International Journal on
“Technical and Physical Problems of Engineering”
(IJTPE)
Published by International Organization of IOTPE

ISSN 2077-3528
IJTPE Journal
www.iotpe.com
ijtpe@iotpe.com

December 2023 Issue 57 Volume 15 Number 4 Pages 34-45

PROPOSED ALGORITHM FOR VIDEO SEQUENCE COMPRESSION

V.Z. Kuklin A.K. Lampezhev L.M. Chervyakov A.A. Tatarkanov

Institute of Design and Technology Informatics, Russian Academy of Sciences, Moscow, Russia
kuklin_vladimir_ran@mail.ru, abas.lampezhev@mail.ru, leonid.chervyakov@bk.ru, as.tatarkanov@yandex.ru

Abstract- Creating a method and simulating video information compression is one of the today’s most urgent tasks, especially given a number of shortcomings inherent in the existing algorithms. For example, the creators of the H.264 format sought to improve compression. The use of the MPEG-4 Visual format is focused on optimizing variability by moving to a subject matter that has the necessary flexibility. For this reason, the mentioned formats are based on completely different principles that provide the ability to compress video information. In turn, this feature has some disadvantages that need to be overcome in the framework of this research. The improvement of techniques that allow compression of various video sequences based on a combination of the proposed ordered chains of operations expresses the purpose of this study. Next, this research will involve an ordered chain of operations to create separation masks, and an ordered chain of operations to provide interpolation of incoming pulse reports. The proposed ordered chains of operations improve the quality of the reconstructed video sequence. In addition, the computational complexity will increase slightly, but it will be possible to increase the compression, slightly reducing the quality. Successful simulation of the video information codec makes it possible to modify the technique for compressing a video sequence, increasing the compression ratio, improving the quality level of the reconstructed video sequence, and optimizing the replacement of motions without resorting to an increase in the boundary level of calculations.

Keywords: Compression Technique, Compression Index, Motion Compensation.

1. INTRODUCTION

A broad spectrum of electric machines is widely used in electromechanical systems. In addition to the required. Information and computer systems are characterized by the active use of multimedia technologies [1]. For this reason, there is a need to create methods, and ordered chains of operations that provide compression of any digital video image and video stream, based on the transformation of each class of images that are characterized by information redundancy [2]. The constantly increasing complexity of calculations that provide algorithms for converting video sequences, and a significant level of costs associated with storing information require significant computing resources [3].

The use of video compression allows the use of various digital video formats in transmission media that are incapable of supporting uncompressed video images, and optimizing the use of any high-speed communication means to transmit the required high-resolution video stream, including the possibility of parallel transmission of a large number of video streams high quality data [4].

The existing developments for processing video information are based mainly on compression techniques [5]. At the same time, in this case, some losses are observed due to the transmission of the main frame (I-frame), which is compressed with respect to volume coordinates. There are also some frames that are compressed relative to spatial and temporal coordinates, and their number will determine the level of compression of any video sequence. Developers must consider the specifics, or the vision of such an option for presenting information by the human eye [6]. At present, the redundancy of the reference frame has been eliminated one way or another. At the same time, such a problem of reference frames does not yet have any definite solution [7]. This trend is based on the evaluation of all available methods for video information compression. Nowadays, a researcher can operate with several formats that provide compression of video information. Such standards formed the basis of the predominant part of the schemes that provide encoding and decoding. Thus, H.264 and MPEG-4 Visual deserve special attention; these formats were created owing to the efforts of specialists from all over the globe [8]. This trend is based on the evaluation of all available methods for video information compression. Nowadays, a researcher can operate with several formats that provide compression of video information. Such standards formed the basis of the predominant part of the schemes that provide encoding and decoding. Thus, H.264 and MPEG-4 Visual deserve special attention; these formats were created owing to the efforts of specialists from all over the globe [9].

The vectors that describe the motion of any individual pixels, and their sets form the basis for the encoding each reference frame. Moreover, the degree of performance of ordered chains of operations providing elimination of temporal redundancy can be optimized by improving the ordered chains of operations that provide encoding. This issue is considered in our research.

2. LITERATURE REVIEW

Before dwelling on any complex capable of performing image chain processing, it is necessary to define terms such as spatial and temporal image chain. These definitions are best illustrated by the following example. We have a certain sensor that in three-dimensional space can record information about changes in the surrounding space. Further, using this information, images are created that are formed in certain sequences. It makes sense to represent this set, covering all frames, with the dependence of brightness $I(x,y,t)$, here x and y are coordinates belonging to the image, t is a coordinate relative to the time axis. Using this approach will help to demonstrate the nature and behavior of any temporal image chains.

Suppose that the sensor can be moved and is able to rotate in three dimensions. It should also be considered that even if the sensor rotates around its axis, this will not affect the resulting image that it will broadcast. Assuming the existence of a chain that has no end, such sensors can cover every possible configuration that it can take. With regard to these assumptions, any image is worth displaying as a unit of an image chain. After some time, a group of these images will be able to create a space that includes all available images. In this case, we will observe a transition to a dependence consisting of 4 main points, or coordinates $I(x,y,t,\vec{s})$, here the 5D vector $\vec{s}$ secures the location of the sensor at any point of three-dimensional space. In other words: $\vec{s} = (\bar{x}, \bar{y}, \bar{z}, \beta, \gamma)$, here, certain coordinates help fix the location of the base point of our sensor, and the angle of its location relative to the considered axis. It is worth regarding the fact that x, y, and t are dependencies of the vector $\vec{s}$. The use of this model, which best describes the behavior of the subject of evaluation, makes the processing of video information more difficult. Consequently, simplified time models are more common in the study of this issue. In this case, spaces covering all images will be created as a generality of temporal chains of images that are captured by sensors, in each of all possible 3D positions.

At the same time, it is worth considering the existence of differences between the chain of images, and the video sequence. The video sequence is only the observed fragment of the electromagnetic spectrum. While considering a chain of images, this limitation is not provided [10]. Consequently, the width of the range of image chains will be greater than that of video sequences; however, if we evaluate the visible part of the spectrum, there are simply no differences between these definitions.

Any of the frames that form a regular video scene is a kind of grid. The values of any sample in its nodes are characterized by a sufficiently high level of interconnection. Therefore, the compression of its initial form is aggravated by additional complexities associated with the computation. Ordered chains of operations that provide video information compression apply the specifics of processing initial information, for example, the redundancy of data, how smoothly it changes, and the specifics of seeing reality with the human eye, in other words, the low degree of sensitivity of human eyes to slight distortion during image reconstruction. The latter is often used in algorithms that provide compression, usually due to a certain percentage of loss in quality [11].

Discussing the specifics of a person's vision of the surrounding reality, the color planes of any images will be characterized by some excess. This is called "redundancy of color spaces". In fact, vision is largely determined by the brightness of the picture. For example, let us replace RGB with YUV with thinning of several components. This replacement opens up the prospect of using digital video formats in an environment that does not support uncompressed video formats [12]. In addition, the use of video compression optimizes the use of high-speed communication channels for the high-resolution video stream transmission, and is also used for parallel transmission of a small amount of video information with a high stream quality. Richardson believes that compressing video sequences will be a problem for quite long time to come [13].

3. METHODS AND EXPERIMENT EFFECTIVENESS

The exclusion of temporal redundancy is based on the hypothesis that in a short time interval, which corresponds to two or three frames, the objects that are part of the video scene will not change significantly. For this reason, a pixel-by-pixel difference in both frames should not be considered, even regarding the fact that the use of compression of the mentioned difference will to some extent limit the ordered chain of operations for decompression and compression, but this principle is applied in most algorithms that provide compression of video information. However, the redundancy of each color plane is caused by the fact that the human eye will primarily perceive the pixel brightness level. The compression process using a video sequence compressor is of interest (Figure 1) [14].

