Technique Medical Image Compression by Linear Algebra Methods

Dmytro Kucherov[1][0000-0002-4334-4175], Galina Rosinska[1][0000-0002-6348-9419], Natalia Khalimon [1][0000-0002-7159-6740], and Ludmila Onikienko[2][0000-0002-2312-583X]

[1] National Aviation University, Kyiv 03058, Ukraine
[2] Central Research Institute of Weapons and Military Equipment of Ukraine's Armed Forces, Kyiv 03049, Ukraine
d_kucherov@ukr.net, rosingala@ukr.net, natalyh1208@ukr.net, 9225foxtrot@gmail.com

Abstract. We consider an important technology of linear algebra, aimed at reducing the size of files intended for storing images and transmitting them over the network. The technology involves the use of elements of tensor analysis based on singular decomposition. A feature of the technology used is the representation of the image by the matrix triad, which includes the tensor core and a pair of unitary matrices containing right and left singular vectors, respectively. Compression is achieved by one recurrent procedure, which involves lowering the rank of the triad to the level of allowable errors while maintaining the original image size. The compression algorithm is an iterative procedure with control of the Frobenius norm by the error matrix of the deviation of the original and the approximation matrix, which ensures the permissible value specified by the designer. The article establishes the basic features of the algorithm, in particular, convergence in a finite number of steps. This technology may be part of some hybrid technology based on different compression methods. The simulation was carried out on typical digital images, confirms the productivity of the procedure, the results obtained are compared with other types of decompositions according to the criteria of the signal-to-noise ratio, the standard error, and visual assessment of perception. Discusses the use of developed technology to store information belonging to big data.

Keywords: Image Compression, Singular Decomposition, Matrix, Algorithm, Codebook.

1 Introduction

In medical practice, non-invasive research methods are widely used, the main advantages of which are the visibility and accuracy of the measurements. Such methods include ultrasound diagnostics, computed and magnetic resonance imaging, radiography and others. The result of the study is shown in the images, the increase of which inevitably causes the problem of storage and transmission of measurement data to the end-user. A natural approach to solving this problem is image compression.

Compression is a popular technique for converting images to a form that is convenient for transmission over wired and wireless communication channels and also for their storage [1]. To date, several effective techniques have been developed based on different approaches. The most commonly used are considered hybrid techniques, including several compression methods, such as JPEG, JPEG2000. The main difference between the two is the replacement of low-performance transforms, such as Huffman and discrete-cosine, with more efficient arithmetic coding and wavelet transform, respectively. Analysis of the development of these methods shows the possibility of replacing individual stages of the method with more effective ones. The main goal of such methods is to improve the compression and recovery performance of images and minimize distortion. The best compression is achieved in the algorithms that involve losses, where frequency separation of the image components is assumed.

One of the effective methods that appeared at the end of the last century is vector quantization. The method assumes a high correlation between the individual elements of the image. However, there are some computational difficulties associated with the definition of a code word. The main output, as in the previously developed methods, is the discarding of several codewords, provided that it does not give significant distortion to the final product. A possible drop option can be constructed using the least-squares method.

The study aim is to study the possibility of linear algebra methods for compressing medical images.

2 Papers Review

There are many works in which the well-known method of linear algebra is applied, which assumes decomposing in singular values (svd), in which the image is decomposed into a matrix triad. By discarding small singular values, the computational cost of presenting the result can be reduced. Currently, there is an increase in work related to the use of singular decomposition in the interests of medical research. The state of the issue of the use of singular decomposition in science and technology is presented in [2]. So, in the work of [3], a singular decomposition is proposed for the detection of brain tumors. The method uses a comparison of the measured parameter with its threshold value. However, there are some problems associated with the application of svd.

In [4], the use of singular decomposition together with the vector quantization algorithm for image compression is proposed, which, in the authors' opinion, seems to be the best technique compared to the discrete cosine transform, which also uses the application of the transform to individual image blocks. The number of vectors involved in quantization is determined by the rank of the matrix. The method involves the elimination of unnecessary (exceeding the rank of the image matrix) computational operations with small singular numbers when the image is restored. Block quantization helps to reduce computational costs when applying singular decomposition to the entire image.

The authors of [5] propose a hybrid compression method based on the sharing of singular decomposition and the Karhunen-Loeve transform. Since the Karhunen-Loeve transform is applied to the entire image, the quality of the reconstructed image when packaged by the singular decomposition of small blocks is higher. The size of large blocks is determined by the bit distribution and is constant for the selected image. Switching between compression methods is proposed to implement the largest functional, determined by the transmission rate and distortion at this speed.

