



UDC 528.2, 004:528

## APPLICATION OF THE THEORY OF WAVELETS FOR COMPRESSION AND FILTRATION OF GEOINFORMATION

Aleksandr S. YARMOLENKO<sup>1</sup>, Olga V. SKOBENKO<sup>2</sup>

<sup>1</sup> Yaroslav-the-Wise Novgorod State University, Veliky Novgorod, Russia

<sup>2</sup> Belorussian State Academy of Agriculture, Gorki, Mogulev region, Republic of Belarus

The purpose of the article is to develop a detailed and accessible technology for the application of wavelets in the processing of geo-information, the subject of research is wavelet-based filtering and compression of geo-information. The research methodology is based on the modern theory of wavelets in the light of linear algebra. Research methods involve study and generalization, abstraction, formalization, mathematical modeling using computer programs compiled by the authors.

After the introduction and formulation of the problem, the basic positions of linear algebra are presented, on which the content of the article is based when constructing orthonormal bases in one- and two-dimensional cases. First, the application of the general theory to the decomposition of the vector of initial data in the Haar and Shannon bases is given. Further, on the basis of the Haar basis, orthonormal bases of wavelet transforms and filtering information are constructed. The procedure for creating wavelet filters by a sequence of convolutions, the use of MSA analysis for constructing an orthonormal basis of the wavelet transform is considered. Implemented the practical possibility of wavelet filtering based on specific programs for modeling geo-information data fields and images, data compression and filtering. The result of the work is the methods of constructing orthonormal bases by various methods of wavelet transform, based on which algorithms and corresponding computer programs for geoinformation compression are compiled using the example of terrain and photographic images. The efficiency of geoinformation compression and noise filtering using wavelets was investigated. A method has been developed for determining the value of a filter depending on the accuracy of the initial geo-information, illustrated by the example of calculating the filter value for compressing information about the heights of the terrain. The same technique is recommended for image filtering.

**Key words:** wavelet transform basis; convolution; MSA analysis; filtration; compression; accuracy; program; picture; data field

**How to cite this article:** Yarmolenko A.S., Skobenko O.V. Application of the Theory of Wavelets for Compression and Filtration of Geoinformation. Journal of Mining Institute. 2018. Vol.234. P. 612-623. DOI: 10.31897/PMI.2018.6.612

**Introduction and formulation of the problem.** Compressing and filtering of geo-information are a topical issue of the theory of mathematical processing of geodetic information (measurements) and images. At the same time, it is necessary to maximize the use of the obtained information and to get the results with satisfactory accuracy and minimal storage cost, which is associated with the compression of information. This is important when working with geo-information in geodesy, land management, environmental engineering, land monitoring, and in conducting precise agriculture [20, p.200]. In this case, the algorithms should be simple and accurate to calculate.

At present, Fourier transformations are widely used to simulate objects of compression and filtering of information [6, 9-11, 16-19, 23]. However, even in its fastest version (fast Fourier transformation – FFT), it is associated with many calculations. In turn, it was noted in [11, 18] that, in contrast to Fourier processing, transformations with other bases are possible, restoring discrete and continuous functions, but significantly reducing computations. One of such bases is wavelet. Currently, it is gradually finding practical application [7-15]. It should be noted here that the works [5-7, 21, 22] are of informative purposes only, they do not present the thorough use of wavelets [21], but mostly describe the noise suppression filters that can be used in modeling processes on wavelet basis. In the works [3, 17] the main theoretical principles on the theory of wavelets, based on foreign studies, are presented. To create technologies for processing geo-information based on these papers, additional research is needed. In [1, 11], significant studies were carried out on the description of the gravitational field of the Earth by wavelets. It is shown, without a detailed description of the technology, that the use of information compression is not always effective when working with wavelets, although when using all the information, the field is quickly and completely restored. In [24, 25], the theory of only the first stage of the wavelet



transform is described in detail, and it is not brought to a universal algorithm. In [14, 15] the original idea of combining the Fourier basis and wavelet basis for modeling the gravitational field of the Earth (GFE) is put forward, existing wavelets are presented, only a general scheme of wavelet analysis synthesis is presented. In [4] in the wavelet transformation, foreign packets are used with an unknown algorithm.

The listed works are of practical and theoretical importance in geo-information processing, but the theory of wavelets is not brought to the engineering level of their application. The following tasks are set and solved on the basis of the abovementioned papers in order to develop a detailed and accessible technology for applying wavelets in geoinformation processing: 1) constructing orthonormal wavelet transformation bases and filtering; 2) the creation of wavelet filters; 3) construction of filters by a sequence of convolutions; 4) application of MSA-analysis for constructing an orthonormal wavelet transformation basis; 5) wavelet filtering and data compression based on specific programs for modeling geoinformation data fields.

**Basic provisions.** In accordance with [18], we give the following definitions. The cyclic shift operator of the sequence  $Z$  [18, p. 125] to the  $k$ -positions to the right is the operator  $R_k$ , creating a new sequence  $R_k Z$  according to the formula

$$(R_k Z)(n) = Z(n - k), \quad (1)$$

where  $n$  – element number in the generated sequence.

Let there be a series (vector of values)

$$Z = (1, 0, 0, 1)^T \quad (2)$$

when  $N = 4$ . Then with  $k = 1$   $(R_1 Z)(0) = Z(0 - 1) = Z(-1) = Z(N - 1) = Z(4 - 1) = Z(3) = 1$ ;  $(R_1 Z)(1) = Z(1 - 1) = Z(0) = 1$ ;  $(R_1 Z)(2) = Z(2 - 1) = Z(1) = 0$ ;  $(R_1 Z)(3) = Z(3 - 1) = Z(2) = 0$ .