In this case, it is worth designating such categories as spatial and temporal models. The second option aims to reduce temporal redundancy. At the same time, the first option operates with the similarity of neighboring frame samples, using the reduction of spatial redundancy.

The start of encoding is determined by the fact that incoming frames from color representations are converted into color difference ones. For example, the RGB color representation of pixels can be represented by three components, which indicate the interdependence of the three basic components that determine visible light. The color representation of YCbCr, and its derivatives (YCC YIQ, YUV,) are based on the brightness index that the human eye can perceive. Because of the representation thus obtained, sampling is performed to increase compression. In other words, the minimum resolution of components is applied, except for such a component as brightness.

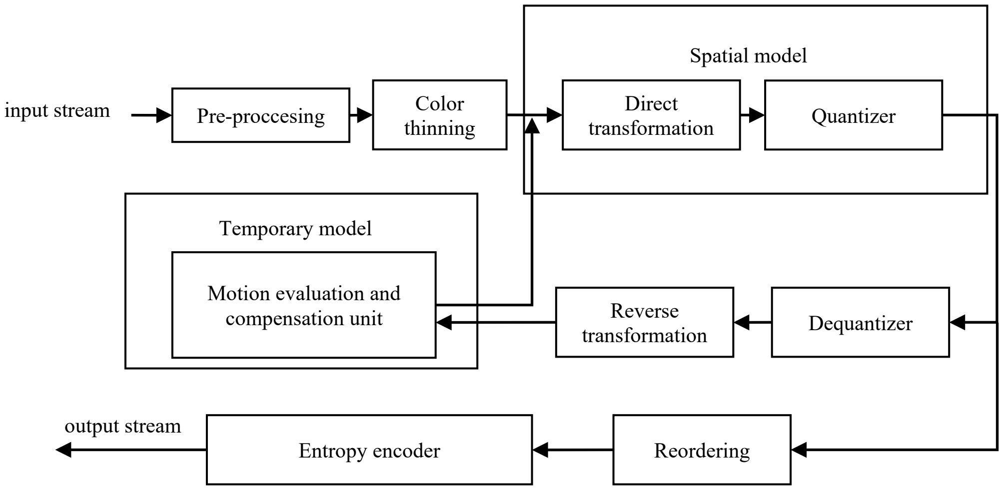


Figure 1. Schematic representation of the method for video sequence compression [14]

Any video series is pre-processed. A whole cascade of filters deprives this chain of irrelevant details. This procedure makes it possible to increase compression, eliminate the negative components of high frequencies, and a number of fast camera movements. Often, in parallel with the methods for primary transformation at the stage when the decoding of the encoded signals is performed, to optimize the reconstructed frames, cascades of filters are used that provide post-processing. The specificity of the functioning of these cascades is covered in [15-17]. Then all stages of gamma correction are implemented. Since a person sees changes in brightness non-linearly, it is necessary to scale pixel brightness using a power law.

If we assume the simplest case, when the next frame is compressed, it ends up on the input channel of the converter. Due to the input stream of the transducers, it is possible to make several samples of the initial pulse. Therefore, on the output channel, we will obtain the indicators of the pulse decomposition relative to the basis. In the opposite case, we will be dealing with a different type of frame; for this a different variant of estimating and compensating for motion will be involved, and only then the transformation will be performed. At this stage, information about frames in a different coordinate plane is displayed. This plane is called the transformation plane [18-20].

There are several types of pulse converters [21], regarding:
1) time
2) frequency, and frequency and time
3) qualities, or characteristics of the impulse

The basic techniques for coding by transformations are divided into a coding variant that does not provide for predictions, and a truncated block coding variant [22], a coding method that provides prediction [23] and delta modulation [19]. The pulse code modulation (PCM) variant assumes that the compression will be performed by sampling the amplitudes of the components. At the same time, the block truncation coding (BTC) is based on splitting the pulse into several blocks in order to select two levels of the pulse, which will allow encoding each of these blocks. The disadvantage of both variants (the BTC and the PCM) is a clearly insufficient level of redundancy exclusion.

The potential differential pulse code modulation (DPCM) that is applied by the AAC format is based on the transformation of the differences obtained between each of its samples, which is distributed linearly. The basic methods for encoding by means of transformation by modeling the initial pulse, include the variant of fractal compression [24], the variant of geometric approximation, etc. These methods for evaluating a frame to establish its key qualities, or characteristics, are characterized by significant computational complexity. The video sequence that has undergone transformation is at the quantizer input.

Ordered chains of operations that ensure the discretization of each pulse sample are:
- Scalar [25]
- Vector
- Grid ones

The variant, expressed as a single number, assumes that any pulse sample will be split into sample spacings. The simplest option implies that values, which appear in the "dead zone" will be rounded to zero values. In the course of sampling, the value of any samples will be replaced by the number of the interval in which they appear. When considering the decoding option, the numbers will change to centroids – the averaged values of each pulse report on any interval. Such sampling may be applied to reduce the fidelity of each image upon completion of encoding with the transformation.

With vector sampling, the initial pulse is split into several rectangular parts. Then these parts are combined regarding their similarities, and some other characteristics. After that, for such a combination, an averaged area is set, which is placed in a certain cell of the dictionary table. When decoding is performed, each span will be replaced by certain dictionary coefficients. If uniform pulse distortion is observed, this ordered chain of operations will reduce the relationship of pulses relative to scalar ordered chains of operations. At the same time, this process is characterized by a significant computational complexity of creating a dictionary, which is a serious factor hindering the use of this method. But considering the specifics of compression, it is possible to use the dictionaries of previous images to create new ones, which helps simplify this operation. The disadvantages that are inherent in vector sampling include the need to broadcast simultaneously code books with encoded messages.

The possibilities of the grid quantizer are of interest. It is a derivative of vector quantizers, but has all the qualities and characteristics of scalar quantizers. The coefficients of the sample values will depend on the values of several nearby samples, and their calculation will be carried out on the basis of a set rule. Upon completion of the sampling, the indicators are redistributed among the classes including elements whose values are greater than zero. The best scan sequence will depend on the distributions of indicators whose values are greater than zero. For example, for standard blocks of images, the normal scanning sequence can be expressed as a zigzag when the scanning starts in the upper left region. The motion evaluation and replenishment reduce the level of interrelationships between samples.

This feature opens up the possibility of compressing a chain of images relative to the initial video frame. These models mainly use ordered chains of operations that provide coding involving prediction. At the same time, encoders predict based only on previous or future frames, from which the current area is subtracted. This option is called forecasting from current areas. The characteristics of various models are fed to entropy encoders. The received information can be in the form of markers, and headers. Further, when video sequences pass through entropy encoders, these chains are subjected to element-by-element compression. This possibility appears due to the use of data on the occurrence of individual chains of signs. The main types of coding are arithmetic coding, and a number of modified codes of variable length, proposed by Huffman [26].

3.1. Creating a Video Codec Model

Currently, researchers have a set of ordered chains of operations that provide compression of video sequences that are described by video compression formats, which form a base for almost all encoding and decoding schemes. Evaluation of common methods for efficient compression of video sequences reveals several interesting patterns. Thus, the MPEG-4 Visual format was developed for high variability; however, at present, only some features of this format are used in practice. In addition, there is a relatively low ability to reduce the averaged information flow densities. The blocking effects, supplemented by high levels of video information compression, should be attributed to the disadvantages of this format. At the same time, it is especially difficult to determine exactly the forms of the object, with respect to which prediction is performed in ordinary video scenes.