Three schemes for the use of singular decomposition were studied in terms of the quality of the compression and the computational complexity of these processes are presented in [6]. It is noted here that those algorithms that allow you to adaptively choose the size of the block to which the singular value decomposition is applied have greater efficiency.

One of the modern trends in the field of processing large data, which should include the image, is the use of tensor networks. The use of tensor networks is used to construct an image compression algorithm proposed in [7]. The image is transformed into the n-dimensional tensor of the real keta coefficients, which determines the quantum state represented by the matrix products. Visual information is represented by the state of a truncated matrix of multiplications. The method is rather complicated due to the uncertainty of finding important values in the matrices of multiplications. To overcome the resulting uncertainty, the authors of [7] propose the construction of a set of tables, the closest table is found by a genetic algorithm.

The method of representation of images based on the singular decomposition of a high order is proposed in [10]. Unlike the usual representation of a tensor by a set of singular values and right and left unitary matrices, orthogonal functions of one variable are used. When scaling images, the main tensor core retains its shape, while the function matrices expand by a given scale factor, the missing elements of the orthogonal function are complemented by the cubic interpolation method. To compress the method must be supplemented by some means.

Compression of hyperspectral images based on the Tucker decomposition is presented in [11]. Following this method, the image is decomposed into three matrices of nuclei and three transformation matrices. The optimal decomposition, which assumes a high compression ratio and a signal-to-noise ratio, is achieved by the block-coordinate descent method, which allows a compressed image to be obtained. To eliminate the dependence of the compression effect on the choice of initial values of the matrices, an algorithm of compressive sensitivity is used. The result set is used for onward transmission. Noises are removed except small singular values from the matrices symbolizing the tensor core.

A parallel computation method using the Tucker decomposition to compress scientific data is proposed in [10]. The composition of various methods for image processing proposed in [11, 12] is also useful.

3   Problem Statement

Image transmission via the communication channel is compressed. A unit of information in digital images is a pixel, and the entire set of image pixels is conveniently represented by a numeric matrix, so it is natural to apply matrix methods to compress images. The greatest interest among the methods of working with matrices with real values is attracted by the tensor representation in the form of a triple of matrices, one of which is diagonal and pairs of orthogonal matrices.

The matrix $A \in R^{m \times n}$ of real values is algebraically represented by a three-dimensional tensor in the form of a bilinear form, obtained from the well-known one, represented by the product of three matrices in the form of a singular decomposition of the form

$$A = M \Lambda N^T , \qquad (1)$$

where Λ is the kernel of a tensor of the same size as A, with nonnegative elements λ_i on the main diagonal (singular values) and the rest zero. The matrices $M \in R^{m \times m}$ and $N \in R^{n \times n}$ are unitary matrices consisting of left and right singular vectors, respectively.

Image compression size shrinks by grouping or dropping part of the pixels. If the grouping algorithms allow you to restore the quality of images, then discarding several pixels leads to a loss of its quality when scaled. Drop pixels in the compression process occurs based on signs of repeatability or frequency. Then the problem of compression is reduced to the definition of a triplet of smaller matrices.

The resulting matrices are sequences of numbers, so the compression process is similar to the coding process, where the input n-dimensional vector $x \in R^n$ is converted by the encoder into a code word $\gamma(x)$ of a smaller size $l = n / K$ ($l < n$), where K is the compression ratio. The discarding of a part of the information is a source of image distortions, which, as a result of compression, with identical image sizes, is evaluated by various criteria, the most common among them are the average square of errors, the signal-to-noise ratio, and visual criteria. The average error square is written as

$$e_{mse} = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} \left| \gamma(x_{ij}) - x_{ij} \right|^2 , \qquad (2)$$

where M, N are the vertical and horizontal dimensions of the image, or the signal-to-noise ratio is represented by the formula

$$R_{s/w} = 20 \log_{10} \left(\frac{\gamma(x_{\max})}{\sqrt{e_{mse}}} \right), \qquad (3)$$

where x_{max} is the maximum pixel value of the image, $R_{s/w}$ is the representation of the signal-to-noise ratio.

The purpose of the paper is to determine the matrices M, N and the kernel of the tensor D, which satisfy the specified image quality defined by formulas (2), (3).

