УДК 621.391

IMPROVEMENT TO MOTION ESTIMATION FOR HIGH-EFFICIENCY VIDEO CODING

Jamal Kamil Kh. Abbas[1], Ali Ihsan Alanssari[1], Patlayenko M.[2], Pilyavskiy V.[2]

[1]Al Nisour University College,
Baghdad, Iraq.
jamal.k.eng@nuc.edu.iq, ali.ih.eng@nuc.edu.iq, nick_msa@ukr.net, v.@ukr.net
[2]O.S. Popov Odesa National Academy of Telecommunications,
1, Kuznechna St., Odesa, 65029, Ukraine.

УДОСКОНАЛЕННЯ ОЦІНКИ РУХУ У ВИСОКОЕФЕКТИВНОМУ ВІДЕО КОДУВАННІ

Джамал К. Кх. Аббас[1], Алі Іхсан Аланссарі[1], Патлаєнко М.О.[2], Пилявський В.В.[2]

[1]Університетський коледж Аль Нісур,
Багдад, Ірак,
jamal.k.eng@nuc.edu.iq, ali.ih.eng@nuc.edu.iq, nick_msa@ukr.net, v.pilyavskiy@ukr.net
[2]Одеська національна академія зв'язку ім. О.С. Попова
65029, м. Одеса, вул. Кузнечна, 1, Україна

УСОВЕРШЕНСТВОВАНИЕ ОЦЕНКИ ДВИЖЕНИЯ В ВЫСОКОЭФФЕКТИВНОМ ВИДЕОКОДИРОВАНИИ

Джамал К. Кх. Аббас[1], Али Ихсан Аланссари[1], Патлаенко М.О.[2], Пилявский В.В.[2]

[1]Университетский колледж Аль Нисур,
Багдад, Ирак,
jamal.k.eng@nuc.edu.iq, ali.ih.eng@nuc.edu.iq, nick_msa@ukr.net, v.pilyavskiy@ukr.net
[2]Одесская национальная академия связи им. А.С. Попова,
65029, г. Одесса, ул. Кузнечная, 1, Україна

Abstract. The paper deals with video coding based on the use of dependent sources, which is caused by the need to accelerate the coding process. The coding time is directly dependent on the computing power, which with the advent of new standards is not enough. Therefore, the article suggests the results of the study, which are designed to improve existing methods, as well as to continue the progress of scientific research in this area. The results consist of using an approach to compression based on coding methods of dependent sources, which can significantly reduce the complexity of processing on the side of the encoder. The results of the study were obtained using two coding algorithms for fixed and dynamic plots. For comparison, JPEG and, respectively, MPEG encoding standards were defined, as well as an improved algorithm that was considered by the authors in previous papers. It solves the problem of efficiently recovering intermediate frames on the decoder side. When evaluating the movement, search styles have a significant impact on the performance of the algorithm, both the search speed and the quality of the search. The search style should be designed considering the characteristics of the distribution of the motion vector probability (MVP) of the sequences of the real plot. The purpose of this study is to improve the coding efficiency of video sequences using the Diamond Search (DS) algorithm. The results show that DS is efficient and reliable, it can always give a higher search speed in different sequences than other algorithms.

Key words: High-Efficiency Video Coding (HEVC), Motion estimation, Motion compensation, Diamond Search, JPEG, MPEG.

