).

$$\mathit{F} = \mathit{H} \mathit{G} \mathit{H}^T, \quad (4)$$

where G – image matrix, H – Hadamard matrix, ordered by Walsh, F – the spectrum of the image.

The inverse transformation takes the form:

$$\mathit{G} = \frac{1}{N^2} \mathit{H}^T \mathit{F} \mathit{H} \quad (5)$$

where G – image matrix, H – Hadamard matrix, ordered by Walsh, F – the spectrum of the image.

The complexity of calculating the spectrum can be reduced if orthogonal functions are applied only to rows or columns of the image, in that case the direct transformation to the spatio-spectral transformation looks like [3]:

$$\mathit{F} = \mathit{H} \mathit{G}, \quad (6)$$

where G – image matrix, H – Hadamard matrix, ordered by Walsh, F – the spectrum of the image.

Since the matrix H is orthogonal and symmetric, the inverse transformation is written as:

$$\mathit{G} = \frac{1}{N} \mathit{H} \mathit{F}, \quad (7)$$

where G – image matrix, H – Hadamard matrix, ordered by Walsh, F – the spectrum of the image.

One of the disadvantages of using the Walsh function basis is a significant unevenness of the energy distribution over the spectral components.

Let us consider a spatio-spectral representation based on pseudorandom sequences (hereinafter referred to as SSR).

Noise similarity (randomness of state change) in such sequences guarantees the absence of sudden power drops in the conversion interval and the absence of overloads in the channel.

SSRs are generated by a shift register with linear feedback [6] and are based on irreducible generating polynomials (8).

$$M(x) = \sum_{i=1}^N a_i x^i \tag{8}$$

where $a_N, a_1 = 1$; $a_i \in [0, 1]$; N - the number of digits of the shift register.

It is necessary to make sure that the polynomial sets the initial state $a_1 = 1$ in the lowest digit, otherwise a null sequence is formed.

Any generating polynomial can be matched with a Fibonacci generator (9).

$$\begin{cases} g_1(k+1) = \sum_{i=1}^N a_i g_i(k), \\ g_i(k+1) = g_{j-1}(t), j = 2, N, \end{cases} \tag{9}$$

where $g_i(t)$ - the state of the i -th digits of the shift register on the clock $k, i = \overline{1, N}, k = \overline{1, 2^N - 1}$.

The algorithm for forming a sequence in matrix form is described by the expression (10).

$$\mathbf{G}(k+1) = \Psi * \mathbf{G}(k), \tag{10}$$

where $\Psi = \begin{bmatrix} a_1 & a_2 & \dots & a_{N-1} & a_N \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}$.

The considered SSRs have certain properties [7] and are called maximum period sequences or M-sequences. The number of different polynomials generating an M-sequence depends on its period $M = 2^n - 1$.

To build a noise-like transformation in order to further transmit images over communication channels, it is necessary to form a transformation matrix by cyclically shifting the original SSR. The number of rows of the matrix is equal to the period, the size of the matrix, respectively, $M \times M$.

The constructed matrix must meet the requirements of orthogonality, it is necessary to align the number of single and zero characters on the period of the SSR and switch to the alphabet $\{-1, 1\}$ in the rows and columns of the matrix. For a matrix formed in this way, a pair of transformations (11) can be written in the following way:

$$\begin{cases} \mathbf{F}_{\mathbf{B}}^{\mathbf{M}} = \mathbf{M}_0^{\mathbf{M}} \mathbf{B}_{\mathbf{M}}^{\mathbf{M}} \\ \mathbf{B}_{\mathbf{M}}^{\mathbf{M}} = \mathbf{F}_{\mathbf{M}}^{\mathbf{M}} \mathbf{M}_0^{\mathbf{M}T} \end{cases} \tag{11}$$

$\mathbf{F}_{\mathbf{M}}^{\mathbf{M}}$ - matrix (matrix-string) of signal conversion coefficients; $\mathbf{B}_{\mathbf{M}}^{\mathbf{M}}$ - the matrix (column matrix) of the signal; $\mathbf{M}_0^{\mathbf{M}T}$ - the transposed transformation matrix.

Figure 1 shows the result of applying the transition to a spatio-spectral representation based on Walsh and SSR functions for transmitting satellite images over a communication channel.

