VIDEO COMPRESSION METHOD ON THE BASIS OF DISCRETE WAVELET TRANSFORM FOR APPLICATION IN VIDEO INFORMATION SYSTEMS WITH NON-STANDARD PARAMETERS

DOI: 10.36724/2072-8735-2020-14-3-47-53

Natalya S. Valitskaya,

Moscow Technical University of Communications
and Informatics, Moscow, Russia,
nvalitskaya@mail.ru

Igor V. Vlasyuk,

Moscow Technical University of Communications
and Informatics, Moscow, Russia,
ru3dlp@yandex.ru

Aleksei M. Potashnikov,

Moscow Technical University of Communications
and Informatics, Moscow, Russia,
nickmikh@gmail.com

Keywords: video information systems, synchronization, parallelization, fault tolerance, video compression, discrete wavelet transform, filtering, additional calculations

The features of image compression standards based on a discrete wavelet transform are considered. Particular attention is paid to the issues of data synchronization and fault tolerance of video information systems, in which, to obtain high spatial resolution, the image is divided into fragments with their subsequent separate decoding and playback. A modification of the codec for low-latency video data transmission systems based on a discrete wavelet transform is presented taking into account the option of changing the filtering direction of the samples from vertical and horizontal to diagonal developed at the TiZV department of MTUCI, allowing to slightly reduce computational complexity by reducing the number of processed extrapolated elements at the edges of the image (or tile), simplify parallel processing of data, and also introduce an additional level of fault tolerance these transmission systems directly to the video codec by generating a pair of independent mutually complementary data streams. A variant of optimizing the parameters of the quantizer of the DVT subband signals is shown, taking into account the spatial frequency response of the human visual system. The possibility and scenarios of using the presented video codec in video information systems with non-standard parameters (as a rule, with ultra-high resolution and often complex shape of the display area) are analyzed.

Для цитирования:

Валицкая Н.С., Власюк И.В., Поташников А.М. Метод видеокompрессии на базе дискретного вейвлет-преобразования для применения в видеотелекоммуникационных системах с нестандартными параметрами // Т-Comm: Телекоммуникации и транспорт. 2020. Том 14. №3. С. 47-53.

For citation:

Valitskaya N.S., Vlasyuk I.V., Potashnikov A.M. (2020) Video compression method on the basis of discrete wavelet transform for application in video information systems with non-standard parameters. *T-Comm*, vol. 14, no.3, pp. 47-53. (in Russian)

Introduction

Currently, there is an active development of video information systems. In case it is necessary use the images of such high resolution that there is no interface for their transmission or there is not enough resource on one video server, it is also possible to use several video servers that are connected using an Ethernet switch [1]. Due to the instability of network technology, there is a problem of signal synchronization when displayed on a video wall.

An analysis of Internet sources, including publications hosted on electronic abstracting platforms such as eLibrary and Scopus, showed that in video information systems, synchronization is currently implemented in the process of data packaging. However, this is not enough: loss of synchronization can occur at the stage of decoding transport streams, since the time required to complete this process is uncontrollable.

Thus, it is necessary to ensure synchronization at the level of a video codec capable of forming tiles of the complete image, because this is the first and, accordingly, the last operation in the process of transmitting media data [2].

In order to implement such a synchronization method, compression standards based on the wavelet transform can be used, since this will allow displaying a single fragment of the low-frequency component on all screen modules reproduced by a single source (video server), which reduces the number of blocks in the decoding process that can potentially occur thread out of sync. At the same time, high-frequency components, although they can be reproduced by different sources, do not significantly affect the integrity of the image perception even in the case of some desynchronization, especially in non-optimal conditions for image control characteristic of video information systems. Therefore, this method provides synchronization built into the stream. However, when using wavelet transforms, a problem arises in the form of additional calculations.

JPEG2000 codec algorithm

An example of a wavelet-based compression standard is JPEG2000. The block diagram of the JPEG-2000 encoder is shown in Figure 1 [3].