The H.264 format can be described as non-variable. For this reason, it has a high level of compression against the background of acceptable quality. The format requires the use of a professional camera, significant calculations, and a royalty. Regarding the advantages of segmented motion compensation, the creators of this format do not consider the data characterizing the commonality of the blocks united by the frame [27]. The key performance indicators of VP8 cannot match the potential of H.264. In fact, the capabilities of this video compressor are still being developed and require the use of special ordered chains of operations.

This research proposes a method for compressing video information, based on a combination of created ordered chains of operations, as shown in Figures 2 and 3. Incoming video streams appear on the input channel of the created model. If this is not the case, the frames will be reference, and temporal models will be involved, after which the video streams will be combined in the form of a combination of the residual frame and the vector describing the motion. The possibilities of entropy coding are manifested due to the group of modified codes of variable length proposed by Huffman [26]. In this case, the dotted arrow shows the relationship between this block and the spatial and temporal models. Let us dwell on the key components of the created method.

3.2. Temporal Models

In this case, the forecast can be made by:

1. Intraframe (intra) prediction
2. Interframe (inter) prediction

When considering intra forecasts, the magnitudes of the pulse samples are predicted relative to the values of adjacent samples. When considering inter forecasts, the prediction is made due to the coding of frames using an ordered chain of operations Pol. Any residual frame appears in the device performing the selection, where the total prediction error is analyzed, relative to the output of all modes, as the sum of the values of each residual frame sample. The result regarding any frames is fed to the decoder. If the frames belong to a group of frames of a separate forecast, then they are transmitted to the input channel of the considered model.

3.2.1. Operation of the Pol Algorithm

An ordered chain of Pol operations was created, enabling to form a separation mask. The appearance of this ordered chain of operations is shown in Figure 4. The functioning of this ordered chain of operations is based on the "non-merging" of interests with respect to a conditional boundary characterized by a width equal to one pixel.

By arranging lines limited in length, boundary lines and areas that have any brightness differences are distinguished. It was proposed to build the execution of an ordered chain of operations regarding a frame with a reduced resolution to optimize the chain complexity. This option, which is confirmed by the study, does not actually affect the quality of splitting.

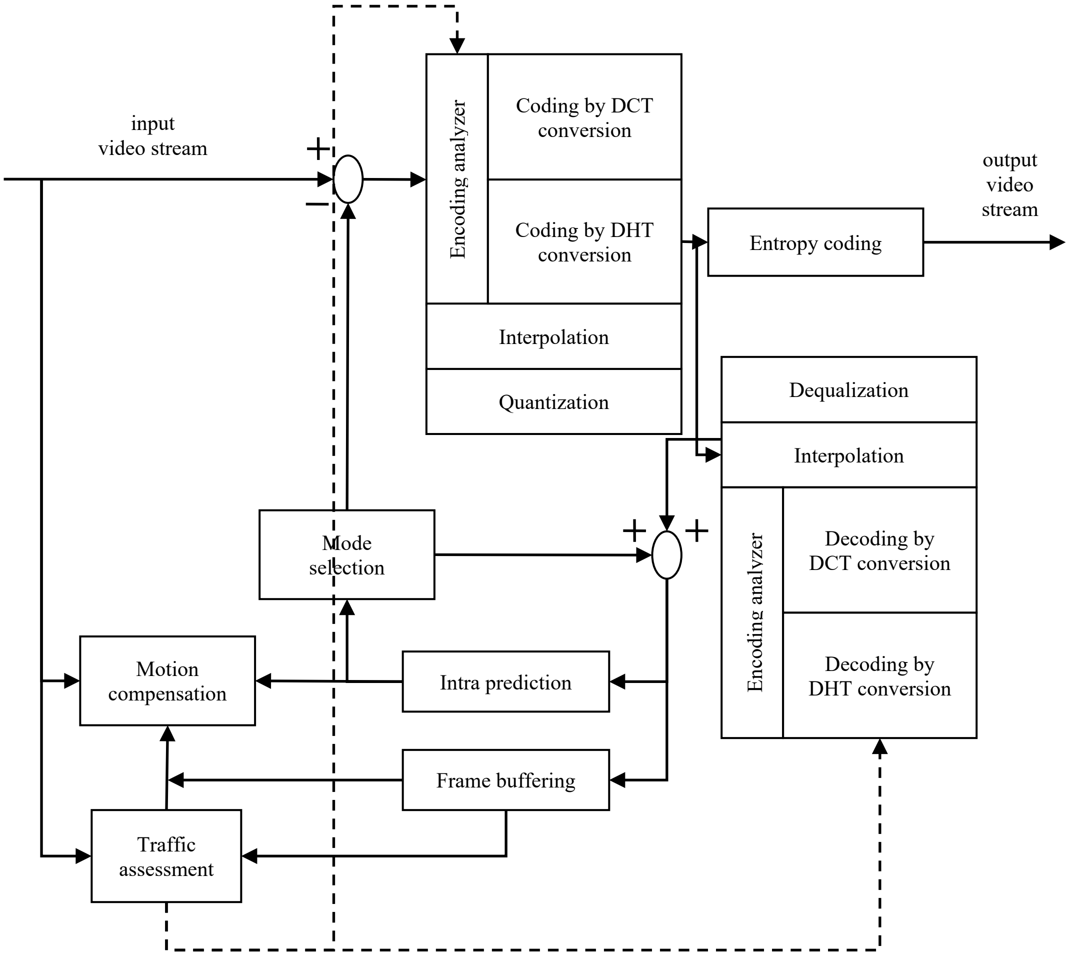


Figure 2. Encoder model schematic

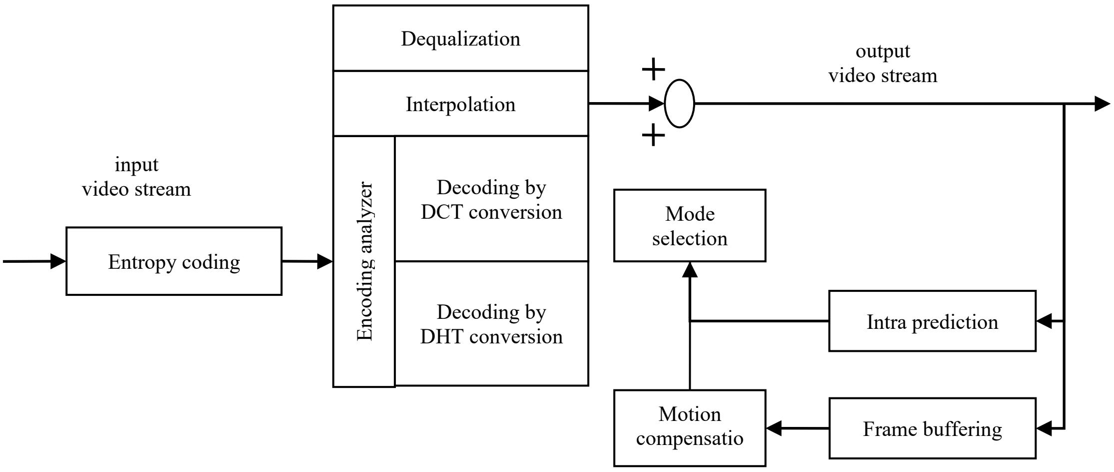


Figure 3. Decoder model schematic

3.2.2. Evaluation of Each Split Algorithm

Today, there are several approaches that make it possible to segment the frames of video sequences. Thus, segmentation can be based on:

1) fixed size block matching (FSBM) [28]
2) variable size block matching (VSBM) [29]

Evaluation of the segmentation conducted using blocks of a certain size shows that all frames of the video sequence are divided into a certain number of blocks.