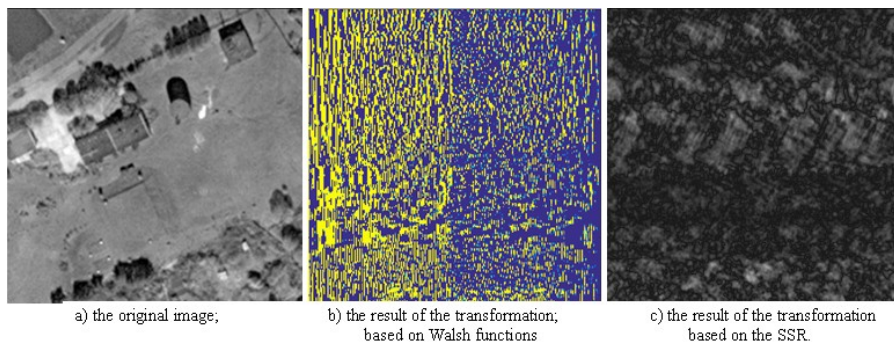


Figure 1 - Using an orthogonal transformation based on a pseudorandom sequence for a satellite image

DOI: <https://doi.org/10.60797/IRJ.2024.146.35.1>

The process of image transmission

When describing the transmission process, it is customary to highlight steps before transmission (preprocessing) and after transmission (postprocessing) [8], [9]. The study dwells on the methods of transformation to the frequency form of images and compression of the obtained data, as well as some recovery methods using machine learning algorithms.

The image transfer process is shown in Figure 2. As input, the model takes an image and transforms it either on the basis of Walsh functions or pseudorandom sequences. Matrices in the spatio-spectral representation are compressed. After the transfer, operations are performed to obtain the original image. The losses that occur during compression, transmission, and performing transformations to the spatio-spectral representation are corrected by machine learning methods.

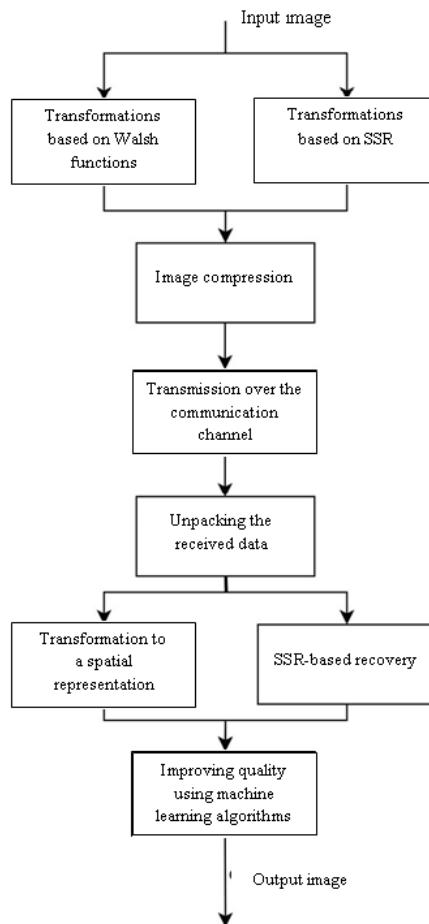


Figure 2 - The process of image transmission
DOI: <https://doi.org/10.60797/IRJ.2024.146.35.2>

Image compression

Compression methods can be used for the convenience of storing and transferring images. In general, they are divided into 2 categories: lossless compression and lossy compression. The first of them assumes the presence of an inverse algorithm that allows getting the original image in its original form.

The simplest and most common example when working with images is the RLE algorithm. Its essence is to replace a sequence of identical values with a pair, where the number of repetitions of the element is indicated first, and then the value itself [10]. In the case of an image in any of its representations (spatial or spatio-spectral), the values can be represented in binary code, which allows encoding sequences consisting of 0 and 1.

Since the RLE algorithm is a lossless compression algorithm, and the transition to a spatio-spectral representation implies rounding operations, which means some information is lost during calculations, it is possible to add to the transformation the clipping of some values below the threshold [11]. Thus, values outside the allowed range take the value 0.

The study of compression with truncation of threshold values was carried out in the Matlab environment [12] on images corresponding to the natural landscape, characterized by a small difference in brightness.