Pre-processing includes dividing each component of the frame into areas (tiles), shifting in brightness to equalize the dynamic range, and converting color space. The wavelet transform is a convolution of the original signal with the wavelet function.

A two-dimensional discrete wavelet transform (DWT) is applied to the image in accordance with formulas (1-2) [4]:

$$W_{\varphi}(j_0, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \varphi_{j_0, m, n}(x, y), \quad (1)$$

$$W_{\psi}^i(j, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \psi_{j, m, n}^i(x, y), \quad (2)$$

where: $f(x, y)$ – function of source tile; $W_{\varphi}(j_0, m, n)$ – coefficients that determine the approximation of a function on a j_0 scale; $i = \{H, V, D\}$ – index for identifying "directed" wavelets; $W_{\psi}^i(j, m, n)$ – coefficients that determine horizontal, vertical and diagonal details when $i = H, V$ and D respectively; $M \times N$ – image tile resolution; j_0 – arbitrary initial scale; $j = 0, 1, \dots, J-1$ – scales; $\varphi_{j, m, n}(x, y)$ – two-dimensional scaling function; $\psi_{j, m, n}^i(x, y)$ – two-dimensional wavelet functions.

Fig. 1. JPEG2000 image compression diagram

Two-dimensional discrete wavelet transform is performed by filtering each row and column of a pre-processed image tile using a low-pass filter and high-pass filter, and it does not matter if these operations are applied to rows or columns in the first place.

Since the number of samples doubles as a result of this process, the samples of the output signal of each filter are subsampled in steps of 2; therefore, the sampling rate remains constant.

An example of a two-dimensional wavelet transform is shown in Figure 2.

Fig. 2. DWT (First level of decomposition)

The lines of the original image are subjected to low-frequency and high-frequency filtering, as a result, the number of samples increases by 2 times. Then, subsampling of the row counts occurs in step 2, which means the removal of every second column of the LF and HF components obtained in the previous step of the wavelet transform. Further, similar actions are performed with respect to the columns: first, the number of column samples doubles as a result of filtering, then it also decreases by 2 times due to subsampling, i.e. removal of every second line.

At this stage, the two-dimensional discrete wavelet transform is completed. The result is the decomposition of the image tile into 4 components (LL, LH, HL, HH) [4]. The value of these four subbands is conveniently summarized in table 1.

This decomposition is the first level of decomposition. As shown in Figure 3, the two-dimensional discrete wavelet transform can be reapplied to the low-frequency component of the image, which will mean an increase in the level of decomposition by one.

Table 1
Names and notation of subbands

Name	Notation		Axis filter	
			X	Y
Low frequency (LF)	$W_{\varphi}(j, m, n)$	LL	LF	LF
Horizontal (HF)	$W_{\psi}^H(j, m, n)$	HL		LF
Vertical (VF)	$W_{\psi}^V(j, m, n)$	LH		HF
Diagonal (HF)	$W_{\psi}^D(j, m, n)$	HH	LF	HF

The coefficients obtained after a two-dimensional discrete wavelet transform are quantized in accordance with formula (4):

$$q_b(i, j) = \text{sign}(y_b(i, j)) \left\lfloor \frac{|y_b(i, j)|}{\Delta_b} \right\rfloor, \quad (4)$$

where: $y_b(i, j)$ – discrete wavelet transform coefficient; $\text{sign}(y_b(i, j))$ – DWT coefficient sign; Δ_b – quantization step.

The quantizer dead zone (“deadband”) is located near zero and is equal to twice the quantization step.

The quantization resolution in each frequency subband is different: the subbands corresponding to the lower sensitivity of the visual system are quantized more roughly: in fact, they adapt according to the average level of sensitivity within each subband.

Before encoding, the tile subbands are preliminarily divided into blocks so that the code blocks corresponding to one subband have the same size. The PCRD-opt Algorithm (Post-Compression Rate-Distortion optimization) algorithm allows to achieve the desired data transfer rate with minimal distortion. To highlight the region of high quality – the region of interest (ROI), a single-bit image mask is used, transformed to cover all the coefficients corresponding to the ROI [6].