Thus we get a new vector sequence

$$R_1 Z = (1, 1, 0, 0)^T. \quad (3)$$

With  $k = 4$  we get the original vector again

$$R_1 Z = (1, 1, 0, 0)^T. \quad (4)$$

According to definition 3.7 and theorem 3.8 [18, p.176] for some given vectors  $U$  and  $V$ , belonging to the same space of elements as the vector  $Z$ , for example (2), it is possible to construct an orthonormal basis of the form

$$B = \{R_{2k} V\}_{k=0}^{M-1} \cup \{R_{2k} U\}_{k=0}^{M-1} = \{V, R_2 V, R_4 V, \dots, R_{N-2} V, U, R_2 U, R_4 U, \dots, R_{N-2} U\}, \quad (5)$$

where  $U$ , according to [18], could be called the father wavelet, and  $V$  – mother wavelet; the symbol  $\cup$  – logical union of sets,  $M = N/2$ .

In this article, the components of the basis, constructed by the vector  $U$ , will be called the paternal one and by the vector  $V$  – maternal.

Orthonormality (5) is possible then only [13, p.250-258, Lemma 7.1; 18, Theorem 3.8], when the matrix system

$$A(n) = \frac{1}{\sqrt{2}} \begin{pmatrix} \hat{U}(n) & \hat{V}(n) \\ \hat{U}(n+M) & \hat{V}(n+M) \end{pmatrix} \quad (6)$$

for  $n = 0, 1, 2, \dots, M-1$  is unitary.

By unitary is meant such a matrix [18, p. 100], for which

$$A^{-1} = A^*, \quad (7)$$

where  $A^{-1}$  – is the inverse of  $A$  matrix, and  $A^*$  is the matrix conjugate to  $A$ , obtained by the complex conjugate value from all elements of the  $A^T$ , transposed to  $A$ . The complex conjugate to the number



$z = x + iy$  is the number  $\bar{z} = x - iy$ . For matrices of real numbers,  $A^* = A^T$ ,  $\hat{U}(n)$ ,  $\hat{V}(n)$  are the Fourier transforms of given vectors  $U$  and  $V$  respectively. It is believed that the pair  $U$  and  $V$  gives rise to an orthonormal basis. In [18] the  $N$ -dimensional space of elements of the vector  $Z$  is designated as  $l^2(Z_N)$  space of values, including complex ones, on which the  $l^2$ -norm is defined [18, p.112]:

$$\|Z\| = \left( \sum_{k=0}^{N-1} |Z(k)|^2 \right)^{\frac{1}{2}}.$$

The set of  $Z_N$  sequence indices is denoted as  $Z_N = \{0, 1, \dots, N-1\}$  [18, p.111]. The generalization of the two-dimensional case is  $N_1, N_2$ -dimensional space  $l^2(Z_{N_1} \cdot Z_{N_2})$ . If we assume that  $\{B_0, B_1, \dots, B_{N_1-1}\}$  is the orthonormal basis [18, p.134]  $l^2(Z_{N_1})$  and  $\{C_0, C_1, \dots, C_{N_2-1}\}$  is the orthonormal basis  $l^2(Z_{N_2})$ , then the orthonormal basis of the space  $l^2(Z_{N_1} \cdot Z_{N_2})$  will be  $D_{m_1, m_2}(n_1, n_2) = B_{m_1}(n_1)C_{m_2}(n_2)$ , where  $0 \leq m_1 \leq N_1 - 1$ ;  $0 \leq n_1 \leq N_1 - 1$ ;  $0 \leq m_2 \leq N_2 - 1$ ,  $0 \leq n_2 \leq N_2 - 1$ .

For example, when

$$B_{m_1}(n_1) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}; \quad C_{m_2}(n_2) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

we get

$$D_{m_1, m_2}(n_1, n_2) = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

The scalar product of (complex) vectors  $z$  and  $w$  is an expression [18, p.91]

$$\langle z, w \rangle = z w = \sum_{j=1}^n z_j \bar{w}_j,$$

where  $\bar{w}_j$  is a number complex conjugated to  $w_j$ . The convolution  $z * w$  is a vector with the following components

$$z * w(m) = \sum_{n=0}^{N-1} z(m-n)w(n).$$

To use the convolution in [18, p.172] they introduce the conjugate reflection  $\tilde{\omega}$  of the vector  $\omega$ :

$$\tilde{\omega}(k-n) = \bar{\omega}(n-k),$$

then

$$z * \tilde{\omega}(k) = \sum_{n=0}^{N-1} Z(n) * \tilde{\omega}(k-n) = \sum_{n=0}^{N-1} Z(n) * \bar{\omega}(n-k).$$

**The construction of orthonormal wavelet transformation bases and filtering.** As an example of constructing an orthonormal basis, we take the Haar vectors [13, p.270; 18, p. 190]:

$$U = \left( \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, \dots, 0 \right)^T; \tag{8}$$

$$V = \left( \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, \dots, 0 \right)^T. \tag{9}$$

Since it was established in [18] that the matrices  $A(n)$  (6) for vectors  $U$  and  $V$  are unitary, then using the rule (5), we construct the orthonormal basis for wavelet transformation of the vector  $Z$ , for example (2). Here  $N = 4$  and  $M = N/2 = 2$ .



