Image Compression as a Variation Calculus Task

Dmitry Kalistratov
Department of Radioelectronics, Tula State University, Tula, Russian Federation
Email: Kalistratow@list.ru

Abstract—This paper is devoted to the problem of increasing the compression ratio and quality of decoded images with high-frequency spectrum. It also addresses the problem of regulating priorities between these parameters. This is very important for communication systems that operate by widescreen video images, because the transmission speed, energy costs and image quality depend on compression ratio. The indicated problems are solved by the methods of wavelet transform in combination with methods of variation calculus. The novelty of the proposed technical solution lies in the fact that the compression of the high-frequency component of the Haar wavelet transform is considered as a variation calculus task. Choice in favor of the compression ratio or the image quality is made by setting the values of the weighting coefficients in the objective function.

Index Terms—Digital video image, compression, code volume, image quality, priorities, functional, variations

I. Introduction

Now video compression has gained wide popularity in communication systems [1]-[4]. Compression is the reduction of the image volume in computer memory [5]-[7]. Devices or programs for image compression are called video codecs [8], [9]. A typical video codec consists of two complementary parts: the encoder and the decoder. The encoder compresses the image and sends its codes to the decoder. Decoder restores the received image.

In real-time communication system the binary stream of codes is usually transmitted via radio signal. Various methods of digital modulation are used here, but the principle remains the same - the higher the compression ratio, the less time and energy will be spent on image transmission.

Thus, the compression efficiency is obvious: compressed images require less memory for storage, they are transferred faster and with less energy. Due to this, the communication system with a built-in video codec is effective from both technical and economic points of view.

However, a large compression ratio reduces the quality of the decoded image. As a result of quantization, rounding, and other compression operations, some part of the information on the decoding side becomes lost. The image becomes blurred, small details disappear.

In other words, the two most important indicators of the effectiveness of a video codec conflict with each other. If greater the compression ratio, then lower the image quality and vice versa. In this view the video codec is a «compromise» between the factor of compression and the quality of the decoded image.

We also should not forget that the minimum required frame rating for the «human eye» is at least 15 images per second. Only in this case, viewing will be nice and comfortable.

II. Compression and Quality

A good example of the «conflict» between the compression ratio and the image quality is a video message about transport movement condition. The Fig. 1 shows the real image, obtained using a quadrocopter. The height of the video capture was about 100 meters. The dotted line shows a rectangular area, for which the modeling results are presented below.

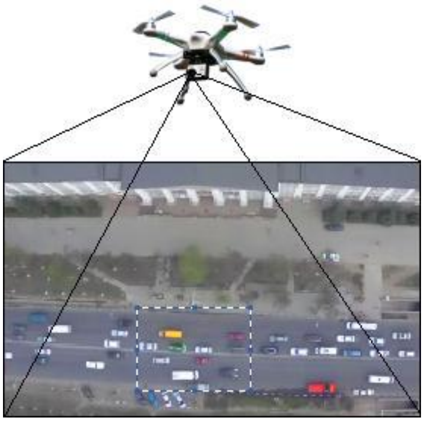

Fig. 1. Widescreen image of traffic flows.

If we just want to show on TV the panorama of «traffic jams» on a city street, the compression ratio is more important. But, if we want to demonstrate in detail the fragment of the video with breach of road movement, then the image quality comes to the fore. In the second case the boundaries of cars projections should be clear. The compression algorithm should not damage them much.

So, increasing the compression ratio leads to lower image quality on the decoding side. The structure and visibility of the image become destroyed. This is especially true for widescreen images with high-frequency spectrum.

III. Haar Wevelet Transform

It is known [10] that the Haar wavelet transform in pairs calculates the half-sums and half-differences for the

Manuscript received April 15, 2019; revised December 6, 2019.
doi:10.12720/jcm.15.1.95-100

image pixel color signals. Let the conversion be performed line by line. Select one image color matrix and one row. Then one half-sum and one half-difference corresponds to each pair of pixels in the row (line). After the transformation, we obtain two matrices: low-frequency *L*, shown in Fig. 2, and high-frequency *H*, shown in Fig. 3. The horizontal dimension on the *x*-axis for each of these matrices is two times smaller than the dimension of the original image fragment.