Fig. 3. DWT (1-3 levels of decomposition)

The required number of levels of decomposition is determined in accordance with the spatial frequency response (SFR) of the human visual system, shown in Figure 4 [5]. Without taking into account the anisotropy of the SFR of the visual system, it can also be represented in a two-dimensional version by rotation about the vertical axis.

Fig. 4. SFR on human visual system

In the figure, $K_{\text{mean}1}$, $K_{\text{mean}2}$, ... $K_{\text{mean}6}$ are the average sensitivity levels within each subband at 6 iterations of DWT (decomposition levels). At a frequency of 1 per./deg. the sensitivity of the visual system is quite low, and further division of the spectrum does not make sense. We pose the problem: when dividing the LF subband by 2 times at each level of decomposition, determine the number of the last iteration J_{max} , at which $f_{\text{min}} = 1$ per. / deg. In this case J_{max} can be calculated by the following formula:

$$J_{\text{max}0} = \left\lceil \log_2 \frac{f_{\text{max}}}{f_{\text{min}}} \right\rceil = \left\lceil \log_2 \frac{60}{1} \right\rceil = 6 \quad (3)$$

Additional calculations

Additional calculations occur when filtering pixels located at the edges of the image. Since the resolution of the frame is greater than the resolution of one screen module, the image is divided into tiles, and the number of extreme pixels increases. To process

one such pixel, additional $\frac{N-1}{2}$ counts are required, where N is

the filter order. With an increase in the level of decomposition, the area of pixels outside the image, which is necessary for processing extreme pixels, expands by 2 times. This is explained by the fact that the wavelet transform is accompanied by a decrease in the resolution of the low-frequency component by 2 times in rows and columns compared to the previous level. Accordingly,

when level J is decomposed, the number of samples $\frac{N-1}{2}$ is

equal to the Dop samples of the original image, which essentially determines half the filter aperture according to formula (5):

$$Dop = \frac{N-1}{2} \cdot 2^{J-1} \quad (5)$$

Fig. 5. Expanding the filter aperture while increasing the level of decomposition

The expansion of the sampling area outside the image necessary for processing the extreme pixels, with an increase in the level of decomposition, is shown in Figure 5.

For example, consider the case of applying an 11th-order filter. At the first level of decomposition, $R1 = 5$ pixels, at the second, after subsampling the samples in steps of 2, the distance $R1' = 10$ pixels of the original image, and at the third, respectively, $R1'' = 20$ pixels.

The greater the filter aperture, the greater the level of decomposition and the order of the filter.

A method to reduce computational complexity

To reduce computational complexity, we will use the method developed by Sedov M.O. In the article [7], the author describes the principle of adaptive discrete wavelet transform, which allows you to change the filtering direction depending on the image structure. The idea is the following: the image is divided into blocks, and if the block contains vertical and / or horizontal borders, classical filtering by rows and columns is applied to it; if the image section has diagonal borders, the filtering of parallel to main diagonal elements is used, and then - perpendicular to the elements of the main diagonal; diagonal filtering can be done using the classic filters, if you rotate the image by 45 degrees.

We take advantage of the latter solution. Filtering at an angle of 45 degrees can be used to reduce computational complexity when processing pixels located at the edges of the image. It is more convenient to carry out this type of filtering by rotating the tile by 45 degrees and applying classical filtering by rows and columns. However, when the image is rotated, the orthogonality of the structure in the form of a shift by half a pixel is violated, and there is no way to process such a tile in simple ways.

The problem can be solved as follows: the tile is divided into two "chess" structures, one of which contains only odd pixels in odd lines and even in even lines (structure A), and the other contains the remaining pixels of the tile (structure B), as shown in Fig. 6, and for each such structure separately after a 45-degree rotation, a two-dimensional fiberboard is applied with the required number of decomposition levels in accordance with the diagram in Fig. 1.

Fig. 6. a) Part of the tile containing structure A;
b) The part of the tile containing structure B

Pixels of the source image $P_{x,y}$ of size X, Y , where $x \in [0, X)$, $y \in [0, Y)$ – the coordinates of the pixels of the

original image, belong to structure A, if the condition $\frac{x+y}{2} \in N$ is met, otherwise – to structure B.