For them, a reference frame of the block is selected, which corresponds to the initial blocks, occupying a certain area, which corresponds to the largest offset applied by the encoder. The standard block configuration corresponds to 16×16 pixels. In this case, the largest offset can be characterized by a value of ±64 pixels. The selection of the optimal variation demonstrates a small forecast error.

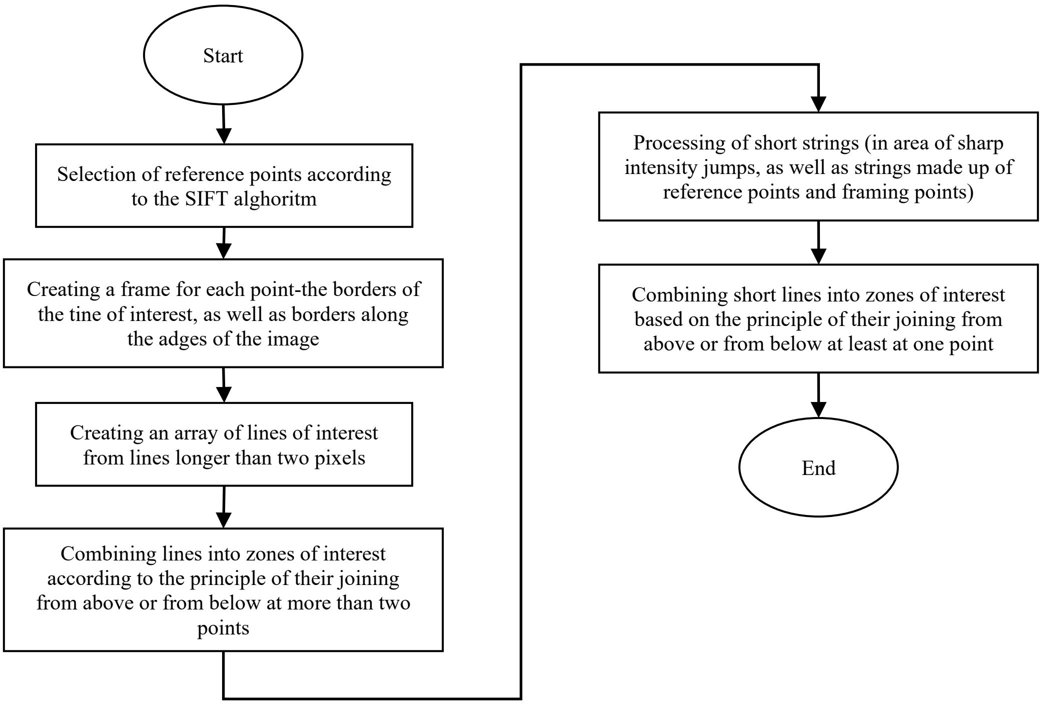


Figure 4. Schematic representation of the sequence of stages in the functioning of an ordered chain of Pol operations

Acceptable compliance during the matching process ensures good prediction accuracy. However, if this is not the case, any modernization of this forecasting method will increase the price for transmitting motion vectors. To obtain an acceptable matching performance using this ordered chain of operations, it is necessary for any block to pass through such a scheme. It is also necessary for it to be characterized by a relatively small shift, which would not overlap objects in the image, characterized by the presence of several degrees of freedom. Considering all shortcomings of the ordered chain of FSBM operations mentioned above, several methods are proposed that can modernize it by flexible changing the configuration of each block to display changes in any part of the images with greater accuracy. This method is referred to as motion compensation because of using variable size blocks (VSBM).

As a result of researching an ordered chain of VSBM operations, and improved chains of VSBM operations by selecting a key pixel and grouping a vector, relative to the direction of motion, several functions were obtained:

- The highest value of signals/noise relative to all configurations of the block under consideration.
- RD is a characteristic demonstrating the degree of pulse change (peak signal-to-noise ratio, PSNR) relative to the indicator that describes the compression (*R*(*D*)).
- The complexity of computing and ordered chain of operations (*Q*), which is measured by the average number of base operations (BO) to blocks, relative to the list of all its configurations.

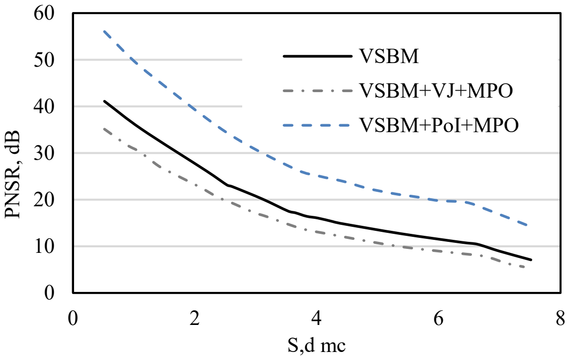
With regard to the tested chain, shown in Figure 5a, we observe the function of the *PSNR* value to the plane £. The dimensions of the plane will increase, and any increase will have a step *d*, relative to any side.

Figure 5b shows the computational complexity of an ordered chain of operations *Q* with respect to *S* and *d*. Figure 5c shows the *R*(*D*) function.

The created ordered chain of operations showed better performance than the known ordered chains of operations, and VSBM. Moreover, this refers precisely to the quality of the reconstructed video sequence and the level of compression, while the complexity of the calculations remained virtually unchanged.

Using this ordered chain of operations makes it possible to:

1) improve the quality of the reconstructed video sequence by 18-22%. Undoubtedly, the computational complexity will increase insignificantly due to the improvement of this chain of operations to eliminate the aperture problem;
2) optimize compression by 7-9% relative to known ordered chains of operations with similar computational complexity;
3) apply various manipulations to the video stream data without alpha masks.



(a)

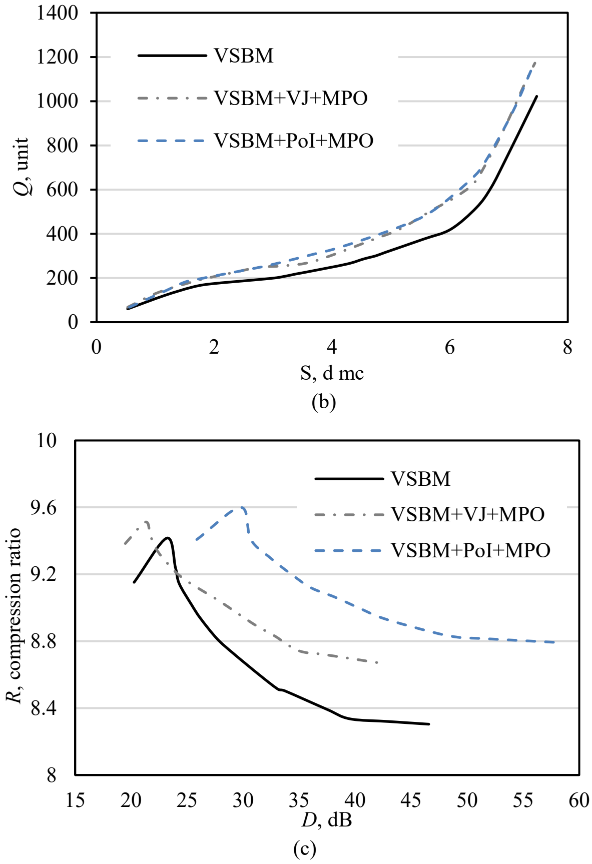


Figure 5. Schematic representation of the function for ordered chains of the split operations VSBM+Pol+MPO; VSBM+VJ+MPO, and an ordered chain of VSBM operations

3.3. Operation of the Spatial Model

In this model, the pulse is on the input channel of the encoder, then each pulse sample is interpolated using the proposed ordered chain of operations. Redistribution is performed using zigzag scanning, sampling is performed by an ordered chain of code sampling operations, which is a kind of grid sampling, where a reliability vector is applied to each bit plane of frames.