After applying the compression algorithm with truncation of values not exceeding the threshold, the restored image will have certain deviations from the original one. Figure 4 shows an image after compression with a truncation threshold of 5 for a spatial spectral representation based on Walsh functions. The performed truncation allowed us to increase the percentage of compression, and due to the conversion to a spatio-spectral representation, the image details were not distorted, and the visual component did not reduce the information content of the data. The increase in standard deviation (SD) is determined by a change in the spectral component and is visually expressed in a change in the tonality of the image: in the original image, most brightness values are in the range from 120 to 130, and the restored image is from 130 to 150.



Figure 3 - Images using compression of the spatial spectral representation based on Walsh functions
DOI: <https://doi.org/10.60797/IRJ.2024.146.35.3>

Experiments with thresholds 3 and 7 were also performed for the spatio-spectral representation based on Walsh functions. In the first case, the SD is 5.78, and the compression ratio is 38%. In the second case, the SD is 15.73, and the compression ratio is 61%. With an increase in the threshold value, the compression ratio increases less compared to an increase in SD, therefore it is advisable to use a threshold of 5, distortions at which are expressed by shifting brightness values to lighter ones, and compression reduces the amount of data by almost 2 times.

An experiment for compression with a cut-off threshold of 20 for a spatio-spectral representation based on the SSR is shown in Figure 5.



Figure 4 - Images using compression of the spatio-spectral representation based on the SSR
DOI: <https://doi.org/10.60797/IRJ.2024.146.35.4>

Experiments with thresholds of 10 and 30 were also carried out. In the first case, the SD is 3.2248, and the compression ratio is 29.2892%. In the second case, the SD is 13.09999, and the compression ratio is 68.1915%.

For the convenience of comparing the results of experiments on image restoration after compression with threshold truncation, it is advisable to choose a truncation threshold for each type of transformation, where the compression ratio is approximately the same. For Walsh functions, the threshold is 5, and for SSR – 20.

Image recovery

The problem of smoothing some areas in the image that arises when using compression algorithms with truncation can be solved by an approach to image restoration based on the use of machine learning algorithms. The models being developed usually use repeating sections of different scales within a single image, or study comparisons between pairs of low- and high-resolution image fragments in external databases. Since these models are trained by minimizing the SD, the results are usually blurred, and do not contain high-frequency textures due to the above regression to the mean value.

The reason for this result is related to the choice of the objective function, which is used by modern methods. Most systems minimize the pixel-by-pixel mean square error (MSE) between the original and the restored image [13]. The results, meanwhile, do not correlate well with the perception of the resulting image, and the task of restoring natural images is an urgent, difficult task.

To solve this problem, models of convolutional neural networks have been proposed, which based on the feature map allow recognizing and working with objects in the image, it leads to clearer results despite the low values of the peak signal-to-noise ratio (PSNR).

Among the algorithms that use convolutional neural networks to restore images, the pixelNN and EnhanceNet-PAT projects are noted. PixelNN tries to recreate the lost image details using a neural network pre-trained on a large collection of images, synthesizing the missing elements based on matching optimal matches with previously processed images [15].

The EnhanceNet-PAT algorithm specializes in performing operations to increase the resolution of images and improve the quality of photos, for example, allowing you to improve the quality of images from web cameras [16]. This neural network is trained on a large number of high-quality images. During the analysis, similar textures are identified and data accumulated in the process of machine learning is used to improve their quality.

The method is based on the nearest neighbor (NN) approach, which reconstructs a high-resolution photo. At the first stage, a convolutional neural network is used to map the input data to an overly smoothed image, and at the second stage, the nearest neighbor pixel method is used to map the smoothed output data to several high-quality high-frequency outputs in a controlled manner.

One of the proposed algorithms, Deep Image Prior, uses two of the most popular image recovery methods, these are machine learning algorithms and methods based on probabilistic repetition of the contents of neighboring areas [14]. Unlike other machine learning algorithms, this model allows restoring images using data directly from the processed image, rather than data from a training sample. The neural network is used as a generator, that is initially initialized with random data, which is then gradually adjusted and based on statistical information extracted from the processed image.