When the image is divided into such structures, the distance between the samples increases by a factor of $\sqrt{2}$. The larger the distance between the samples, the lower the filter cutoff frequency. In this case, a smaller number of DWT iterations will be required to achieve the selected frequency $f_{\min}$. This is equivalent to working with a tile that has a resolution in each of the original (horizontal and vertical) directions that is $\sqrt{2}$ less. The filter becomes lower frequency, and the number of additional horizontal and vertical samples remains the same, or equivalently, at the same cutoff frequency lower than the filter aperture, as shown in Figure 7.

Fig. 7. Filter aperture during classical filtration and filtration at an angle of 45 degrees at one $f_{\min}$.

Let's show the dependence of the gain achieved by diagonal filtering, depending on the resolution of the screen module of the video wall. First, we determine the values of the areas of additional pixels during classical filtering D1 and when filtering at an angle of 45 degrees D2 as a percentage of the resolution of the image in one direction.

In formulas (6-7), double distances appear when counting D1 and D2, because $X1$ and $X2$ characterize the half-aperture of the filter aperture, i.e. the area of additional calculations when the filter window is located is symmetrical relative to the edge of the tile, and when considering the gain in relation to the image resolution on one side, it must be taken into account that the tile has two edges in the horizontal and vertical directions.

$$D1 = \frac{X1 * 2}{I} * 100, \% \quad (6)$$

$$D2 = \frac{X2 * 2}{I} * 100, \% \quad (7)$$

The gain is determined in accordance with the formula (8):

$$V = D1 - D2, \% \quad (8)$$

Graphs of the dependence of the gain on the resolution of the screen module with the filter order $N = 11$ for different levels of decomposition are shown in Figure 8.

Fig. 8. Graphs of the dependence of the gain on the resolution of the screen module at different levels of decomposition in the case of filtering at an angle of 45 degrees

Thus, if we consider the percentage of the filter aperture to the image resolution (on one side), then the gain is greater, the lower the resolution of the screen module and the greater the level of decomposition.

Parallel processing

As mentioned above, for filtering at an angle of 45 degrees, the tile is divided into two fragments with a checkerboard pattern of samples, each of which is independently applied by a discrete wavelet transform. Parallel processing is a very common practice, in which the same task can be performed by a large number of devices of lower processing power.

There is an alternative solution of Dvorkovich V.P. [8]. The method consists in dividing the signal spectrum into three bands: low-frequency, mid-frequency and high-frequency. Filtering rows and columns in this way gives a spectrum decomposition

into 9 regions, then it is proposed to apply the usual two-band wavelet transform to the LF and MF components and, depending on the depth of such a conversion, get 11, 13 or 16 components. The structure of such transformations and the frequency response of three-band wavelet filters is shown in Figure 9. To compare the spectrum decomposition methods in the classical way, when filtering at an angle of 45 degrees and when using a three-band wavelet for a one-dimensional version of the conversion, it is convenient to summarize in table 2.

Table 2

Spectrum decomposition methods in the classical way, when filtering at an angle of 45 degrees and when using a three-way wavelet

Classical filtration									
Tile									
1			L						HF
2		L			HF				
45 degrees filtration									
Tile									
1			L						HF
2		L			HF				
Three-band wavelet									
Tile									
1			L						HF
1a		LF		HF		LF	MF	HF	

The coding method described above with the image converted into a pair of chess structures allows us to additionally parallelize the processing process into twice as many threads as compared to the classical processing scheme in Fig. 1.

The method considered above [8] also allows you to parallelize the conversion process to a better degree than the classical processing scheme. Moreover, the transformations performed with the approach described in the present work and in the method [1] can be combined, adding their advantages.