Then, in accordance with (1), (5), based on (8), (9), it is possible to make an orthonormal wavelet basis:

$$B = (V, R_2V, U, R_2U) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & -1 & 0 & 1 \end{pmatrix}. \quad (10)$$

Then the wavelet transformation at this stage will be

$$\hat{Z} = B^* Z. \quad (11)$$

In the example it is

$$\hat{Z} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix}, \quad (12)$$

and the inverse transformation, following (11), has the form

$$Z = B^{-1} \hat{Z} = B^* \hat{Z} = B^T \hat{Z}. \quad (13)$$

In the example it is

$$Z = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Now let's filter the signal, i.e. its decomposition into paternal and maternal components of the wavelet basis. According to [13, c.26; 18, p.99] in space with a complex scalar product  $\langle \cdot, \cdot \rangle$  and an orthonormal basis  $R = \{u_1, u_2, \dots, u_n\}$  for any  $v$  of this space

$$v = \sum_{j=1}^n \langle v, u_j \rangle u_j. \quad (14)$$

Where  $v$  corresponds to the data vector  $Z$  and basis vectors  $u_j$  for all columns of the matrix (10). After substituting them into (14) we get

$$Z = Q(Z) + P(Z), \quad (15)$$

where

$$Q(Z) = \frac{1}{2} \left[ (1001) \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} + (1001) \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \right] = \frac{1}{2} \left[ 1^* \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} + (-1) \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \right] = \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix};$$

$$P(Z) = \frac{1}{2} \left[ (1001) \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + (1001) \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right] = \frac{1}{2} \left[ 1^* \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + 1^* \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right] = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

It is clear from the example that  $Q(Z)$  and  $P(Z)$  are the components of the high and low frequencies, respectively. Following [18, p.180], the high frequency component of the signal is represented as the formula



$$Q(Z) = \sum_{k=0}^{(N/2)-1} \langle Z, R_{2^k} V \rangle R_{2^k} V = (\langle Z, R_0 V \rangle) R_0 V + (\langle Z, R_2 V \rangle) R_2 V + \dots + \left( \left\langle Z, R_{\frac{N}{2}-1} V \right\rangle \right) R_{\frac{N}{2}-1} V, \quad (16)$$

and the low-frequency component as

$$P(Z) = \sum_{k=0}^{(N/2)-1} \langle Z, R_{2^k} U \rangle R_{2^k} U = (\langle Z, R_0 U \rangle) R_0 U + (\langle Z, R_2 U \rangle) R_2 U + \dots + \left( \left\langle Z, R_{\frac{N}{2}-1} U \right\rangle \right) R_{\frac{N}{2}-1} U. \quad (17)$$

**Examples of filters based on Shannon bases with real and complex numbers.** If we use the components of the vectors  $\hat{U}$  and  $\hat{V}$  (Fourier transforms) in [18, p.179, 181], then for  $N=4$  using the inverse Fourier transform, by analogy with (14)-(17), we can determine the high and low frequency components for these examples and in these bases (Table 1).

Table 1

**Decomposition of the Z vector in paternal and maternal wavelets in various bases**

Original vector Z	Bases		
	Haar	Shannon 1	Shannon 2
$P(Z)$			
1	$\frac{1}{2}$	$\frac{3}{4} - \frac{1}{4}i$	$\frac{1}{2}$
0	$\frac{1}{2}$	$\frac{1}{4} - \frac{1}{4}i$	0
0	$\frac{1}{2}$	$\frac{1}{4} + \frac{1}{4}i$	$\frac{1}{2}$
1	$\frac{1}{2}$	$\frac{3}{4} + \frac{1}{4}i$	1
$Q(Z)$			
1	$\frac{1}{2}$	$\frac{1}{4} + \frac{1}{4}i$	$\frac{1}{2}$
0	$-\frac{1}{2}$	$-\frac{1}{4} + \frac{1}{4}i$	0
0	$-\frac{1}{2}$	$-\frac{1}{4} - \frac{1}{4}i$	$-\frac{1}{2}$
1	$\frac{1}{2}$	$\frac{1}{4} - \frac{1}{4}i$	0

From the comparison of the bases given in Table 1, as well as the Meyer, Bartlett-Lemarié, Daubechies bases [13, p.207], the Haar basis must be given priority for the following reasons:

- Haar basis is simple to calculate;
- the signal filtering is clear on its basis; so, the low-frequency part at the first stage is equal to the zero Fourier transform coefficient, the high-frequency part corresponds to the deviations of the signal from its middle;
- in bases other than Haar, additional requirements are imposed on the number  $N$ . For example, in the Shannon bases it must be a multiple of 4, and in the Haar basis of multiplicity of  $2^p$  with  $p = \log_2 N$ ; the number  $N$  determines the Daubechies basis as well [18].



**Constructing filters by a sequence of convolutions.** In the wavelet expansions (16), (17), it is noted that the coefficients of  $R_{2k}V$ ,  $R_{2k}U$  are corresponding scalar products, which in the form of convolutions can be written as [13, 18, 19]:

$$\langle Z, R_{2k}V \rangle = Z^* \tilde{V}(2k), \quad \langle Z, R_{2k}U \rangle = Z^* \tilde{U}(2k).$$

These expressions are the theoretical basis for the fast wavelet transform. Then a filter based on such a convolution is constructed as follows [13, p.279; 18, p. 187]: 1) forms the vectors  $\tilde{U}$ ,  $\tilde{V}$  (described at the end of the basic provisions); 2) constructs convolutions  $Z^* \tilde{U}$ ,  $Z^* \tilde{V}$ ; 3) introduces the decimation operator for removal of components with odd numbers  $D(Z\tilde{U})$ ,  $D(Z\tilde{V})$ ; 4) implements the sparse sample operator for doubling the size of the vector by inserting a zero between two adjacent values  $U(D(Z^* \tilde{U}))$ ,  $U(D(Z^* \tilde{V}))$ ; 5) performs filtering – the vector of the low-frequency component is constructed in the form of a convolution  $P(Z) = U^*U(D(Z^* \tilde{U}))$  and high-frequency  $Q(Z) = V^*U(D(Z^* \tilde{V}))$ ; 6) restores the signal  $Z = P(Z) + Q(Z)$ .