Анотація. У статті розглядається кодування відео на основі використання залежних джерел, що викликано необхідністю прискорити процес кодування. Час кодування знаходиться в прямій залежності від обчислювальних потужностей, яких з появою нових стандартів не вистачає. Тому в статті пропонуються результати дослідження, які покликані удосконалити існуючі методи, а також продовжити прогрес наукових досліджень цієї галузі. Результати полягають у використанні підходу до стиснення на основі методів кодування залежних джерел, що дозволяє значно знизити складність обробки на стороні кодера. Результати дослідження отримані з використанням двох алгоритмів кодування для нерухомих і динамічних сюжетів. Для порівняння були визначені стандарти кодування JPEG і відповідно MPEG, а також вдосконалений алгоритм, який розглядався авторами в попередніх роботах. Він вирішує завдання ефективного відновлення проміжних кадрів на стороні декодера. При оцінці руху стилі пошуку істотно впливають на продуктивність алгоритму, як швидкість пошуку, так і якість пошуку. Стиль пошуку повинен бути розроблений з урахуванням характеристик розподілу ймовірності вектора руху (MVP) послідовностей реального сюжету. Метою даного дослідження є підвищення ефективності кодування відеопослідовностей з використанням алгоритму Diamond Search (DS). Результати показують, що DS ефективний і надійний, він завжди може дати більш високу швидкість пошуку в різних послідовностях, ніж інші алгоритми.

Ключові слова: високоефективне кодування відео (HEVC), оцінка руху, компенсація руху, Diamond Search, JPEG, MPEG.

Аннотация. В статье рассматривается кодирование видео на основе использования зависимых источников, что вызвано необходимостью ускорить процесс кодирования. Время кодирования находится в прямой зависимости от вычислительных мощностей, которых с появлением новых стандартов не хватает. Поэтому в статье предлагаются результаты исследования, которые призваны усовершенствовать существующие методы, а также продолжить прогресс научных исследований данной области. Результаты заключаются в использовании подхода к сжатию на основе методов кодирования зависимых источников, что позволяет значительно снизить сложность обработки на стороне кодера. Результаты исследования получены с использованием двух алгоритмов кодирования для неподвижных и динамических сюжетов. Для сравнения были определены стандарты кодирования JPEG и соответственно MPEG, а также усовершенствованный алгоритм, который рассматривался авторами в предыдущих работах. Он решает задачу эффективного восстановления промежуточных кадров на стороне декодера. При оценке движения стили поиска оказывают существенное влияние на производительность алгоритма, как скорость поиска, так и качество поиска. Стиль поиска должен быть разработан с учетом характеристик распределения вероятности вектора движения (MVP) последовательностей реального сюжета. Целью данного исследования является повышение эффективности кодирования видеопоследовательностей с использованием алгоритма Diamond Search (DS). Результаты показывают, что DS эффективен и надежен, он всегда может дать более высокую скорость поиска в разных последовательностях, чем другие алгоритмы.

Ключевые слова: Высокоэффективное кодирование видео (HEVC), оценка движения, компенсация движения, Diamond Search, JPEG, MPEG.

The transfer of video from mobile devices via wireless communication networks has become increasingly common. This is due to the significant development of wireless data transmission, and the development of mobile devices: wireless sensors, wireless video surveillance cameras, user mobile devices, etc. In such systems, the transmitter is usually characterized by low computational capabilities and limited battery capabilities. In this case, significant limitations are imposed on both the processor power and the amount of memory. The existing video compression technology, primarily the approaches described in ITU-T H. 26x and ISO / IEC MPEG standards, do not take into account the specificity of the source of information in similar systems. Therefore, the development of new methods for compressing visual data that do not require large computational capabilities on the transmitter side is an important and urgent task. The document discusses distributed video coding, an approach to compression based on the coding methods of dependent sources, which can significantly reduce the complexity of processing on the transmitter side. It solves the problem of efficiently recovering intermediate frames on the decoder side. Various aspects of coding dependent sources (or distributed coding sources) are presented in the works of well-known domestic and foreign authors (S.I. Gelfand, V.D. Kolesnik, B.D. Kudryasov, M.S. Pinker, G.Sh. Poltirev, A. Weiner, D. Wolfe, D. Slepian).