3.3.1. Determining an Ordered Chain of Operations by Inserting Pulse Samples

In addition to the main ordered chains of operations that provide interpolation, several ordered chains of operations based on flexible pulse interpolation were created (Figure 6). Such an ordered chain of operations functions as follows: each indicator of the main sample corresponds to a single dependence:

$$f^0(2m+1.2n+1)=\frac{[f(2m,2n)+f(2m+2.2n+2)]}{2} \quad (1)$$

$$f^1(2m+1.2n+1)= \\ =\frac{[f(2m,2n)+f(2m+2.2n)+f(2m,2n+2)+f(2m+2.2n+2)]}{4} \quad (2)$$

$$f^2(2m+1.2n+1)=\frac{[f(2m+2.2n)+f(2m,2n+2)]}{2} \quad (3)$$

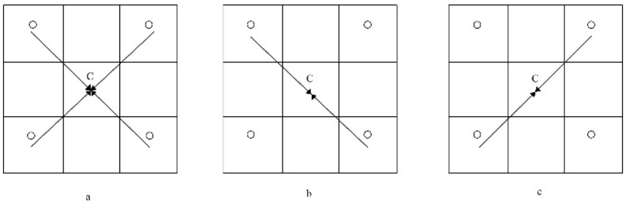


Figure 6. Schematic representation of the flexible interpolation of the main sample, here o is a reference sample, c: basic sample, and the remaining ones are extreme samples

Dependency matching will consider the characteristic:

$$\mu(2m+1.2n+1)=\left|f(2m,2n)-f(2m+2.2n+2)\right|- \\ -\left|f(2m,2n+2)-f(2m+2.2n)\right| \quad (4)$$

Based on the values α_1 and α_2 of the decision rules that determine the indices of interpolating dependencies relative to the value of this characteristic:

$$f(2m+1,2n+1)=\left\{\begin{matrix} 0\ ,\ \mu<\alpha_1 \\ 1\ ,\ \alpha_1\le\mu\le\alpha_2 \\ 2\ ,\ \mu>\alpha_2 \end{matrix}\right\} \quad (5)$$

The need for complex training helping to find these characteristics is the disadvantage of this ordered sequence of operations. At the same time, the values of the decision rules are matched, regarding the reduction of the averaged errors:

$$\sum\nolimits_{(m,n)\in I}\left|f'(m,n)-f(m,n)\right|\to\min_{\alpha_1,\alpha_2} \quad (6)$$

where, I is a series of coefficients corresponding to each reference sample.

Schematic representation of the created ordered chain of operations is shown in Figure 7. Considering the specifics of video information processing and aiming to minimize the complexity of calculations, we proposed to report the numbers of interpolating dependencies to each block as a two-bit code, which is determined during encoding. An ordered chain of operations that provides interpolation using flexible kernel sizes is a recursive ordered chain of operations. In it, all basic, main, and extreme samples will be calculated by a flexible ordered chain of operations. Furthermore, the direction of the forecast will be set for the entire block, regarding the decrease in averaged errors. Then they are assigned code words having a length of two bits - for each of the blocks.

We proposed to create an interpolation residual only for the reference frame. In this case, the shooting frames should be converted to zero. It is important to consider that blocks with different configurations, which are obtained during the estimation, and the motion compensation, must use certain interpolation kernels. Their use will provide necessary compatibility in process of joining several blocks characterized by different sizes.

3.3.2. Evaluation of an Ordered Chain of Operations Providing Pulse Sample Interpolation

By carefully evaluating the ordered chains of operations that provide interpolation, as shown in Figure 8, the best compression was confirmed, even regarding a slight decrease in the quality of the reconstructed video sequence.

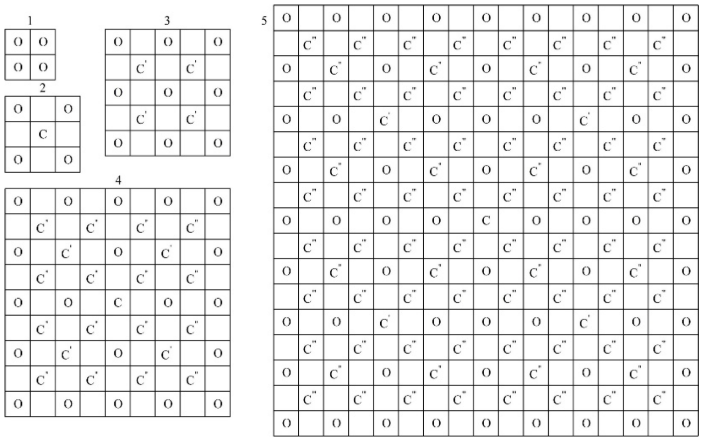
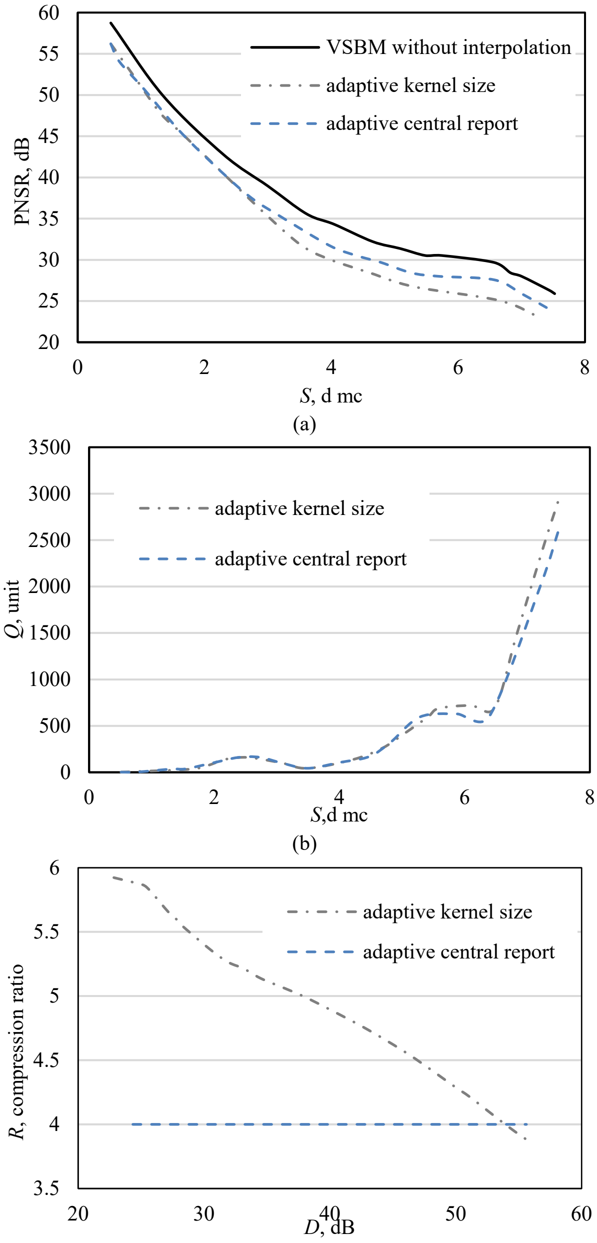


Figure 7. Schematic representation of interpolation with flexible kernel sizes: 1: 2×2; 2: 4×4; 3: 8×8; 4: 16×16; 5: 32×32



(a)

(b)

(c)

Figure 8. Schematic representation of function for ordered chains of operations that provide interpolation of flexible base samples, of VSBM without interpolation, regarding the flexible kernel size

Finally:

- Pulse matching opens up the prospect of blocking reduction.
- It becomes possible to increase the level of video sequence compression by 18-22% using ordered chains of operations that provide interpolation of pulse samples.
- The created ordered chain of operations, which provides interpolation of pulse samples, enables to increase the compression level by 28-30%.