This ends the first stage of the decomposition of the signal into high- and low-frequency components. The number of all stages is determined by the formula  $p = \log_2 N$ . Each subsequent stage consists of an analysis phase and a synthesis phase. In the analysis phase, at stage  $n$  the following is performed: 1) input of the vector  $Z_{n-1}^T * \tilde{U}_{n-1}$  and its decimation  $Z_n = D(Z_{n-1}^T * \tilde{U}_{n-1})$ ; 2) input of vectors  $U_n$ ,  $V_n$  of the normalized basis of dimension  $N_1 = N/2^{n-1}$ ; 3) decimation and thinning of convolutions  $Z_n^T * \tilde{U}_n$ ,  $Z_n^T * \tilde{V}_n$ ;  $UDU_n = UD(Z_n^T \tilde{U}_n)$ ,  $UDV_n = UD(Z_n^T \tilde{V}_n)$ ; 4) the high-frequency and low-frequency components of the  $Z_n$  vector are calculated:  $Q(Z_n) = V_n^*U(D(Z_n^* \tilde{V}_n))$ ;  $P(Z_n) = U^*U(D(Z_n^* \tilde{U}_n))$ .

In the synthesis phase, the following is performed: 1) thinning the vectors  $Q(Z_n)$ ,  $P(Z_n)$  (the thinning is performed  $n - 1$  times until the original signal is reached):  $U(Q(Z_n))$ ,  $U(P(Z_n))$ ; 2) the convolution operations receive the high and low frequency components of the signal at stage  $n$ :  $Q_n(Z) = U(Q(Z_n)) * U_1$ ;  $P_n(Z) = U(P(Z_n)) * U_1$ .

The end result of the synthesis analysis is:

$$Z = P_1(Z) + \sum_{i=0}^{p-1} Q_{p-1}(Z).$$

This algorithm is implemented in a specially designed software Sub Wavelet Analysis Synthesis () created by the authors.

$Z$  signal can be represented as decomposition

$$Z = c_0 b_0 + c_1 b_1 + \dots + c_{N-1} b_{N-1}, \quad (18)$$

basis  $B$ , which is a set of orthonormal bases [18, p.176, p.185-186]  $B = \{b_i\}_{i=0}^{N-1}$ . It is obvious that in this case  $c_i = (Zb_i^T)$ . The task is to determine the orthonormal vectors in decomposition (18). In the theory of wavelets [18, p.209-225, definition 3.28] instead of (18), the following representation is accepted:

$$Z = \sum_{k=0}^{(N/2^{h-1})} c_{1,k} \Phi_{-p,k} + \sum_{k=0}^{(N/2^{h-1})} c_{2,k} \Psi_{-p,k} + \sum_{k=0}^{(N/2^{p-1}-1)} c_{3,k} \Psi_{-(p-1),k} + \dots + \sum_{k=0}^{(N/2^{p-(p-1)-1})} c_{N/2,k} \Psi_{-(p-(p-1),k)}. \quad (19)$$

If we take  $p = \log_2 N$ , then (19) will be

$$Z = c_{1,0} \Phi_{-p,0} + c_{2,0} \Psi_{-p,0} + c_{2,0} \Phi_{-p,1} + \dots + c_{N/2,0} \Psi_{-1,0} + c_{N/2,1} \Psi_{-1,1} + \dots + c_{N/2,N/2-1} \Psi_{-1,N/2-1}. \quad (20)$$

Here, all the vectors of the basis  $B \psi_{-j,k}$  are written from left to right according to the degree of detail of the vector  $Z$ . In [18], they are written from right to left. The construction of basis vectors  $\psi_{-j,k}$  is performed in the following order [18, p.225]: 1) the sequence of wavelet filters  $U_1, V_1; U_2, V_2; \dots; U_p, V_p$  is used; here  $U_l, V_l \in l_2(Z_{N/2^{l-1}})$ , for example, on the basis of (8), (9) with  $l=2$  we have

$$\begin{pmatrix} U_2 \\ V_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \end{pmatrix}; 2) \text{ each of the bases } \psi_{-j,k} \text{ is constructed by formula [18, p.225]}$$



$$\Psi_{-j,k} = R_{2^j k} f_j,$$

where  $f_l = g_{l-1} U^{l-1}(V_l)$ , and  $g_l = g_{l-1} U^{l-1}(U_l)$  with initial values equal to  $f_1 = V_1$ ,  $g_1 = U_1$ . In the above expressions, the operator  $U^{l-1}(V_l)$  means  $l-1$ -times thinning of vector  $V_l$  or  $U_l$ . The convolution elements  $f_l(n)$ ,  $g_l(n)$  для  $n = 0, 1, \dots, N-1$  are calculated by the formulas

$$f_l(n) = \sum_{m=0}^{N-1} (g_{l-1}(m) U^{l-1}(V_l)(n-m)); \quad g_l(n) = \sum_{m=0}^{N-1} (g_{l-1}(m) U^{l-1}(U_l)(n-m)).$$

**MSA-analysis for constructing an orthonormal basis of the wavelet transform.** In fundamental works [13, 18, 19], the wavelet decomposition assumes the presence of scaling and shifting functions. In [18, definition 5.30], the scaling function is also called the paternal wavelet. In [13, p.103; 19, p.44] the term «paternal wavelet» in the definition of the scaling function does not apply. In the same works [18, p.198; 19, p.39-42, ch.5] the shifting function (mother wavelet) is simply called wavelet.