However, until recently, the practical implementation of these ideas was not realized. Only in the late 1990s were tasks applied in which the use of distributed coding could give advantages over the existing approaches of the time. One of the significant advantages of using distributed coding in this system is that the procedure for estimating and eliminating temporal redundancy, based on the prediction of intermediate frames, can be performed at the receiver, which significantly reduces the complexity of processing at the transmitter. The inter-prediction procedure in such systems is called the generation of additional information. Recovery of intermediate frames is carried out in the decoder, and an increase in compression is possible only through the modification of the receiver. Within this class of application problems, codec concepts have been developed based on the principles of distributed coding sources. In recent years, much attention has been paid to the expansion and modification of these concepts in view of the emerging new application problems, such as the efficient encoding of multidimensional and multi-resource production. Despite this, a number of open questions remain at the level of basic principles. These include the features of the input data for the intercoder prediction procedure on the decoder side, as well as the estimation of the intercoder prediction error parameters.

It should also be noted that in most modern papers, heuristic approaches are described for solving practical problems arising from the distributed coding of visual data. At the same time, theoretical data models are not fully studied, which makes it impossible to effectively develop new and improve existing compression algorithms using this approach.

High-Efficiency Video Coding. Intraframe processing is performed using the JPEG compression algorithm and an improvement in the encoding method. Increasing the quantization coefficient allows you to reduce the speed of the video stream, but at the same time, a coarser quantization of the values of the transformation coefficient, with the result that most of them become zero. The signal decoded images with large values of the compression ratio, containing a certain kind of distortion, are noticeable in the subjective assessment of the image of the frame, also reduce the objective assessment.

It should also be noted that the entropy encoding of quantized coefficients was performed by the length of a series of zigzag scans, and then by archiving the resulting series to the .rar format with the maximum compression ratio. The following are the results for various values of the compression ratio, while the evaluation of the results of the algorithms was performed according to the metrics:

- $PSNR_{AM}$ - peak signal-to-noise ratio for advanced method, main image;
- $PSNR_{AM+CI}$ - peak signal-to-noise ratio for advanced method, a main image with addition corrects information;
- $PSNR_{JPEG}$ - peak signal-to-noise ratio for the comparative method (JPEG);
- compressed file size in bytes;

Comparative versions of the results of these algorithms with the same source data and the requirements for the resulting image are presented in Table 1, 2. The improved method is present (1, 2).

Table 1 – The results of the algorithms developed a method (main image - MImg) and JPEG

	94 frame video sequence “foreman”						
$PSNR_{MImg}$, dB	34,4196	32,1348	30,8439	29,4944	28,4559	27,9485	27,4832
$PSNR_{JPEG}$, dB	34,4349	32,1389	30,8410	29,4953	28,4915	27,9429	27,4814
improved	9 685	6 213	4 590	3 497	2 763	2 462	2 253
JPEG	10 377e	6 545	4 953	3 730	2 897	2 590	2 336

Table 2 – The results of the algorithms of the developed method (the main image with the addition of corrective information) and JPEG

	94 frame video sequence "foreman"						
$PSNR_{MImg}$, dB	34,5	32,5	31,8	31,2	30,8	30,7	30,6
$PSNR_{JPEG}$, dB	34,5	32,5	31,8	31,2	30,8	30,7	30,6
improved	10 116	7 613	6 842	6 530	6 311	6 212	6 322
JPEG	10 446	7 076	6 041	5 329	4 939	4 762	4 710

The above data allows you to build graphs of image quality, expressed as the PSNR value in dB, on the amount of compressed data in bytes. These graphics are shown in Fig. 1, 2.