3.3.3. Motion Analyzer

If the frames in their group are static, only the reference frames can be transmitted to the decoder. The remaining frames will be restored by duplication by the decoder. But if the frames in their group are characterized by dynamism, it is necessary to apply a two-dimensional ordered chain of operations. We propose to use a device that performs motion analysis. Its functioning can be described by the relation:

$$cube = \left\{ \begin{array}{l} \text{static, if } M_1 < T_1 \text{ and } M_2 < T_1 \\ \text{dynamic, if } M_2 > T_2 \\ \text{otherwise of overage motion} \end{array} \right\} \qquad (7)$$

where, T_1 and T_2 are the threshold values of the device that performs motion analysis, they are defined as follows:

$$M_1 = \max(m_1(0,0), m_1(0,\frac{M}{2}), m_1(\frac{N}{2},0), m_1(\frac{N}{2},\frac{M}{2})) \qquad (8)$$

$$M_1 = \max(m_2(0,0), m_2(0,\frac{M}{2}), m_2(\frac{N}{2},0), m_2(\frac{N}{2},\frac{M}{2})) \qquad (9)$$

$$m_1(a,b) = \max_{t\in[0...P-1]} \sum_{u=a}^{a+\frac{N}{2}-1} \sum_{v=b}^{b+\frac{M}{2}-1} \left| c'(u,v,7) - c(u,v,t) \right| \qquad (10)$$

$$m_2(a,b) = \max_{t\in[0...P-1]} \sum_{x=a}^{a+\frac{N}{2}-1} \sum_{y=b}^{b+\frac{M}{2}-1} \left| c(u,v,0) - c(u,v,t) \right| \qquad (11)$$

where, c, and c' are the pixel brightness values recorded on the current and previous frames.

The current encoding analysis (AC) device has many useful functions. With regard to the threshold values corresponding to 4, and 14, 2D or 3D models are matched. At the same time, several auxiliary thresholds T_3, and T_4 can be involved. In this case, the formation of the encoding analyzer will consider the requirement (5).

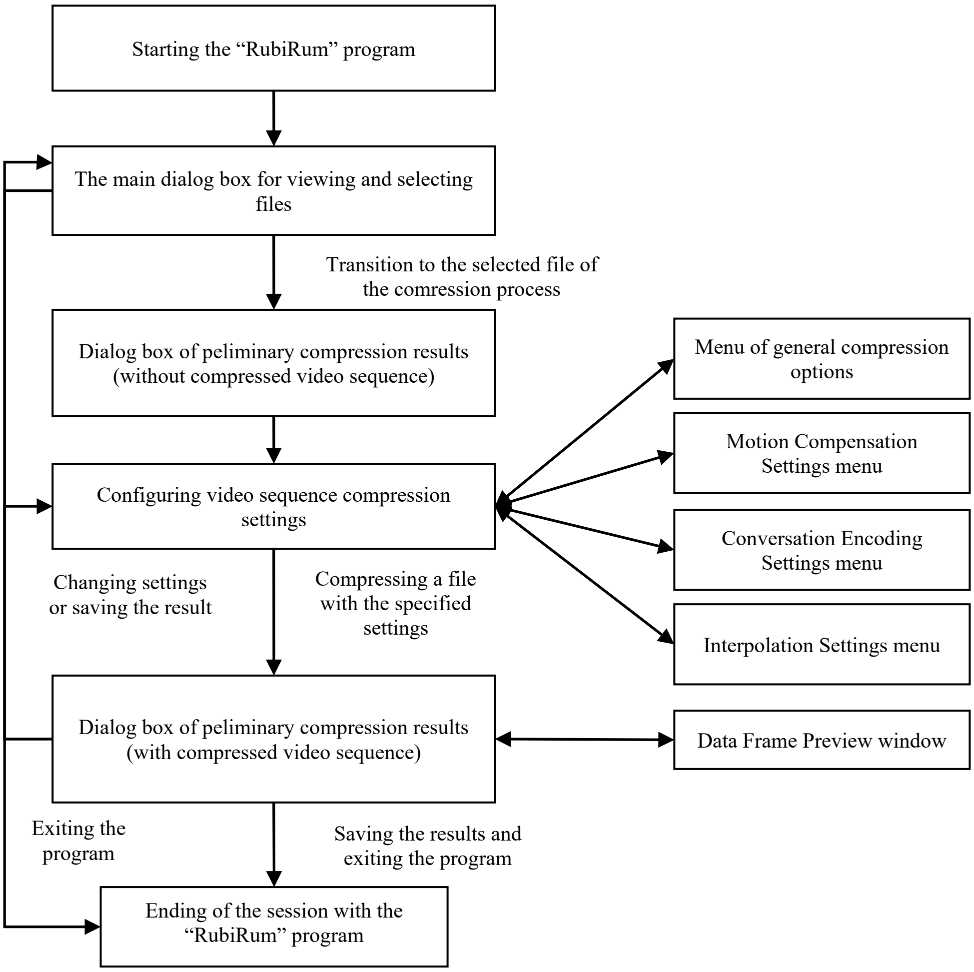


Figure 9. The interaction schematic

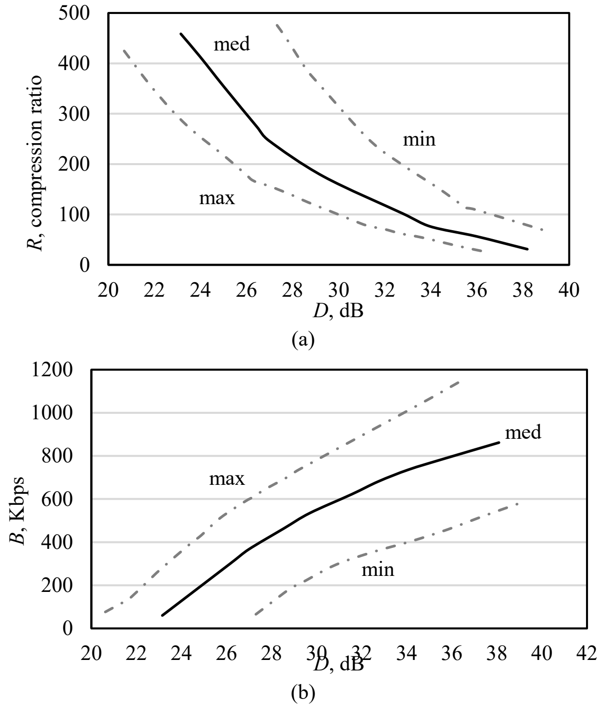
3.4. Evaluation of Codec Software

Several Visual C++ tools were developed and created. The order of their functioning is shown in Figure 9. The use of these software tools opens the possibility of:
1) selecting key coding characteristics;
2) visualizing the results of evaluation, and motion compensation by frames;
3) evaluating key compression characteristics;
4) performing video sequence compression relative to the specified list of characteristics.

4. EVALUATION OF THE RESULTS

Each result of the operation of the proposed method for compressing video sequences is shown in Figures 10, 11, and 12. At the same time, the functions describe two extreme cases with respect to the chain, characterized by minimum (min) and maximum (max) dynamism.