4   Preliminary notes

Let the image A be possible to represent the original in the form (1) and there are orthogonal matrices M, N, which represent sets of orthonormal vectors $M_k = [m_1, m_2, \ldots, m_k]$ and $N_k = [n_1, n_2, \ldots, n_k]$, defining singular vectors, such that the matrix Λ contains on its main diagonal singular numbers λ_i placed in this order

$$\lambda_1 \ge \lambda_2 \ge \ldots \ge \lambda_r > \lambda_{r+1} = \ldots = \lambda_n = 0, \tag{4}$$

where r is the rank of the matrix A. As is known, they are the non-negative square roots of the eigenvalues of the symmetric AA^T matrix, T is the transposition index.

Theorem. Suppose there are matrices A, $\hat{A}$, $E \in R^{m \times n}$ such that $\| A \| \ge \| \hat{A} \|$, $\| E \| \le e \in R^+ \setminus 0$. For a matrix A of rank $k > 0$, there is decomposition in the form (3) and a recurrent procedure

$$E_{i+1} = E_i - \hat{A}_i \tag{5}$$

with $E_0 = A$ and $\hat{A}_i = \sum_{j=1}^{i} \lambda_j m_j n_j$, where m_j, n_j are the vectors of the matrix M, N, respectively, $i = 1, 2, \ldots$, allows to get the matrix A of the rank p such that

$$\left\| A \right\| - \left\| \hat{A} \right\| \le e \tag{6}$$

for a finite number of $p < k$ steps.

Proof. Using the property for unitary transformations (formula 5.7 out of [9]), we obtain the representation

$$\hat{A} = \sum_{i=1}^{p} \sum_{j=1}^{q} \lambda_i m_i n_j^T. \tag{7}$$

Since the vectors m_i, n_j in (7) are orthogonal,

$$\lim_{\substack{p \to M, \\ q \to N}} \sum_{i=1}^{p} \sum_{j=1}^{q} \lambda_i m_i n_j^T = 0, \tag{8}$$

then

$$E = E_{i+1} = A - (\lambda_1 m_1 n_1 + \ldots + \lambda_i m_i n_i) = E_i - \lambda_i m_i n_i. \tag{9}$$

Since for the applied decomposition, (4) is fulfilled, the condition for stopping the procedure is the condition

$$\left\| \lambda_i m_i n_i \right\| < e. \tag{10}$$

This proves the first part of the theorem, i.e. statement about the convergence of the computational procedure. If $i = p$, then i is the number of decomposition steps, if $i = k$, then $e = 0$. The theorem is proved completely.

5   Simulation

It is assumed that the image signal x displayed by the grayscale color model is a numeric matrix, which is further decomposed into the kernel $\Lambda = (\lambda_1, \ldots, \lambda_N)$ and the codebook D. The codebook is a memory of coefficients $\gamma(x) = (M, N) \in D$, which remain unchanged during transformations, encoding implies a reduction in the size of the core relative to the original size with permissible distortions. A threshold c is preliminarily set, according to which the distortions are considered permissible. The coding system iteratively selects the size of the kernel $\hat{\Lambda}$ that satisfies the specified threshold according to the selected quality criterion, for example (2) or (3). At the output of the comparator, we obtain the image $\hat{x}^*$, which satisfies the specified quality. In the general case, the coding system is supplemented with auxiliary elements that perform normalization operations, splitting the image into blocks. The image coding scheme is shown in Fig. 1.
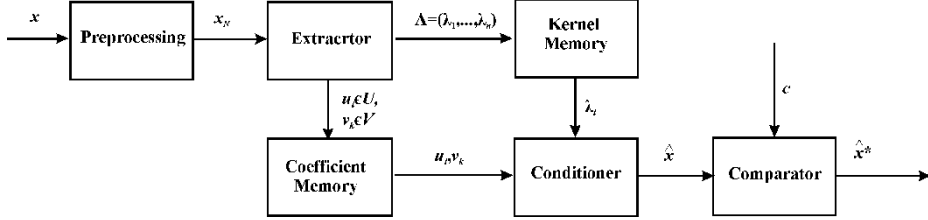

Fig. 1. Image encoder block diagram

If necessary, the compressed image can be restored by the inverse transform, if it is provided for by the compression algorithm. In this case, the dictionary is also available to the decoder. Decoding is the inverse mapping of the code word to the input vector. In loss algorithms, the original image is not restored.