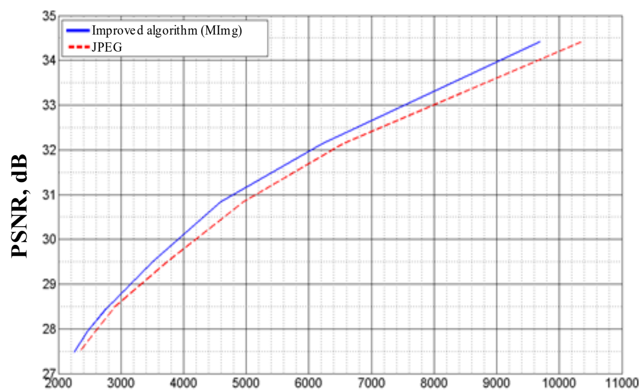

Figure 1 – Graphs of image quality as a function of the size of the received data when encoding 94 frames of a foreman video sequence using the developed method (main image) and the JPEG algorithm

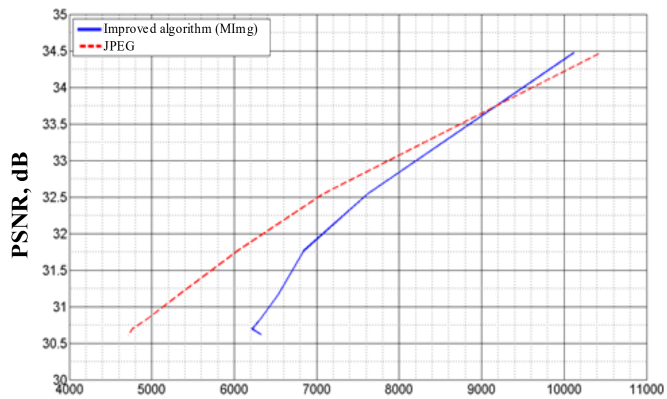

Figure 2 – Graphs of image quality as a function of the size of the received data when encoding 94 frames of the foreman video sequence using the developed method (main image with correction information added) and the JPEG algorithm

It can be noted that when compressing images of high-definition frames using the improved method using only the main image, the volume occupied and required for data transmission is an order of magnitude smaller than when encoding the same images using JPEG algorithms. When compressed with minimal losses in the quality of the decoded image, the difference between the results of the algorithms in the occupied volume of compressed data is 24.9%. When compressing with a high degree of loss in the quality of the decoded image, the difference in the amount of compressed data is 31%. For standard definition formats, the improved method using only the main image also showed its efficiency in comparison with the JPEG method, but the coding efficiency becomes less pronounced with decreasing format size values. So for the video sequence "foreman", which has a format of 1920x1080 pixels, the gain in comparison with JPEG will be from 3.6% to 6.7%.

Consequently, the improved method, using the main image, allows you to create data for transmission that serves to decode the transmitted image frame, the volume of which will be 31% less compared to the volume occupied by the data obtained by the JPEG method in the video information signal of high definition 1280x800 pixels.

In addition to the main image data, the data of the correction information is also transmitted, which allows achieving greater efficiency compared to the JPEG method only when encoding with low degrees of loss in the quality of the decoded image. Increasing the degree of losses to high values leads to the fact that the JPEG coding efficiency is already higher. So, at low losses, the gain in compression by the developed method is up to 25%, while at high values of loss, the loss is 22%.

Results of simulation of interceded compression with the introduction of advanced coding method. The improved method provides a small time to search for motion vectors and a small amount of data during intercoder processing. The simulation was carried out using the basic compression algorithm, MPEG family standards, with a full search of the search area and using an improved coding method, which also takes a full search of the search area. Evaluation of the effectiveness of the new method was carried out by comparing the operation time of the listed algorithms and the number of obtained motion vectors. In order to fully evaluate the performance of the algorithms and increase the accuracy of the data obtained, eleven consecutive frames were processed. For the Foreman video sequence, these are frames 95 - 104. The results of the work are presented in Table 3.

Table 3 – Results of the work of the algorithms of the developed method and MPEG for the video sequence "Foreman"

Frame number	95	96	97	97	99	100	101	102	103	104
Time, sec. (improved algorithm)	15.23	14.60	13.18	10.91	12.39	12.43	11.00	13.26	14.02	14.00
Time, sec. (MPEG)	16	16	16	16	16	16	16	16	16	16
The number of motion vectors (Improved algorithm)	1473	1254	871	575	686	752	555	629	908	1197
The number of motion vectors (MPEG)	1553	1385	1009	626	822	894	686	785	1057	1337

The results of the experiment, given above, allow you to create graphs for a visual reflection of the effectiveness of the developed method. Graphs are presented in Fig. 3, 4.