The Haar wavelet transform can be complemented by level quantization. The quantization coefficients may be different for *L* and *H*. In the encoder matrixes *L* and *H* are divided into quantization coefficients after the direct transform. In the decoder, they are multiplied back by the same coefficients before the inverse transform. Quantization and dequantization operations occur with rounding. The number of discrete signal levels and the bit scale are reduced. As a result, quantization gives an additional gain in the compression ratio, but degrades the quality of the decoded image on the receiving side.

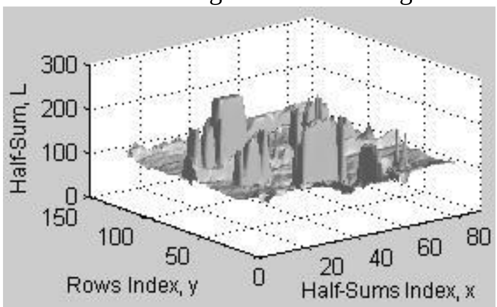

Fig. 2. Low-frequency component of the Haar wavelet transform.

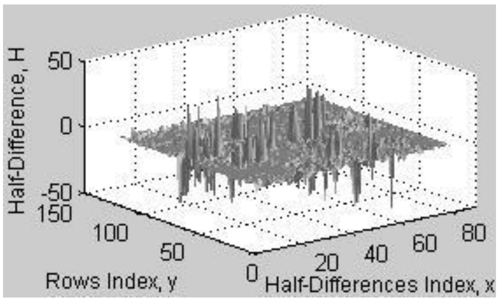

Fig. 3. High-frequency component of the Haar wavelet transform.

As a whole, compression in this type of wavelet transform is achieved mainly due to the high-frequency component, because it has small absolute values and requires a smaller bit scale for storing and transmitting information. But, if the image has a high-frequency spectrum (as in our example), then the effect of compression becomes minimal. So, the Haar wavelet transform is weak in relation to the compression of images with a high-frequency spectrum. This is the main disadvantage of the Haar wavelet transform, which the proposed solution aims to fix.

IV. Methodology

In the proposed method, the component *L* is processed in a standard way. The changes concern only processing of component *H*. Therefore, the block-scheme of methodology in Fig. 4 contains additional optimization and filtering blocks only for the *H*. Encoder blocks are located in the left column of the circuit, and decoder blocks in the right. Note that the decoder performs the conversion in the reverse order. The compression task is solved as a variation calculus problem [11, 12] with the distribution of priorities between the compression factor and image quality.

Like the standard approach, the method allows quantization and dequantization. However, it is recommended to set a large quantization coefficient only for the *L* wavelet transform component. In the particular case, quantization coefficients can also be set equal to one.

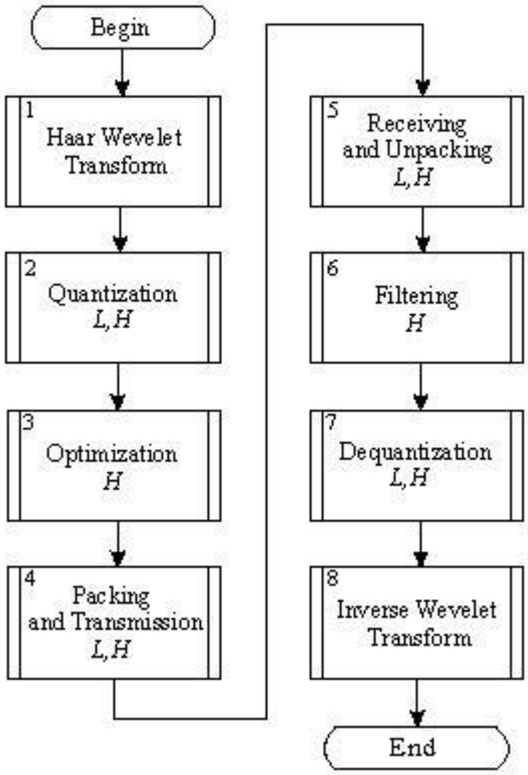

Fig. 4. The block-scheme of methodology.