In the Haar system, in accordance with [18, (5.49)], the father wavelet is written as

$$\varphi(x) = \begin{cases} 1, & 0 \leq x < 1; \\ 0, & \text{otherwise.} \end{cases} \quad (21)$$

Maternal wavelet according to the formula [18, p.380]

$$\psi(x) = \varphi(2x-1) - \varphi(2x), \quad (22)$$

and that is easy to show has the form

$$\psi(x) = \begin{cases} -1, & 0 \leq x < \frac{1}{2}; \\ 1, & \frac{1}{2} \leq x < 1; \\ 0, & \text{otherwise.} \end{cases} \quad (23)$$

Paternal and maternal wavelets allow us to construct an orthonormal basis of the wavelet transform to represent discrete geoinformation for processing. For this, the so-called multiple-scale analysis (MSA) has been developed in wavelet theory. On its basis, an easy-to-use algorithm for constructing an orthonormal wavelet transform basis is created. MSA analysis based on functions

$$\varphi_{j,k} = 2^{-\frac{j}{2}} \varphi(2^{-j} x - k); \quad (24)$$

$$\psi_{j,k} = 2^{-\frac{j}{2}} \psi(2^{-j} x - k). \quad (25)$$

These formulas are given in [5, p.113, p.128; 13, p. 241; 18, p. 19, p. 193]. Moreover, in the works cited, the exponent in (24), (25) is attributed to both positive and negative signs, as in our case. In the case of a negative degree, the graph of the function is stretched along the x axis, and if it is positive, it is compressed. In our work, we are interested in stretching along the x axis, therefore, we have taken the notation of the degree with a negative sign. With such a record, the values of the functions are refined depending on the number of orthonormal vectors of the basis of the wavelet decomposition, i.e. there is an increase in the details of the analyzed information or an increase in the resolution. Therefore, in foreign literature [5, p.113], the MSA is rightly called multiresolution analysis. The construction of wavelet bases in the Haar system will be based on (21) – (25) in the following order:

- 1) construction of the basis vector of the zero approximation  $\varphi$ ;
- 2) construction of the subsequent specifying basic vectors.

1. *Construction of the basis vector of the zero approximation.* The basis for the construction of all basis vectors is the wavelet basis of the form [18, (3.67), (3.84)], which we use in the expansion (20). Although in the theory of wavelets [3, 5, 13, 17-19], the possibility of constructing several basis vectors of the zeroth approximation is allowed, in this paper we confine ourselves to only one



in accordance with this decomposition. In this case, such a decomposition is most often applicable in various rows, including the decomposition in a Fourier series. In accordance with decomposition (20), adopted here as the basis, the basis vector  $\varphi_{-p,k}$  is unique, since  $k$  «moves» through values from 0 to  $((N/2^p) - 1)$  [18, (3.67)]. (In MSA the notations  $\varphi_{-p,k}$ ,  $\psi_{-p,k}$  of decomposition (20) correspond to  $\varphi_{-p,k}$ ,  $\psi_{-p,k}$ ). Since  $N = 2^p$ , then  $k = 0$ . Hence, in (24)  $j = p$ ,  $k = 0$ . According to (21)  $0 \leq 2^{-p}x - k < 1$ , or  $k \leq x < (1 + k)/2^p$ . With  $k = 0$  there will be  $0 \leq x < 1/2^p$  and the components of this vector will be determined by the formula

$$\varphi(x) = 2^{-\frac{p}{2}} \begin{cases} 1, & 0 \leq x < 2^{-p}; \\ 0, & \text{otherwise.} \end{cases} \quad (26)$$

Thus, based on (26) we obtain the vector

$$\varphi_{p,0} = 2^{-\frac{p}{2}} (1, 1, \dots, 1) \quad (27)$$

with the number of identical members  $2^p$ .

2. *Construction of the subsequent specifying basic vectors.* Subsequent specifying basis vectors are calculated by (25) with (23) taken into account. The order of the set of non-intersecting basis vectors [18, (3.67)]  $U\{\psi_{-j,k}\}_{k=0}^{(N/2^j)-1}$  is determined by the value of  $j$ . His value changes from  $p - 1$  to 0 with a step of  $-1$ . Then the number of basis refinement vectors of order  $j$  will be a value of  $N/2^j$ . The value of  $k$  will change from 0 to  $N/2^j - 1$ . For example, for  $j = p$  we have the following specifying vector (unique):  $\psi_{-p,0}$ . For  $j = l$  the specifying vectors of order  $l$  will be:  $\psi_{-l,0}$ ;  $\psi_{-l,1}$ ;  $\psi_{-l,(N/2^l)-1}$ . Thus for  $N = 8$   $p = 3$  the specifying vectors of order  $l = 1$  with the upper limit  $k$ , equal to  $N/2^l - 1 = 2^p/2 - 1 = 3$ , will be:  $\psi_{-1,0}$ ;  $\psi_{-1,1}$ ;  $\psi_{-1,2}$ ;  $\psi_{-1,3}$ . Each of these vectors with a certain  $k$  is also determined by (25) with regard to the maternal wavelet (23). Then you can write

$$\psi_{l,k} = 2^{-\frac{l}{2}} \begin{cases} -1, & 0 \leq 2^{-l}x - k < \frac{1}{2}; \\ 1, & \frac{1}{2} \leq 2^{-l}x - k < 1; \\ 0, & \text{otherwise} \end{cases} \quad \text{or} \quad \psi_{l,k} = 2^{-\frac{l}{2}} \begin{cases} -1, & k2^l \leq x < \left(\frac{1}{2} + k\right)2^l; \\ 1, & \left(\frac{1}{2} + k\right)2^l \leq x < (1 + k)2^l; \\ 0, & \text{otherwise.} \end{cases} \quad (28)$$

As an example, take  $p = 3$ ,  $l = 1$ ,  $k = 2$ .

Then

$$\psi_{1,2} = \frac{1}{2} \begin{cases} -1, & 4 \leq x < 5; \\ 1, & 5 \leq x < 6; \\ 0, & \text{otherwise.} \end{cases}$$

Table 2 gives an example of wavelet decomposition using both the convolution method and the MSA method.