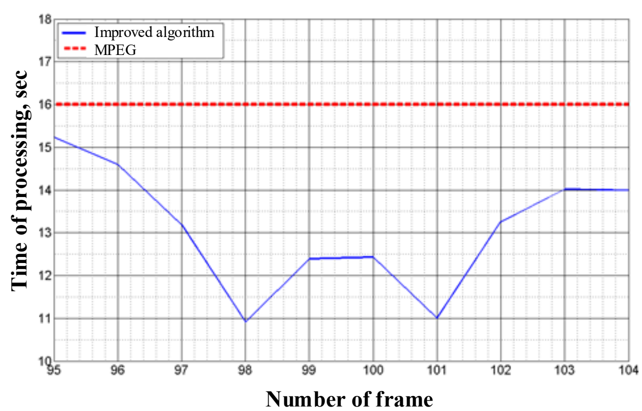

Figure 3 – Time frames processing frames of the video sequence "Foreman '" for the developed method and the MPEG method

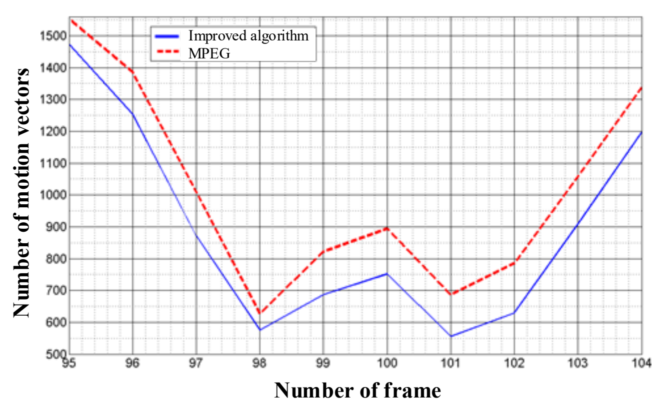

Figure 4 – Graphs of the number of motion vectors required to restore the processed frames of the Foreman video sequence for the developed method and the MPEG method

To demonstrate the results of the work below, Fig. 5 and 6 show the processed and decoded frames of test video sequences, in which the blocks for which the motion vector search was performed were marked with squares, and the beginning of the white dot marked with blue lines.

Analyzing the data obtained and the data of the above graphs, we can note the following. The time required to process a frame with interframe coding is significantly reduced when using the advanced method. The dynamic nature of the change of time spent turns out to be from the frame, indicates the presence in the improved method of adaptability to the space-time structure of each frame processed.