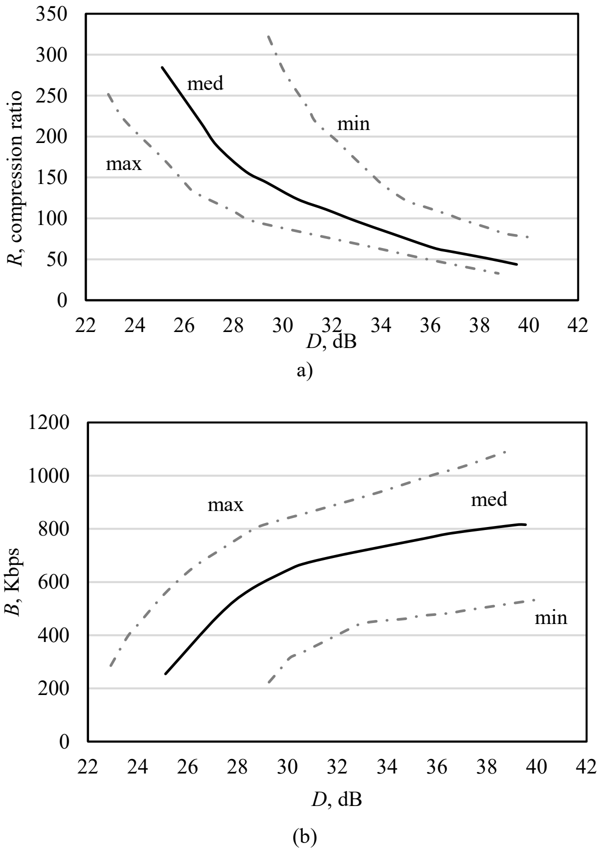
A number of created software tools were applied to compress and reconstruct artificial and conventional video sequences.



(a)

(b)

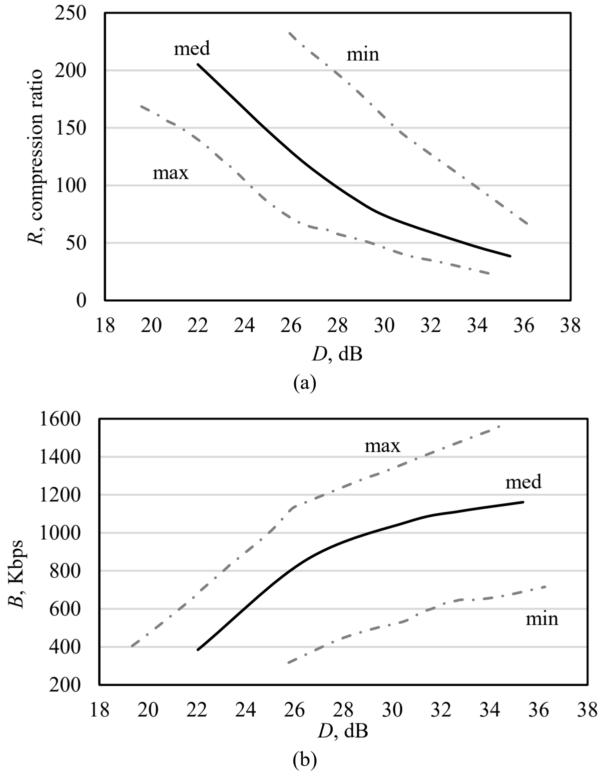
Figure 10. Schematic representation of the function of the compression index (a) and bitrate (b) relative to the quality level of the reconstructed video sequence as applicable to the H.264 codec (ffmpeg)



a)

(b)

Figure 11. Schematic representation of the function of the compression index (a) and bitrate (b) relative to the quality level of the reconstructed video sequence as applicable to the VP8 codec (WebM)



(a)

(b)

Figure 12. Schematic representation of the function of the compression index (a), and the bitrate (b) relative to the quality level of the reconstructed video sequence as applicable to the created method (RubiRum)

5. CONCLUSION

The main scientific result of the research is the modernization of methods for compressing artificial and conventional video sequences. This development provides an increase in the compression rate, improves the quality of the reconstructed video series, with regard to an acceptable level of calculations. Finally, this can be considered as a solution to several urgent scientific and technical problems in the field of telecommunications.

1. A method was created for compressing video sequences, based on a combination of several ordered chains of operations. The method allows to improve the quality indicators of the reconstructed video sequence by 4-5%, its compression level by 28-32% relative to the VP8 technique, reduce the bitrate by 28-32%, and also reduce the total cost of computing by 18-22%.
2. The research results show that the efficiency of the created video in some respects exceeds the capabilities of the video compressor, made considering the H.264 format.
3. Several software tools were created to implement the proposed methodology. A number of experiments were conducted, during which images with different informational meanings were compressed.

The results of the conducted research show that the created video compression method demonstrates a number of the best quality indicators and compression level relative to VP8. However, it is not symmetrical. The time required to perform decompression is lower than the cost to compress an image. Thus, this video compression method can be recommended for use in digital television, and in video databases.

ACKNOWLEDGEMENTS

The results presented are a part of a larger work being completed in accordance with a research grant titled Subsidy Agreement No. 075-11-2022-029 issued on April 8, 2022 on the subject: "Creation of a digital cluster for data storage and processing with consequent formation of a dynamic stream using AI technologies" with Ministry of Science and Higher Education of Russia.

REFERENCES

[1] V.Z. Kuklin, I.A. Alexandrov, A.A. Umyskov, A.K. Lampezhev, "Analysis of the Prospects for Developing Storage and Processing Complexes for Multiformat Media Data", Journal of Computer Science, Vol. 18, No. 12, pp. 1159-1169, 2022.

[2] R.R. Lopez, A.L. Sandoval Orozco, L.J. Garcia Villalba, "Compression Effects and Scene Details on the Source Camera Identification of Digital Videos", Expert Systems with Applications, Vol. 170, p. 114515, 2021.

[3] K. Mariappan, G.N.K. Suresh Babu, "Efficient Cloud Materials Storage and Cost Minimization for Video Using Block Compressive Sensing Based on Double Density Wavelet Transform", Materials Today, Vol. 81, pp. 415-422, 2023.

[4] S. Afzal, C.E. Rothenberg, V. Testoni, P. Kolan, I. Bouazizi, "Multipath MMT-Based Approach for Streaming High Quality Video over Multiple Wireless

Access Networks", Computer Networks, Vol. 185, p. 107638, 2021.

[5] V. Kuklin, I. Alexandrov, D. Polezhaev, A. Tatarkanov, "Prospects for Developing Digital Telecommunication Complexes for Storing and Analyzing Media Data", Bulletin of Electrical Engineering and Informatics, Vol. 12, No. 3, pp. 1536-1549, 2023.

[6] G.S. Lebedev, E.Y. Linskaya, A.A. Tatarkanov, A.K. Lampezhev, "Recent Solutions in the Field of Automated Monitoring and Quality Control of Telemedical Services", International Journal of Engineering Trends and Technology, Vol. 71, No. 1, pp. 62-78, 2023.

[7] A.A. Jeny, M.B. Islam, "Optimized Video Compression with Residual Split Attention and Swin-Block Artifact Contraction", Journal of Visual Communication and Image Representation, Vol. 90, p. 103737, 2023.

[8] P. Sharma, V. Neema, S.K. Vishvakarma, S.S. Chouhan, "MPEG/H256 Video Encoder with 6T/8T Hybrid Memory Architecture for High Quality Output at Lower Supply", Memories - Materials, Devices, Circuits and Systems, Vol. 4, p. 100028, 2023.

[9] J. Gutierrez Aguado, R. Pena Ortiz, M. Garcia Pineda, J.M. Claver, "Cloud-Based Elastic Architecture for Distributed Video Encoding: Evaluating H.265, VP9, and AV1", Journal of Network and Computer Applications, Vol. 171, p. 102782, 2020.

[10] N. Shyamala, S. Geetha, "Improved Integer Wavelet Transform (IIWT) Based Medical Image Compression Method", International Journal on Technical and Physical Problems of Engineering (IJTPE), Issue 51, Vol. 14, No. 2, pp. 339-346, June 2022.