Table 2

Decomposition of vector  $Z$  by components  $\varphi$  and  $\psi$

Original vector	Components							
	$c_{1,0} \cdot \varphi_{-3,0}$	$c_{1,0} \cdot \psi_{-3,0}$	$c_{1,0} \cdot \psi_{-3,0}$	$c_{3,1} \cdot \psi_{-2,1}$	$c_{4,0} \cdot \psi_{-1,0}$	$c_{4,1} \cdot \psi_{-1,1}$	$c_{4,2} \cdot \psi_{-1,2}$	$c_{4,3} \cdot \psi_{-1,3}$
4	7.25	-3.25	-1	0	1	0	0	0
2	7.25	-3.25	-1	0	-1	0	0	0
3	7.25	-3.25	1	0	0	-2	0	0
7	7.25	-3.25	1	0	0	2	0	0
10	7.25	3.25	0	-1.5	0	0	1	0
8	7.25	3.25	0	-1.5	0	0	-1	0
10	7.25	3.25	0	1.5	0	0	0	-2
14	7.25	3.25	0	1.5	0	0	0	2





Each component was calculated by the formula:  
for the low frequency

$$P_{i,j} = 1/2^{p-j} \sum_{i=1}^{N/2^{p-j}} P_{i,j-1}$$

and high frequency

$$Q_{i,j} = P_{i,j-1} - P_{i,j}$$

Here  $j$  – number of the stage;  $i$  – number of the signal component in the group (the group consists of 2, 4, ...  $2p$  elements, depending on the sequence number of the stage, respectively):  $p$  is the number of stages. The end result:  $Z = P_1 + \sum_{i=0}^{p-1} Q_{p-i}$ , where  $P_1$  is zero frequency component (vector  $\psi_{-3,0}$  in Table 1);

$Q_{p-i}$  are specifying components. In the above transformations at each stage with non-integer  $p$  the remainder of elements is possible, the number of which is less than the number  $2^{ip}$  with  $ip = 1, 2, \dots, 2p$ . This balance also contains the average  $TZ$  value, which is written in the highest  $p_{ip}$  line in place of this remainder elements. The last line immediately following the line with the number of the integer part of  $p$ , denoted by  $pf$ , is the average of all elements of the previous line  $P_{pf}$ . This average is the same for the  $P_{pf} + 1$ . The subline  $q_{pf} + 1$  is calculated in the general order:

$$q_{pf+1} = P_{pf} - P_{pf} + 1.$$

**Investigation of the efficiency of geoinformation compression and noise filtering using wavelets.** The research is based on specially compiled by the authors of the program in the languages VISUAL BASIC Excel (VBE) and IDL systems ENVI for compression and filtering geo-information. As the first object of research, the model of relief, given in [23], was adopted. In VBE-program Sub MacrosWAVELsYstOsh() the true heights of the points are represented by the  $cc1()$  array, and the heights weighed down by random errors are represented by  $cm()$  array. As in [23], the connection of real arrays is defined by the formula  $cm(i) = cc1(i) + \text{delta}$ , where  $i$  changes from 0 to  $N - 1$ , and  $\text{delta} = \text{Randbetwen}(-t, +t)\text{Std}$  is the function of VBE for generating a random number in the range of quantile values from  $-t$  to  $+t$ , Std is a standard for random error (noise). The heights of the  $cm()$  array are subject to wavelet decomposition by frequencies. With wavelet compression and, accordingly filtering, the expansion terms remained the largest in amplitude at all frequencies. The filtration vector is represented by the array  $\text{Filt} = \text{Array}(1; 0.8; 0.6; 0.5; 0.25; 0.15; 0.1; 0)$ . For example, if the  $\text{Filter} = \text{Filt}(1) = 0.8$  value at all decomposition frequencies, the values of more than 0.8 remained, the rest values were reset. At the output a filtered array of heights  $Tw()$  was formed. The standard deviation (SD) of the filtered heights was determined by the formula

$$CKO = \sqrt{\frac{\sum_{i=1}^N (Tw(i) - cc(i))^2}{N}}$$

Table 3

Standard deviations (SD) for each filter (in meters) depending on the standards (Std) of the distribution of random height errors

Filter	Residual information, %	Std values				
		0.05	0.1	0.3	0.4	0.5
1	25	0.66	0.88	0.71	0.93	0.94
0.8	33	0.55	<b>0.35</b>	<b>0.63</b>	<b>0.86</b>	<b>0.88</b>
0.6	40	0.23	0.26	0.63	0.86	0.94
0.5	48	0.23	0.26	0.64	0.86	0.94
0.25	51	0.19	0.24	0.63	0.86	0.99
0.15	66	0.15	0.20	0.64	0.86	0.99
0.1	70	0.11	0.17	0.63	0.86	0.99
0	<b>100</b>	<b>0.09</b>	<b>0.18</b>	0.63	0.86	0.99

Table 3 shows the SD values depending on the filter value – Filter and the accuracy of the modeled heights – Std values. The remaining information (in percent) is calculated as the percentage ratio of the number of the largest remaining coefficients to their total number before compression, equal to 15.

The effect of the systematic part of the wavelet compression on the accuracy of the final result after zeroing the details is given below.

SD for each filter at heights not burdened with random errors:

Filter	0.55	0.34	0.25	0.23	0.15	0.10	0.06	0
SD	1	0.8	0.6	0.5	0.25	0.15	0.1	0

From the experimental studies it follows that:

- in order to maintain the highest possible accuracy of the initial heights when wavelet decomposition is performed, data compression is unacceptable; even in the absence of compression (Filter = 0) and the presence of random errors, the standard deviation of the resulting heights is greater than their standard deviation at the input;
- data compression can be allowed, but the threshold for reducing the accuracy of the relief image should be taken into account; for example, with a threefold compression (Filter = 0.8) and Std = 0.1, the standard deviation of the heights at the output of the standard deviation = 0.35 is greater than Std by 3.5 times, in other cases it is 2 times greater. Figure 1 shows a relief built from true heights, and Figure 2 shows a relief obtained with triple wavelet compression (Filter = 0.8) and an error height standard of 0.3 m (Std = 0.3 m);
- regularities of the effect of wavelet compression on the accuracy of the information obtained at the output are the same as when filtering information in Fourier series [23], but here the amount of calculations is negligible compared to Fourier transforms.