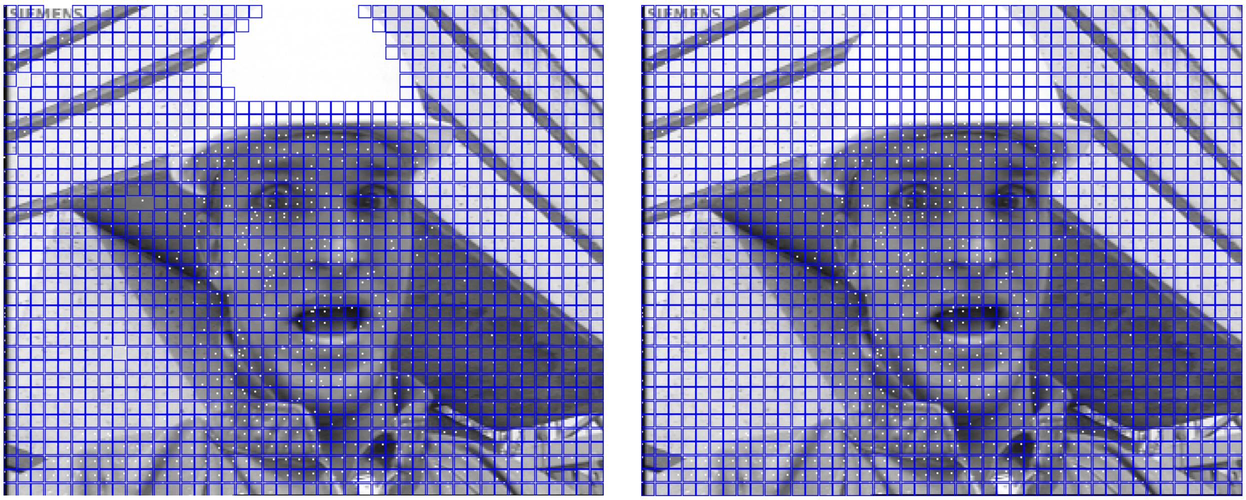

Figure 5 – Image 95 of the frame of the "Foreman" video sequence with highlighted blocks for which the motion vector search will be performed. Left - improved method, right - MPEG

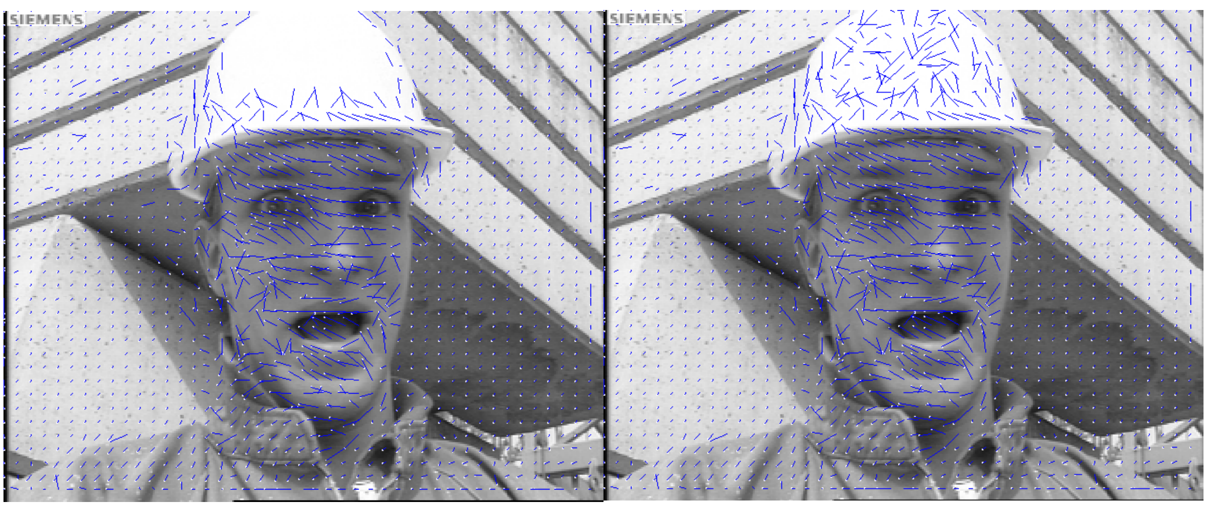

Figure 6 – Frame 95 of the video frame sequence "Foreman" with highlighted motion vectors

The result of processing is the motion vector, the number of which is an order of magnitude lower when using the improved method. Reducing the number of motion vectors, which must necessarily be transmitted to the receiving side to ensure the decoding procedure, leads to a reduction in the total volume of the video information signal. Thus, the obtained results and their analysis is a confirmation of the effectiveness of the improved method and its ability to make significant improvements to the existing video information signal processing algorithms.

Experiments and Results. The proposed method can estimate the motion between consecutive frames with the shortest time. Spacious tests have been executed to assess the performance of the proposed method. Evolutions have been performed about to 2D-HEVC codec version HM-KTA [3]. Each experiment has been performed in accordance with the JCT-VC Common Test Conditions (CTC), which dub encoder configuration for valuation of 2D codec execution used by the International Organization for Standardization (ISO). Main parameters of the encoders have been collected in Table 4.