After transmitting an image that has been subjected to lossy compression, it is possible to use machine learning to improve the quality of the received data. For images restored to a spatial representation, the best result is shown by using the Deep Image Prior algorithm. This network is suitable to perform standard image processing tasks such as noise reduction, rendering, and improvement of detail and resolution.

The described algorithm is a parametric function $x = f(z)$, which maps the code vector z to the image x . The network has a standard structure, the architecture is presented in the form of a U-Net-like\hourglass with several million parameters.

Figure 6 shows the results of quality improvement for images subjected to compression in the spatio-spectral representation based on Walsh and SSR functions. The use of machine learning methods allows you to improve the result of the received image, using compression with clipping by a threshold value.

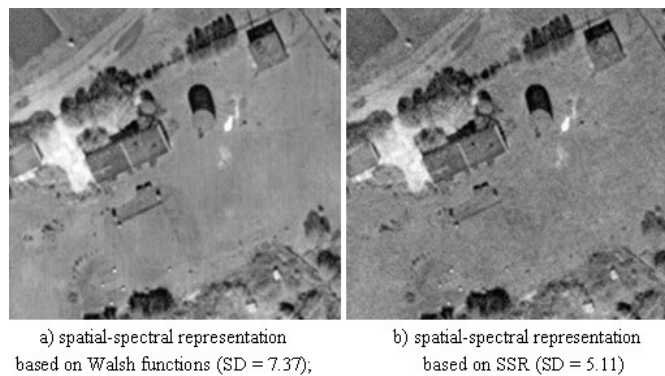


Figure 5 - Improved image by machine learning methods
DOI: <https://doi.org/10.60797/IRJ.2024.146.35.5>

Conclusion

The experimental part was conducted in the Matlab development environment [12]. For this purpose, separate program modules have been developed for converting to a spatio-spectral representation, truncating values according to a set threshold, compressing and restoring images, including decompression and conversion to a spatial representation. The model can be constructed depending on the experiment with different frequency forms of representation and variable threshold values.

For the experimental part, an array of different images containing different types of landscape was compiled. This type of image is interesting because despite the presence of a large number of details (for example, a forest area where each tree can be considered as a separate object), the brightness distribution on them ranges from 50 to 150 in shades of gray.

As a metric of the efficiency of using truncation during compression, the number of bits per pixel was calculated (12). This expression takes into account the decrease in the array of initial values of the matrix and determines the allocation of memory for each element.

$$\text{count} = \frac{N*N*8-K*8}{N*N}, \quad (12)$$

where N – the length and width of the image, K – the number of identical values in the image, 8 – the constant value of the number of bits allocated per value.

The metric is presented only for the spatial-spectral representation, since truncation of values in the spatial representation does not give effective results during reconstruction.

For a spatial spectral representation based on Walsh functions, we present the calculation of the number of bits per pixel before and after truncation. Initially, 8 bits are allocated for each element. After truncation by the threshold value for the spatial spectral representation based on Walsh functions, we get a value of 3.9 bits per pixel, which is about 2 times lower than before truncation. At the same time, being in the frequency domain avoids the appearance of anomalies that reduce the information content of the image.

$$\text{count} = \frac{256*256*8-33516*8}{256*256} = 3,9$$

After truncation by the threshold value for the spatio-spectral representation based on the SSR, we get a value of 3.87 bits per pixel, which is also about 2 times lower than before truncation.

$$\text{count} = \frac{256 \cdot 256 \cdot 8 - 33801 \cdot 8}{256 \cdot 256} = 3,87$$

Acknowledgment

In the course of the conducted research, it was revealed that the constructed model of the image transmission process allows reducing the information load on the channel through the use of preprocessing methods. At the same time, using a lossless compression algorithm preserves the values of the original matrix.

Truncating the threshold values does not reduce the amount of information obtained during visual image analysis.

Machine learning methods allow correcting anomalies that may occur when restoring an image to a spatial representation.

Further development of research is aimed at exploring the possibilities of lossy compression algorithms, as well as improving the performance of machine learning algorithms that take into account statistical relationships in spatial-spectral representation, in the tasks of increasing the detail and resolution of images after their transmission with distortions.

Конфликт интересов

Не указан.

Рецензия

Все статьи проходят рецензирование. Но рецензент или автор статьи предпочли не публиковать рецензию к этой статье в открытом доступе. Рецензия может быть предоставлена компетентным органам по запросу.