The method for calculating the filter value for compressing elevation information can be adopted as follows:

1. According to the known method, the evaluation of the standard of the elevation of the relief (Std) is determined in the form of the standard error of the relief survey.

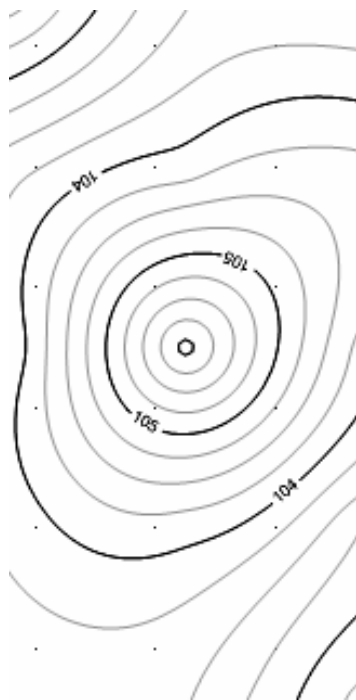


Fig. 1. Relief in horizontals, built on the heights adopted as initial [23] (cross-section of the relief 0.25 m)

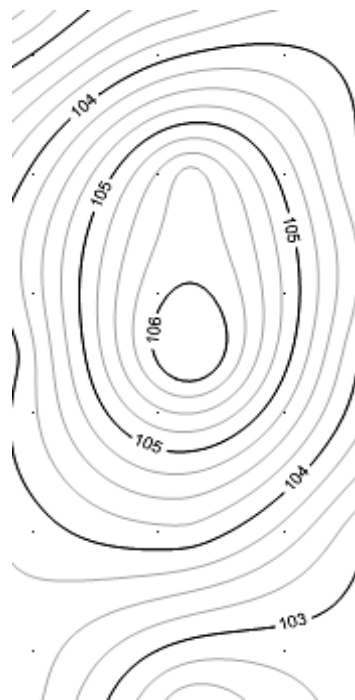


Fig. 2. Filtered terrain in case of triple wavelet information compression (Filter = 0.8, Std = 0.3 m)

2. For this object, the standard deviations for each filter are determined.

3. If the value of the average error of the relief, obtained as  $0.8Std$  [2], is less than one-third of the height of the cross-section of the relief, then the threshold obtained (experimentally) for reducing the accuracy of compressed information about the relief for a certain filter value (Filter). If the average error of the relief is more than a third of the height of the cross section of the relief and approaches the magnitude of the standard deviation, obtained at 25-30 % compression of information (i.e. three- and four-times compression), then the threshold value (standard deviation) reducing the accuracy of the compressed information about the relief is not important.

At the same time, the authors compiled a program *pro\_oi\_data\_corr24bitWAVE* in the algorithmic

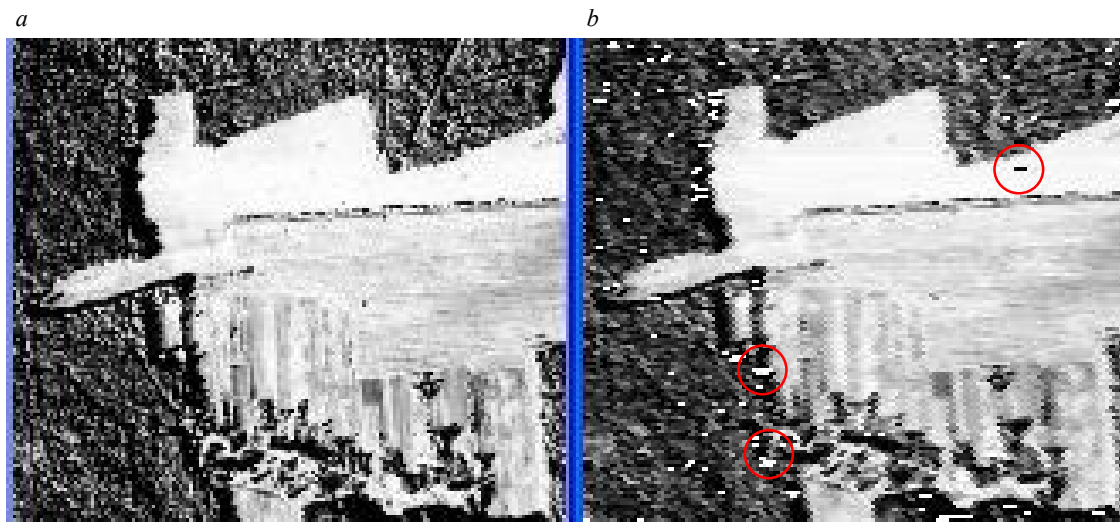


Fig.3. The images are original (reconstructed over all wavelet elements) (a) and compressed 2.5 times (b) with filtered values (less than 4 units) constituting the wavelet decomposition modulo

language IDL of the ENVI system of wavelet filtering (compression) of images. Figure 3 shows the images — the original (and reconstructed over all wavelet elements) and compressed 2.5 times with filtered values (less than 4 units) constituting the wavelet expansion modulo.

Circles mark discrepancies with the original image. With a greater compression of various mismatches, the circles are larger and are already noises at which the image is unsuitable for use. Thus, when compressing an image, it is also recommended to first select a filter in accordance with steps 1-3 in the case of a relief, or in which there are no visible discrepancies and implement wavelet compression.