Table 4 – Main configuration parameters of the encoders utilized in the tests.

Parameter	**Value**
Profile	Main
GOP size	4
Intra period	24
FrameSkip	0
SEIDecodedPictureHash	1
MaxCUWidth	256
MaxCUHeight	256
MaxPartitionDepth	6
QuadtreeTULog2MaxSize	6
LoopFilterTcOffset_div2	-2
InternalBitDepth	8
SearchRange	64

The coding efficiency is measured at four operating points (QP = 27; 22; 37; 32), calculating time reduction and the average rate reduction in comparison to HEVC with full search algorithm (FS) following the method described in [4] and the results can be summarized in Table 5. The PSNR values and the corresponding total bitrates for all test sequences for HEVC with diamond search algorithm and HEVC with full search algorithm and presented in Table 6.

Table 5 – BD-rate calculated by using Bjontegaard rates for PSNR introduced by DS algorithm against the FS algorithm for sequences.

Sequences	Random Access Configuration			
	BD-Rate			Encoder time
	Y	U	V	
BasketballDrill_832x480	0.40%	-0.59%	0.31%	20.43%
BQSquare_416x240	-0.20%	0.41%	-0.20%	34.92%
BlowingBubbles_416x240	0.78%	-0.11%	0.33%	22.72%

Sequences	Low Delay Configuration			
	BD-Rate			Encoder time
	Y	U	V	
BasketballDrill_832x480	0.21%	0.84%	1.22%	42.88%
BQSquare_416x240	-0.19%	-0.48%	-0.65%	24.51%
BlowingBubbles_416x240	0.56%	1.19%	2.06%	20.50%

Table 6 – Experimental results for all test sequences for DS algorithm and FS algorithm.

Sequences	QP	Random Access Configuration									
		Diamond Search					Full Search				
		Bitrate [kbit/s]	YPSNR [dB]	UPSNR [dB]	VPSNR [dB]	Time [sec]	Bitrate [kbit/s]	YPSNR [dB]	UPSNR [dB]	VPSNR [dB]	Time [sec]
Basketball Drill	22	3196.4	41.1	43.6	44.4	1616.8	3184.9	41.1	43.6	44.4	7102.0
	27	1542.4	38.1	41.4	41.9	1508.9	1539.3	38.1	41.4	41.9	6966.6
	32	760.8	35.1	39.6	39.9	1197.0	763.00	35.2	39.6	39.9	6370.9
	37	408.0	32.6	38.4	38.5	1129.5	407.9	32.6	38.4	38.6	6101.8
BQSquare	22	1748.7	39.5	43.9	44.9	324.8	1757.7	39.6	43.9	44.8	673.5
	27	864.9	36.1	41.7	42.4	184.2	865.5	36.1	41.7	42.4	618.3
	32	461.9	32.8	39.9	40.7	164.7	460.6	32.8	39.9	40.7	541.3
	37	263.4	29.9	38.7	39.4	158.6	263.3	29.9	38.7	39.5	507.8
Blowing Bubbles	22	1664.3	38.8	41.5	43.6	410.8	1663.2	38.8	41.5	43.5	1633.6
	27	771.2	35.6	39.0	41.0	281.8	769.4	35.7	38.9	40.9	1291.8
	32	372.5	32.6	37.1	39.0	277.2	371.9	32.6	37.2	39.1	1321.3
	37	183.6	29.9	35.6	37.6	258.6	183.2	29.9	35.7	37.6	1127.9