In the proposed method, the decoder does not transmit the *H* values themselves, but the values of its first line by line derivatives on *x*. Moreover, before transfer, *H* is additionally smoothed by adding to him the second artificial component. Thus, the first line (row) derivatives of the *H* become smaller in absolute values.

Consider the encoding process in detail. Select one color and denote the signal of one row of *H* as *w*. In the process of optimization, a second artificial component *q* (with lower frequency spectrum than itself *w*) is added to the *w*. On the encoding side, this component *q* in some measure smoothes *w* (that increases the compression ratio), but on the decoding side *q* is eliminated by a high-pass filter (that preserves the image quality).

In other words, small values of the smoothed first derivative s' of total component *s* (calculated as a sum

$w+q$) are convenient for compression, since they are represented by small numbers and require a small bit scale. Contradictory requirement is that on the decoding side the output of the filter f should be similar to the original w in encoder. Priorities between these goals are regulated by weighting coefficients c_1, c_2.

From a mathematical point of view, this is a task of variation calculus with two quality criteria and one limitation function (for decoder filter). The functional can be written as:

$$J=\int_{l_1}^{l_2}\left(c_1\left((w+q)'\right)^2+c_2(w-f)^2+\lambda\left(ak(w'+q')-af'-f\right)\right)dx, \quad (1)$$
$$J(q,f)\in\mathbb{R},\ x\in\mathbb{R},\ w(x),q(x),f(x),\lambda(x)\in\mathbb{R},$$

where J is target functional, x is argument (associated with pixels along a row), l_1, l_2 are integration limits (associated with the beginning and end of the row), c_1, c_2 are weight coefficients, w is initial signal, q is additional artificial signal, f is the filter output signal (w equivalent in decoder), λ is the Lagrange multiplier function, a is the filter coefficient, k is the gain coefficient. In the theoretical formulation of the variation task, all dependent parameters are assumed to be continuous and belong to the set of real numbers. In practice, the solution of the task is translated into a discrete form.

The first summand of the functional with c_1 represents the compression ratio criterion. The second summand with c_2 corresponds to the image quality. If c_1 is greater than c_2, then priority is given to the compression ratio. If c_2 is greater than c_1, then image quality becomes more important. In fact, the larger the coefficient c_1, the stronger the additional q distorts the original w in favor of compression and vice versa.

The third summand is the equivalent of a high-frequency RC-filter, the output voltage of which is removed from the resistor. The coefficient a sets the cutoff frequency of the filter. The coefficient k is used to amplify the filter signal (it also affects the quantization in the encoder).

The solution of the variation task is found by the Euler equations. These equations are known. Thus, finding two variations of the functional by the Euler equations and adding one limitation function, we obtain a system of three ordinary differential equations:

$$\begin{cases} 2c_1 w''+2c_1 q''+ak\lambda'=0, \\ -a\lambda'+2c_2 w-2c_2 f+\lambda=0, \\ akw'+akq'-af'-f=0. \end{cases} \quad (2)$$

In practice, this system is solved by numerical methods at the program level. In computer processing, the mathematical model is discretized, and differential relations are used in conjunction with derivatives. The algorithm runs through the current row and then moves on to the next.

Together with the s' codes, the first column of H is also transmitted to the decoder to set the same boundary conditions as in the encoder. Decoder restores s, and then passes it through a high-frequency filter, the equation of which is written in the functional with the Lagrange multiplier. The additive q is eliminated, it remains only signal w.

The method is intended mainly for images with a high-frequency spectrum, since the focus is on the H-component of the Haar wavelet transform. The more high-frequency harmonics contained in the input image, the more pronounced the effectiveness of the method in the regulation and improvement of video codec quality indicators.

Important nuances of the proposed method are: the necessary and sufficient conditions of the functional extremum, the stability of the solution, and the software implementation of the high-pass filter in digital form. Let's discuss them in more detail.

The necessary and sufficient conditions of functional extremum - mathematical aspects, which peculiar not only to this task, but also to all variation tasks in general. The functional is treated as a function of functions. The necessary conditions for an extremum are formed from the requirement that the first variation of the functional must be equal to zero. The first variation in optimization task is the semantic analogue of the first differential in ordinary algebraic analysis.