## Conclusions

1. Wavelet filtering and compression lead to the loss of some information. However, such filtering and compression are very effective at a known acceptable threshold (the value of the standard deviation) of reducing the accuracy of the compressed information. At the same time, the amount of computation is negligible compared to the filtering and compression in the Fourier series and the cosine-sine transforms of the JPEG compression.

2. When filtering and compressing images, it is recommended first to select a filter in accordance with steps 1-3 in the case of relief. It is necessary to set the SD of the pixel, at which compression is acceptable and there are no visible discrepancies.

## REFERENCES

1. Bagrov A.A., Bagrova A.S. Decomposition of spherical functions on Haar wavelets. *Izvestiya vuzov. Geodeziya i aerofotos'emka*. 2008. N 4, p. 6-8 (in Russian).
2. Bol'shakov V.D., Gaidaeв P.A. The theory of mathematical processing of geodetic measurements. Moscow: Nedra, 1977, p. 367 (in Russian).
3. Vorob'ev V.I., Gribunin V.G. Theory and practice of wavelet transform. St. Petersburg: VUS. 1999, p. 204 (in Russian).
4. Gonzha E.A. About wavelet filtering of a digital image of the earth's surface. *Izvestiya vuzov. Geodeziya i aerofotos'emka*. 2017. N 3, p. 105-110 (in Russian).
5. D'yakonov V.P. Wavelets. From theory to practice. Moscow: Solon-R, 2002, p.448 (in Russian).
6. Zhurkin I.G., Shaven'ko N.K. Automated remote sensing data processing. Moscow: Diona, 2013, p. 456 (in Russian).
7. Krasil'nikov N.P. Digital processing of 2D and 3D images. St. Petersburg: BKhV-Peterburg, 2011, p. 608 (in Russian).
8. Lapshin A.Yu. Development and research of methods for calculating the gravimetric height of a quasigeoid and components of the plumb evasion based on the wavelet transform: Avtoref. dis. ... kand. tekhn. nauk. Moskovskii universitet geodezii i kartografii. Moscow, 2011, p. 207 (in Russian).



9. Mazurova E.M. Fast Fourier Transform Algorithms. *Izvestiya vuzov. Geodeziya i aerofotos'emka*. 2004. N 3. p. 18-35 (in Russian).
10. Mazurova E.M. Two-dimensional and matrix representation of the fast Fourier transform. *Izvestiya vuzov. Geodeziya i aerofotos'emka*. 2004. N 4, p. 3-12 (in Russian).
11. Mazurova E.M., Bagrova A.S. To the question of calculating the height anomaly on the basis of the wavelet transform and the fast Fourier transform in a flat approximation. *Izvestiya vuzov. Geodeziya i aerofotos'emka*. 2008. N 4, p. 6-8 (in Russian).
12. Malinnikov V.A., Uchaev D.V. Analysis of methods for the formation of a multifractal measure based on the wavelet processing of experimental data. *Izvestiya vuzov. Geodeziya i aerofotos'emka*. 2007. N 6, p. 57-61 (in Russian).
13. Mala S. Wavelets in signal processing: Per s angl. Moscow: Mir, 2005, Moscow, p. 671 (in Russian).
14. Neiman Yu.M., Sugaipova L.S. Adaptation of the global geopotential model to regional features. Part.1. *Izvestiya vuzov. Geodeziya i aerofotos'emka*. 2014. N 3, p. 3-12 (in Russian).
15. Neiman Yu.M., Sugaipova L.S. Basics of multi-scale geopotential approximation. Moscow: Izd-vo MIIGAiK, 2016, p. 218 (in Russian).
16. Piskunov N.S. Differential and integral calculus: In 2 volumes. Moscow: Nauka, 1978. Vol.2. p. 575 (in Russian).
17. Uelstid S. Fractals and wavelets for compressing images in action. Moscow: Izd-vo Triumf, 2003, p. 320 (in Russian).
18. Freizer M. Introduction to wavelets in the light of linear algebra: Per. s angl. Moscow: Binom. Laboratoriya znanii. 2008, p. 487 (in Russian).
19. Chui K. Introduction to wavelets: Per. s angl. Moscow: Mir, 2001, p. 412 (in Russian).
20. Shparr D., Zakharenko A.V., Yakushev V.P. Precision agriculture. Federal'noe ministerstvo prodovol'stviya, sel'skogo khozyaistva i zashchity prav potrebitelya Germanii. St. Petersburg – Pushkin. 2009, p. 398 (in Russian).
21. Shovengerdt R.A. Remote sensing, models and image processing methods. Moscow: Tekhnosfera, 2010, p. 560 (in Russian).
22. Yakovlev A.N. Introduction to wavelet transform. Novosibirsk: Izd-vo NGTU, 2003, p. 104 (in Russian).
23. Yarmolenko A.S, Skobenko O.V. Filtering geo-information in Fourier series. *Izvestiya vuzov. Geodeziya i aerofotos'emka*. 2016. N 1, p. 107-113 (in Russian).
24. Yarmolenko A.S. The use of wavelets in the analytical representation of discrete functions of graphic information. *Izvestiya vuzov. Geodeziya i aerofotos'emka*. 2008. N 3, p. 20-30 (in Russian).
25. Yarmolenko A.S. Wavelet transform in encoding graphic information. *Izvestiya vuzov. Geodeziya i aerofotos'emka*. 2010. N 4, p. 18-25 (in Russian).

**Authors: Aleksandr S. Yarmolenko, Doctor of Engineering Sciences, Professor, yarmolenko\_alex@mail.ru (Yaroslav-the-Wise Novgorod State University, Veliky Novgorod, Russia), Olga V. Skobenko, Assistant Lecturer, olga-skobenko@mail.ru (Belorussian State Academy of Agriculture, Gorki, Mogulev region, Republic of Belarus).**

*The paper was received on 19 April, 2018.*

*The paper was accepted for publication on 21 September, 2018.*