The size of the dictionary is determined by the set of codewords L, using binary coding. The size of the code word is $V = \log_2 L$, bit. To simplify the computing system, the matrix of images is pre-divided into submatrices, with which it is convenient to perform calculations. The minimum size of the image blocks is 8×8 pixels. When a block of information is compressed by K times, the code word is also reduced by K times.

The simulation was carried out on a personal computer with an image of 384×384 pixels. The reference image was previously divided into sub-images of 8×8 pixels. Then, the procedure (5) was applied successively to each sub-image, to get an approximation to the original image with the specified quality, i.e.

$$x = \Lambda\gamma(x) = \sum_{i=1}^{p}\lambda_i m_i n_i^T + E = \hat{x} + E \,. \qquad (11)$$
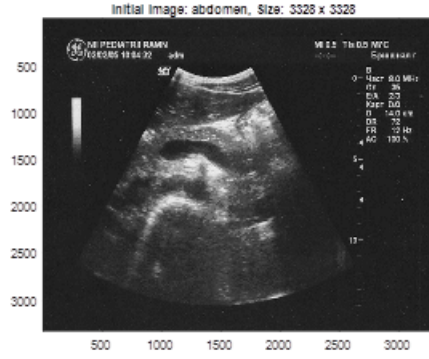
Here, E is the residual matrix obtained after obtaining p essential components of the *svd* decomposition, estimated by the Frobenius norm

$$e = \|E\|_F = \sqrt{\left\|S - M_p \Lambda_p N_p^T\right\|^2} \,. \qquad (12)$$

Image quality was controlled by criteria (2), (3).

6   Experimental Results and Discussion

The proposed algorithm was estimated for the original image obtained as a result ultrasound examination of the abdominal cavity, the view of which is shown in Fig.2.
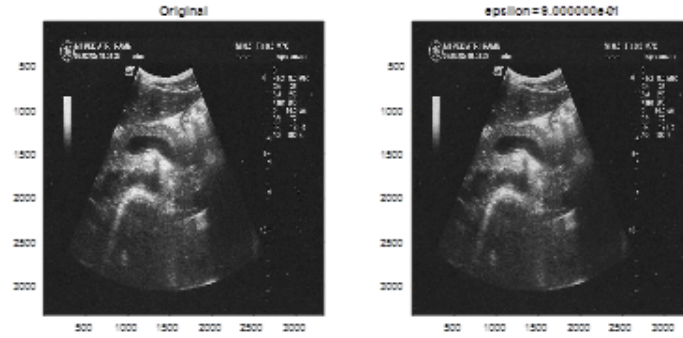

Fig. 2. Original image

Algorithm (5) was studied to test its effectiveness at different values ε image Fig. 2, calculated criteria (2), (3), made a comparison with full decomposition and partial decomposition. Images on different ε are shown in Fig. 3-6.

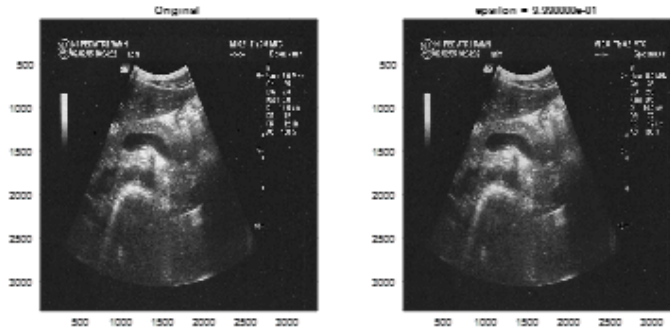**Fig. 3.** Compressed image, ε=0.1

Fig. 4. Compressed image, ε=0.9

Fig. 5. Compressed image, ε=0.999

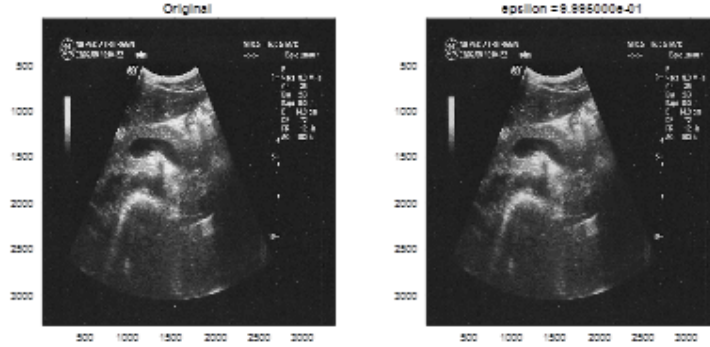**Fig. 6.** Compressed image, ε=0.9995

Comparison data are presented in Table 1. Here R is the rank of a compressed matrix, K is the compression ratio.