Conflict of Interest

None declared.

Review

All articles are peer-reviewed. But the reviewer or the author of the article chose not to publish a review of this article in the public domain. The review can be provided to the competent authorities upon request.

Список литературы на английском языке / References in English

1. Kostrov B.V. Parallel computations in problems of reconstruction of distorted images in spatial-spectral form / B.V. Kostrov, N.N. Grinchenko, A.A. Vyugina et al. // Proceedings of the Institute for System Programming of the RAS. — 2023. — Vol. 35, No. 2. — P. 157-168. — DOI 10.15514/ISPRAS.2023.35(2).11. — EDN ZHDSZD.
2. Kostrov B.V. Method of eliminating distortions in images presented as bit planes / B.V. Kostrov, A.A. Vyugina, S.N. Baranova // XXI century: results of the past and problems of the present plus. — 2023. — Vol. 12. — No. 3(63). — p. 93-99. — EDN UICDXK.
3. Asaev A.S. The use of spectral images / A.S. Asaev, B.V. Kostrov, V.V. Kostrov // Radio Engineering. — 2008. — No. 9. — p. 99-102.
4. Ahmed N. Orthogonal transforms in digital signal processing / N. Ahmed, K.R. Rao; edited by I. B. Fomenko. — M.: Svyaz. — 1980. — 248 p.
5. Zlobin V.K. Processing of aerospace images / V.K. Zlobin, V.V. Eremeev. — M.: FIZMATLIT. — 2006. — 288 p.
6. Sumenkov N.A. Application of noise-like orthogonal transformation for image transfer and processing / N.A. Sumenkov, B.V. Kostrov, N.N. Grichuk et al. // Bulletin of the YAHMS AAW. — 2021. — No. 2. — P. 168-175.
7. Sumenkov N.A. Construction of an orthogonal basis based on pseudorandom sequences / N.A. Sumenkov, B.V. Kostrov, N.N. Grinchenko et al. // Bulletin of the Concern VKO "Almaz — Antey". — 2020. — No. 4. — P. 95—100. — DOI: 10.38013/2542.0542.2020.4.95.100.
8. Grinchenko N.N. Mathematical model of the image transmission process based on bit planes / N.N. Grinchenko, S.N. Baranova, M.A. Lobachev et al. // Bulletin of the concern VKO "Almaz — Antey". — 2023. — No. 1. — P. 82-89. — DOI: 10.38013/2542.0542.2023.1.82.
9. Kostrov B.V. Ortogonal encoding of binary images / B.V. Kostrov, N.N. Grinchenko, S.N. Baranova et al. // Bulletin of the YAHMS AAW. — 2023. — No. 2. — P. 82-87.
10. Gonzalez R. Digital Image Processing / R. Gonzalez, R. Woods. — Edition 4th. — Moscow: Technosphere. — 2018. — 1184 p.
11. Kostrov B.V. Protocol of image transmission over communication channel / B.V. Kostrov, A.S. Bastrychkin, B.A. Kostrov et al. // Intelligent and Information Systems: Materials of All-Russian Scientific and Technical Conference. — Tula State University. — Tula. — 2016. — p. 220-225.
12. Amos Gilat MATLAB: Theory and Practice. 5-th edition / Translated from English Smolentsev N.K. — M.: DMK Press. — 2016. — 416 p.
13. Sajjadi M.S.M. Enhancenet: Single image super-resolution through automated texture synthesis / M.S.M. Sajjadi, B. Scholkopf, M. Hirsch. // The IEEE International Conference on Computer Vision (ICCV). — 2017.
14. Lempitsky V. "Deep Image Prior," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition / V. Lempitsky, A. Vedaldi and D. Ulyanov. — Salt Lake City, UT, USA. — 2018. — p. 9446-9454. — DOI: 10.1109/CVPR.2018.00984.
15. Bansal Y. PixelNN: Example-based Image Synthesis / Y. Bansal, D. Sheikh, I. Ramanan. — 2017. — abs/1708.05349. — CorpusID: 6990380.
16. W. Shi J. Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network / W. Shi J., F. Caballero, J. Huszar. // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — 2016. — p. 1874-1883. — CorpusID: 7037846.