LLF	MLF	HLF	LLF		MLLF		HLLF
					MLHF		HLHF
LMF	MMF	HMF	LMLF	LMHF	MMLLF		HMF
					MMHLF	MMHHF	
LHF		HHF	LHLF	LHHF	MHF		HHF

a)

c)

Fig. 9. a) One iteration of a three-way wavelet transform; b) Additional conversion; c) Frequency response of three-way wavelet filters

Fault tolerance

Splitting a tile into 2 parts and their parallel processing gives increased fault tolerance. In the case of failures at the network and software level, part of the image may be lost, for example, individual transforms. With the loss of low-frequency components, this can lead to extremely noticeable artifacts in the image.

When parallel processing of two chess structures formed from pixels of the common part of the frame with their independent further entropy encoding and transmission, the amount of transmitted data, of course, increases, but not twice. At the same time, a tile can be restored only by one of the components, using one of the known methods of local restoration of image elements, for example, used in demosaicing, for example [9].

This will lead to almost imperceptible distortions in the image due to a $\sqrt{2}$ decrease in resolution in the worst case, vertically or horizontally, which is quite acceptable for emergency operation of the system.

Conclusion

In the course of the study of the features of the use of DWT-based image compression standards in video information systems, along with advantages, for example, the convenience of synchronizing streams due to the possibility of outputting low-frequency transformants from a single source with a short delay time, scalability and the absence of the need for post-processing of images due to the specifics of distortions, shortcomings were also identified, such as often insufficient code parallelism, instability to errors in the data stream for low-frequency transformations, necessity to perform extrapolation processing of images near its boundaries.

The proposed scheme for dividing an image into two “chess” structures applying wavelet transforms to them in the direction $\pm \pi/4$ to the vertical allows eliminating or reducing the influence of several of the above disadvantages: to form two independent data streams in a natural way, each of which allows restoration of

the image with high subjective quality by simple means, reduce the size of extrapolated areas at the edges of the image, and improve the parallelization of the video compression algorithm by half.

The research results can be applied to create video information systems with increased requirements for synchronization accuracy and fault tolerance.

References

1. Gorodilov M.A., Dolgovesov B.S., Khramtsov I.D., Radostev A.H. (2015). Features of multi-screen display of distributed multimedia data. *Bulletin of NSU. Series: Physics*. vol. 10, no. 2, pp. 91-98. (in Russian)
2. Valitskaya N.S., Vlasjuk I.V. (2019). Synchronization methods for video information systems. *Telecommunications and Information Technologies*, no. 2, pp. 51-57. (in Russian)
3. Vlasjuk I.V., Romanova E.P., Sidorova A.I. (2010). Coding of high-quality areas in the JPEG 2000 image compression standard. *T-Comm*, no. 9, pp. 53-55. (in Russian)
4. Gonzalez R., Woods R. *Digital image processing*. Translated from English by P.A. Chochia. Moscow: Technosphere, 2005. 1072 p. (in Russian)
5. Fairchild Mark D. *Color appearance models*. Chichester: John Wiley & Sons, Ltd., 2005. 385 p.
6. Acharya T., Tsai P. *JPEG2000 standard for image compression: concepts, algorithms and VLSI architectures*. Hoboken: John Wiley & Sons, Inc., 2005. 274 p.
7. Sedov M.O. (2012). Adaptive discrete wavelet transform. *T-Comm*, no. 9, pp. 127-128. (in Russian)
8. Dvorkovich V.P., Gilmanshin A.V. (2008). A new approach to the use of wavelet filters in image processing. *Digital signal processing*, no. 1, pp. 37-42. (in Russian)
9. Bezrukov V.N., Vlasjuk I.V., Romanov S.G. Patent No. 2557261 Russian Federation, H04N 5/335, H04N 11/24 2013138922/07; declared 08/20/2013; publ. 07/20/2015, Bull. No. 20. Method and apparatus for generating image signals in standard and high definition digital television systems. (in Russian)

МЕТОД ВИДЕОКОМПРЕССИИ НА БАЗЕ ДИСКРЕТНОГО ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ ДЛЯ ПРИМЕНЕНИЯ В ВИДЕОИНФОРМАЦИОННЫХ СИСТЕМАХ С НЕСТАНДАРТНЫМИ ПАРАМЕТРАМИ