To provide the necessary conditions, the method of indefinite Lagrange multipliers is widely used. This method takes into account all variables in the functional construction (including Lagrange multipliers near limitation functions). Necessary conditions are usually formed as systems of differential equations that include the Euler equations and limitation equations (here we have 2 Euler equations and 1 limitation equation).

The necessary conditions must be supplemented by sufficient conditions of the extremum. And if, in the case of necessary conditions, the first variation of functional was used, then the sufficient conditions are based on the analysis of its second variation. The second variation is a semantic analog of the second differential. There are several ways to determine the sufficiency of the functional extremum.

The most common approach is to analyze the sufficiency “in terms of variations”. For example, in our task, to ensure the minimum of the functional, the second variation of the functional must be positive. In our functional this is achieved thanks to the using of quadratic degree in the first two members. With such a construction, the value of the functional cannot be negative and it tends to a zero minimum. In the analysis of the sufficient condition we should not take into account the limiting function on the right side of the functional, since the member at the Lagrange multiplier will always be equal to zero.

Another important aspect is the stability of the solution. If the solution is unstable, the divergent processes distort the parameters of the mathematical model in amplitude to infinity in theory. Such distortion is unacceptable. Stability depends on the combination of the differential equations coefficients. One of the most well-known criteria for checking the stability of the solution is the Hurwitz criterion. If the solution can not be made stable, then it must at least be reduced to the boundary of stability (as in the considered model).

Now let's move on to the digital filter description. We can implement the decoder filter in digital form by various ways. But the improvement in compression ratio and image quality will in some measure depend on the specific discrete realization of the filter. Therefore, it is important to discuss in detail the features of the software implementation of the digital high-pass filter in decoder.

The organization of the digital filter is convenient to explain with its electronic counterpart in Fig. 5. The equivalent electrical scheme of our digital filter is a serial connection *RC*-circuit. First of all, we emphasize, that the equivalent scheme of the analog filter is given here only for explanation. An analog filter is a real circuit, while a digital filter is a program.

Input voltage U_1 is a complete signal, that corresponds to the sum *s*, U_2 is its low-frequency part, and output voltage U_3 is its high-frequency part, that corresponds to the digital filter output *f*. The constant *RC* in analog filter is determined by the value of the *a* coefficient in the digital filter. At the end we must extract U_3.

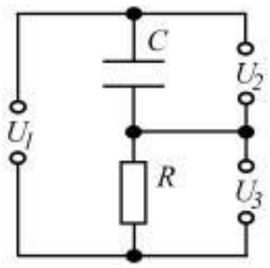

Fig. 5. The equivalent analog scheme for the digital filter.

The voltages of the analog filter in the equivalent scheme depend on the continuous time, while the corresponding parameters of our discrete filter are functions of the discrete coordinate (since they depend on the serial number of the wavelet transform elements). So, filters have similar working principle, but different nature of arguments. In addition, the definition area of the digital filter is discrete.

At a high sampling rate the signals of the analog and digital filters will be the same. However, if the sampling rate for the digital filter definition area is not enough (counts are taken with long intervals), then the destruction of the digital signals begins to appear. This is due to the Nyquist theorem, which states that the sampling frequency of the harmonic signal must be at least 2 times higher than the frequency of the signal itself.

If we program the digital filter so that it immediately takes the voltage from the resistor in the equivalent scheme, then in the digital version we will get a strong negative effect of the Nyquist theorem. High-frequency harmonics of *s* in our task will be damaged, decoding accuracy will decrease, and image quality will be lost.

It is much more profitable to implement a digital filter in two steps. In the first stage, we extract the low-frequency component by taking the voltage U_2 from the capacitor in the equivalent circuit, and then in the second stage, subtract this component from the full original signal U_1. The end result is the desired high frequency filter component U_3. In such filter implementation, thanks to the intermediate step, we will get a practically unaffected high-frequency filter component, because the negative impact of the Nyquist theorem will have little effect at low frequencies. This kind of filtering approach was used to simulate the method.