Table 1. Experimental results.

ε	R	e_{mse}	$R_{s/w}$	K
0	3328	$2.61 \cdot 10^{-5}$	93.971	1
0.1	2995	0.0058	70.49	1
0.3	2330	0.0058	70.49	1
0.5	1664	0.0058	70.49	1
0.7	998	0.0058	70.49	1
0.9	333	0.0058	70.49	0.96
0.999	33	0.0058	70.49	0.83
0.9995	17	0.0158	66.1364	0.786

Analysis of the obtained results allows you to confirm the performance of the algorithm. Acceptable image quality was maintained by reducing the size of the decomposition core matrix by 200 times. Low compression ratios were obtained because of additional conversion of the reconstructed image to the format JPEG. Better compression can be achieved by preserving the triad of transformation matrices in numerical form, using, for example, the TXT format. The best image quality when visually evaluating the quality of compression is obtained at small values of ε, as it should be. The proposed algorithm is useful for compressing medical images.

7 Conclusion

The study showed the effectiveness of using singular decomposition for storing and transmitting in the network medical images typical of non-invasive research methods. As a result of the study, it has been found that matrix methods based on decompositions that can be combined with other methods, such as grouping, traditionally used in hybrid compression methods, are quite effective for image compression. Higher image quality is achieved by increasing the accuracy of the truncated description approaching the original image. The developed algorithm has the property of adaptability. Further research is planned to be devoted to searching methods for finding the unknown tensor core for given unitary matrices M, N.

References

1. Salomon, D., Motta, G.: Handbook of Data Compression. 5th edn. Springer-Verlag, London (2010).
2. Sadek R.A.: SVD Based Image Processing Applications: State of The Art, Contributions and Research Challenges. International Journal of Advanced Computer Science and Applications, 3(7), 26 – 34 (2012).

3. Abbadi N. K. El., Kadhim N. E.: Brain Tumor Classification Based on Singular Value Decomposition. International Journal of Advanced Research in Computer and Communication Engineering, 5 (8), 553 - 557 (2016).

4. Yang, J.-F., Lu, C.-L.: Combined Techniques of Singular Value Decomposition and Vector Quantization for Image Coding. IEEE Transactions on Image Processing 4 (8), 1141 – 1146 (1995).

5. Waldemar, P., Ramstad, T.: Hybrid KLT-SVD Image Compression. In: EEE International Conference on Acoustics, Speech, and Signal Processing, pp. 2713 – 2716, Munich, Germany (1997).

6. Tian, M., Si-Wei, L., Ling-Zhi, L.: An Investigation into Using Singular Value Decomposition as a Method of Image Compression. In: Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, pp. 5200 – 5204, Guangzhou (2005).

7. Boque A. T. Tensor Networks for Image Compression. http://diposit.ub.edu/dspace/bitstream/2445/96365/1/TFG_Fis_Trujillo_Boque_Alex.pdf, last accessed 2019/02/08.

8. Rövid, A., Szeidl, L., Rudas, I.J., Várlaki, P.: Image Processing on Tensor-Product Basis. Óbuda University e-Bulletin 2 (1), 247 – 258 (2011).

9. Hassanzadeh, S., Karami, A.: Compression and noise reduction of hyperspectral images using non-negative tensor decomposition and compressed sensing. European Journal of Remote Sensing, 49 (1), 587 – 598 (2016).

10. Austin, W., Ballard G., Kolda T.G.: Parallel Tensor Compression for Large-Scale Scientific Data. In: IEEE International Parallel and Distributed Processing Symposium (IPDPS), pp. 912 – 922, Chicago, IL, USA (2016).

11. Kucherov, D.P., Katsalap, R.G., Zbrozhek, L.V.: Improving Images Quality by Combination of Filtering Methods. The International Journal of Engineering and Science (IJES) 4 (7), 69 – 77 (2015).

12. Kucherov D.P.: A computer system for images compression. http://er.nau.edu.ua/handle/NAU/38657, last accessed 2019/09/27.

13. Golub, G.H., Van Loan, C.F.: Matrix Computations. The John Hopkins University Press, 4th edn. Baltimore, USA (2013).

14. Jain, A.K.: Fundamentals of digital image processing. Prentice Hall, USA (1989).