Sequences	QP	Low Delay Configuration									
		Diamond Search					Full Search				
		Bitrate [kbit/s]	YPSNR [dB]	UPSNR [dB]	VPSNR [dB]	Time [sec]	Bitrate [kbit/s]	YPSNR [dB]	UPSNR [dB]	VPSNR [dB]	Time [sec]
Basketball Drill	22	3497.6	41.1	43.4	44.2	7932.0	3491.9	41.1	43.4	44.3	11475.6
	27	1660.3	37.9	41.1	41.8	7881.5	1663.3	37.9	41.2	41.8	11094.2
	32	807.1	35.0	39.3	39.7	1449.4	807.9	35.0	39.3	39.8	8095.3
	37	425.9	32.4	37.9	38.2	1075.2	424.2	32.5	37.9	38.3	7992.8
BQSquare	22	2184.6	39.2	43.5	44.3	294.8	2182.9	39.3	43.5	44.3	1058.3
	27	1005.8	35.4	41.3	42.1	213.7	1006.9	35.4	41.3	42.1	926.0
	32	499.3	32.1	39.5	40.3	188.0	499.1	32.1	39.5	40.3	886.8
	37	270.1	29.1	38.5	39.1	169.7	271.3	29.1	38.5	39.1	655.4
Blowing Bubbles	22	1973.5	38.8	40.9	42.9	686.7	1971.2	38.7	40.9	42.9	2327.1
	27	864.6	35.2	38.4	40.5	452.9	861.4	35.2	38.4	40.6	2186.0
	32	399.1	32.1	36.6	38.6	343.1	393.9	32.0	36.6	38.6	2059.8

The proposed algorithm allows reducing the amount of information by 24.9% under compression with minimal loss in the quality of the decoded image. When compressing with a high degree of loss in the quality of the decoded image, the difference in the amount of compressed data is 31%. The winnings may differ due to the complexity of the image being encoded, the degree of quality required, but the advantage will be preserved compared to existing methods. It is not possible to bring the data for different types of images in this work.

In the paper, a comparison between HEVC with DS algorithm and HEVC with FS algorithm, that allows reducing time, has been presented. The performed experiments showed that the HEVC with DS algorithm led to average time reduction to over 26% in random access configuration and 29% in low delay configuration for video sequences recorded with the HEVC with DS algorithm comparing to the HEVC with FS algorithm.

It should be noted that the presented proposals for the improvement of compression algorithms in the complex of the general compression algorithm will give a predominant compression ratio in comparison with the existing methods.

REFERENCE:

1. Oleg Gofaizen, Olena Osharovska, Mikola Patlayenko, Volodymyr Pyliavskyi "IEEE The Second International Conference on Informations and Telecommunication Technologies and Radio Electronics." *Adaptive decomposition of TV image*. Odessa, 2017.
2. Patlaenko M.O. " Analysis of image compression methods on the basis of wavelet transformation in applications of television production " *International scientific and technical journal VOTTP* 2016.
3. Joint Collaborative Team on Video Coding, “HMKTA reference software 16.6”.
4. G. Bjontegaard, “Calculation of Average PSNR Differences between RD-curves,” ITU-T SG16, Doc. VCEG-M33, Austin, USA, Apr. 2001.

ЛІТЕРАТУРА:

1. IEEE The Second International Conference on Informations and Telecommunication Technologies and Radio Electronics (УкрМіКо-2017) [Конфереція] / Гофайзен О.В. Ошаровська О.В., Патлаєнко М.О., Пилявський В.В. // Adaptive decomposition of TV image, 11-15 вересня. – Одеса, 2017.
2. Аналіз методів стиснення зображень на основі вейвлет-перетворення у застосуваннях телевізійного виробництва [Стаття] / auth. М.О. Патлаєнко // Міжнародний науково-технічний журнал ВОТТП, 2016. – № 2. – 55 с.
3. HMKTA reference software 16.6 / auth. Coding Joint Collaborative Team on Video.
4. Calculation of Average PSNR Differences between RD-curves [Документ] / Bjontegaard G. // ITU-T SG16. - Austin : VCEG-M33, 2001

DOI 10.33243/2518-7139-2019-1-1-112-120