Валицкая Наталья Сергеевна, МТУСИ, Москва, Россия, nvalitskaya@mail.ru

Власюк Игорь Викторович, МТУСИ, Москва, Россия, ru3dip@yandex.ru

Поташников Алексей Михайлович, МТУСИ, Москва, Россия, nickmikh@gmail.com

Аннотация

Рассмотрены особенности стандартов сжатия изображений на базе дискретного вейвлет-преобразования. Особое внимание уделено вопросам синхронизации данных и отказоустойчивости видеоинформационных систем, в которых для получения высокой пространственной разрешающей способности применяется разделение изображения на фрагменты с их последующим раздельным декодированием и воспроизведением. Представлена модификация кодека для систем передачи видеоданных с малой задержкой на базе дискретного вейвлет-преобразования с учетом разработанного на кафедре ТиЗВ варианта с изменением направлений фильтрации отсчетов с вертикального и горизонтального на диагональные, позволяющая несколько снизить вычислительную сложность за счет уменьшения количества обрабатываемых экстраполированных элементов на краях изображения (или тайла), упростить параллельную обработку данных, а также внедрить дополнительный уровень отказоустойчивости системы передачи непосредственно в видеокодек за счет генерации пары независимых взаимодополняющих друг друга потоков данных. Показан вариант оптимизации параметров квантователя сигналов субполос ДВП с учетом пространственной частотно-контрастной характеристики зрительной системы человека. Проанализирована возможность и сценарии использования представленного видеокодека в видеоинформационных системах с нестандартными параметрами (как правило, со сверхвысоким разрешением и часто сложной формой области отображения).

Ключевые слова: видеоинформационные системы, синхронизация, распараллеливание, отказоустойчивость, видеокомпрессия, дискретное вейвлет-преобразование, фильтрация, дополнительные вычисления.

Литература

1. Городилов М.А., Долговесов Б.С., Храмцов И.Д., Радостев А.Х. Особенности построения систем для полиэкранного отображения распределенных мультимедийных данных // Вестник НГУ. Серия: Физика, 2015. Т. 10, № 2. С. 91-98.
2. Валицкая Н.С., Власюк И.В. Методы синхронизации потоков в видеоинформационных системах // Телекоммуникации и информационные технологии, 2019. №2. С. 51-57.
3. Власюк И.В., Романова Е.П., Сидорова А.И. Кодирование областей повышенного качества в стандарте сжатия изображений JPEG 2000 // T-Comm: Телекоммуникации и транспорт, 2010. № 9. С. 53-55.
4. Гонсалес Р., Вудс Р. Цифровая обработка изображений / Пер. с англ. П.А.Чочиа. М.: Техносфера, 2005. 1072 с.
5. Fairchild Mark D. Color appearance models. Chichester: John Wiley & Sons, Ltd., 2005. 385 p.
6. Acharya T., Tsai P. JPEG2000 standard for image compression: concepts, algorithms and VLSI architectures. Hoboken: John Wiley & Sons, Inc., 2005. 274 p.
7. Седов М.О. Адаптивное дискретное вейвлет-преобразование // T-Comm: Телекоммуникации и транспорт, 2012. №9. С. 127-128.
8. Дворкович В.П., Гильманшин А.В. Новый подход к использованию вейвлет-фильтров при обработке изображений // Цифровая обработка сигналов, 2008. №1. С. 37-42.
9. Безруков В.Н., Власюк И.В., Романов С.Г. Патент №2557261 Российская Федерация, H04N 5/335, H04N 11/24 2013138922/07; заявл. 20.08.2013; опубл. 20.07.2015, Бюл. № 20. Способ и устройство формирования сигналов изображений в системах цифрового телевидения стандартной и высокой четкости.

Информация об авторах:

Валицкая Наталья Сергеевна, МТУСИ, студент, Москва, Россия

Власюк Игорь Викторович, МТУСИ, к.т.н., доцент кафедры ТиЗВ, Москва, Россия

Поташников Алексей Михайлович, МТУСИ, мнс отдела ЦТиВ, Москва, Россия