[11] S. Chen, J. Zeng, "Processing of Volumetric Video", Immersive Video Technologies, Academic Press, pp. 445-468, 2022.

[12] K.S. Kumar, S.S. Kumar, N.M. Kumar, "Effcient Video Compression and Improving Quality of Video in Communication for Computer Encoding Applications", Computer Communications, Vol. 153, pp. 152-158, 2020.

[13] F. Chiariotti, "A Survey on 360-Degree Video: Coding, Quality of Experience and Streaming", Computer Communications, Vol. 177, pp. 133-155, 2021.

[14] H. Liu, H. Shen, L. Huang, M. Lu, T. Chen, Z. Ma, "Learned Video Compression via Joint Spatial-Temporal Correlation Exploration", The AAAI Conference on Artificial Intelligence, Vol. 34, No. 07, pp. 11580-11587, 2020.

[15] Y. Jeong, I. Kim, H. Kang, "A Practical Projection-Based Postprocessing of Block-Coded Images with Fast Convergence Rate", The IEEE Transactions on Circuits and Systems for Video Technology, Vol. 10, No. 4, pp. 617-623, 2000.

[16] A. Nosratinia, "Postprocessing of JPEG-2000 Images to Remove Compression Artifacts", IEEE Signal Processing Letters, Vol. 10, No. 10, pp. 296-299, 2003.

[17] H. Paek, R.C. Kim, S.U. Lee, "A DCT-Based Spatially Adaptive Post-Processing Technique to Reduce the Blocking Artifacts in Transform Coded Images", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 10, No. 1, pp. 36-41, 2000.

[18] G. Qiu, "MLP for Adaptive Postprocessing Block-Coded Images", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 10, No. 8, pp. 1450-1454, 2000.

[19] H. Song, S. Yu, C. Wang, L. Song, H. Xiong, "A New Deblocking Algorithm Based on Adjusted Contourlet Transform", The IEEE International Conference on Multimedia and Expo, Toronto, Canada, 09-12 July 2006.

[20] D. Sun, W.K. Cham, "Postprocessing of Low Bit-Rate Block DCT Coded Images Based on a Fields of Experts Prior", IEEE Transactions on Image Processing, Vol. 16, No. 11, pp. 2743-2751, 2007.

[21] Z. Liu, L. Meng, Y. Tan, J. Zhang, H. Zhang, "Image Compression Based on Octave Convolution and Semantic Segmentation", Knowledge-Based Systems, Vol. 228, p. 107254, 2021.

[22] L. Pearlstein, S. Maxwell, A. Aved, "Adaptive Prediction Resolution Video Coding for Reduced DRAM Bandwidth", Integration, Vol. 62, pp. 382-394, 2018.

[23] X.S. Zhan, J.W. Hu, J. Wu, H.C. Yan, "Performance Analysis Method for NCSs with Coding and Quantization Constraints", ISA Transactions, Vol. 107, pp. 287-293, 2020.

[24] D.K. Jang Bahadur Saini, et al., "Fractal Video Compression for IOT-Based Smart Cities Applications Using Motion Vector Estimation", Measurement: Sensors, Vol. 26, p. 100698, 2023.

[25] H.B. Kekre, P. Natu, T. Sarode, "Color Image Compression Using Vector Quantization and Hybrid Wavelet Transform", Procedia Computer Science, Vol. 89, pp. 778-784, 2016.

[26] D.A. Huffman, "A Method for the Construction of Minimum-Redundancy Codes," Resonance, Vol. 11, No. 2, pp. 91-99, 2006.

[27] A. Dhandayuthapani, J. Lawrence, "Plant Disease Recognition Using Optimized Image Segmentation Technique", International Journal on Technical and Physical Problems of Engineering (IJTPE), Issue 50, Vol. 14, No. 1, pp. 211-218, March 2022.

[28] K.S. Rao, A.V. Paramkusam, N.K. Darimireddy, A. Chehri, "Block Matching Algorithms for the Estimation of Motion in Image Sequences: Analysis", Procedia Computer Science, Vol. 192, pp. 2980-2989, 2021.

[29] G.R. Martin, M.K. Steliaros, R.A. Packwood, "Efficient Motion Estimation and Coding for Arbitrary-Shaped Video Objects", Journal of Visual Communication and Image Representation, Vol. 12, No. 1, pp. 66-83, 2001.

BIOGRAPHIES

Name: **Vladimir**
Middle Name: **Zhanovich**
Surname: **Kuklin**
Birthday: 18.02.1952
Birthplace: Kazan, Russia
Master: Mathematics, Department of General Mathematics, N.I. Lobachevsky Institute of Mathematics and Mechanics, Kazan State University, Kazan, Russia, 1971
The Last Scientific Position: Leading Researcher, Institute of Design and Technology Informatics, Russian Academic of Sciences, Moscow, Russia, Since 2022
Research Interests: Modeling and Automation of Technological Processes in Mechanical Engineering
Scientific Publications: 29 Papers, 2 Books, 4 Patents, 1 Thesis

Name: **Abas**
Middle Name: **Khasanovich**
Surname: **Lampezhev**
Birthday: 07.07.1997
Birthplace: Murmansk, Russia
Bachelor: Navigation and Ballistic Support for Use of Space Technology, Department of Dynamics and Flight Control of Rockets and Spacecraft, Faculty of Special Mechanical Engineering, Bauman Moscow State Technical University, Moscow, Russia, 2016
Master: Navigation and Ballistic Support for Use of Space Technology, Department of Dynamics and Flight Control of Rockets and Spacecraft, Faculty of Special Mechanical Engineering, Bauman Moscow State Technical University, Moscow, Russia, 2020
The Last Scientific Position: Researcher, Institute of Design and Technology Informatics, Russian Academy of Sciences, Moscow, Russia, Since 2020
Research Interests: Multi-Parameter Optimization and Mathematical Modeling, Multi-Parameter Optimization and Mathematical Modeling
Scientific Publications: 14 Papers, 2 Patents, 2 Theses

Name: **Leonid**
Middle Name: **Mikhailovich**
Surname: **Chervyakov**
Birthday: 30.06.1950
Birthplace: Kursk, Russia
Master: Mechanical Engineering, Department of Mechanical Engineering Technologies and Equipment, Faculty of Mechanics and Technology, Kursk Polytechnic Institute, Kursk, Russia, 1973
The Last Scientific Position: Leading Researcher, Institute for Design and Technological Informatics, Russian Academy of Sciences, Moscow, Russia, Since 2018
Research Interests: Modeling and Automation of Technological Processes in Mechanical Engineering
Scientific Publications: 82 Papers, 28 Books, 31 Patents, 8 Theses

Name: **Aslan**
Middle Name: **Adalbievich**
Surname: **Tatarkanov**
Birthday: 30.08.1991
Birth Place: Nalchik, Russia
Bachelor: Tracked and Wheeled Vehicles, Department of Multi-Purpose Tracked Vehicles and Mobile Robots, Faculty of Special Mechanical Engineering, Bauman Moscow State Technical University, Moscow, Russia, 2012
Master: Tracked and Wheeled Vehicles, Department of Multi-Purpose Tracked Vehicles and Mobile Robots, Faculty of Special Mechanical Engineering, Bauman Moscow State Technical University, Moscow, Russia, 2014
The Last Scientific Position: Researcher, Institute for Design and Technological Informatics, Russian Academy of Sciences, Moscow, Russia, Since 2019
Research Interests: Simulation Modeling in Multiphysics Modeling systems and CAD Systems, Mathematical Modeling
Scientific Publications: 26 Papers, 28 Patents, 6 Theses