V. Results

To additionally explain the principle of the proposed method, let's at first present the graphical simulation results. Fig. 6 shows how the additional component smoothes the *H*. The gradual character of this graph is confirmed by the small values of its first line by line derivative in Fig. 7. The graph of Fig. 8 refers to the decoder. It represents the distortion of decoded *H*, which tell us about image quality on the receiving side. In this example, we can see that the image quality will be quite good. Capital letters on the graphs corresponds to lowercase in a mathematical model by nature, but represent a complete matrixes composed of all rows.

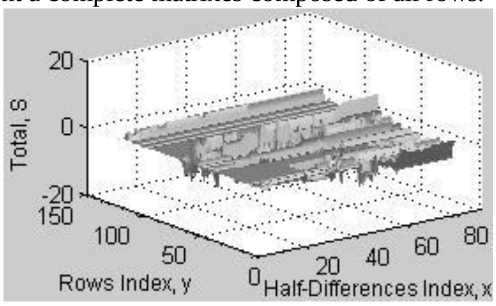

Fig. 6. Total signal *S* as a sum of *W* and *Q* after smoothing.

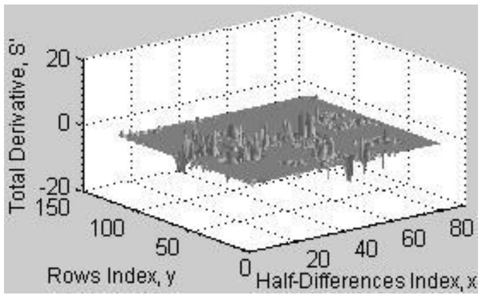

Fig. 7. First line by lint derivative of the total signal *S*.

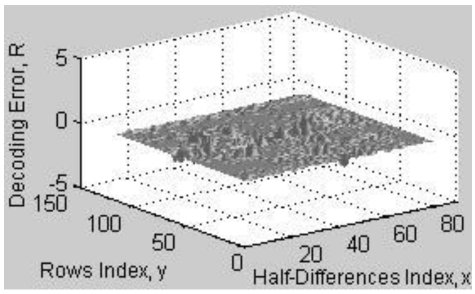

Fig. 8. The difference of *H* in the encoder and in the decoder.

The set of zero values on the graph of the first derivative S' is the reason of the compression ratio increasing in the proposed method. If the optimization block will be supplemented by algorithm of Run Length Encoding (RLE), then the compression ratio will increase even more.

Table I shows the values of the simulation parameters at which the proposed method was compared in effectiveness with the base method (without optimization). Through k_1 and k_2 are indicated the quantization coefficients of the L and H prior to optimization. The designations of other parameters have already been disclosed above.

The sum of weight coefficients c_1 and c_2 here is taken equal to one. However, this requirement is not necessarily. Mutual ratio of the values of these coefficients is more critical, because it shows how many times one of them is more important than the other.

TABLE I: SIMULATION PARAMETERS

Parameters	Value
Size of Image Fragment	160×120 pixels
Coefficient k_1	4
Coefficient k_2	4
Coefficient k	8
Coefficient a	0.5
Coefficient c_1	0.3
Coefficient c_2	0.7

Two criteria of video codec efficiency were used for a comparative analysis of the proposed and base methods: the degree of compression and the quality of the decoded image. The compression efficiency was determined using the Compression Factor (*CF*), which was calculated by the formula:

$$CF=\frac{N_1}{N_2} \tag{3}$$

where *CF* is compression factor, N_1,N_2 are the total quantities of bits, required for storage of uncompressed and compressed image correspondently.

To assess the quality of the decoded image, the Signal to Noise Ratio (*SNR*) was used. The signal was the original encoded image. The noise equivalent was represented as a discrepancy between the original image in the encoder and the same recovered image in the decoder (as if the image was distorted not by compression, but by noise). Used formula that gives the result in decibels:

$$SNR=20\lg\left(\frac{U_1}{U_2}\right) \tag{4}$$

where *SNR* is signal to noise ratio, U_1,U_2 are root mean squares for original signal and noise equivalent correspondently.

Table II presents the values of the efficiency parameters characteristic of the proposed method. Both parameters were calculated programmatically, taking into account the three color matrixes of the image.

TABLE II: EFFICIENCY PARAMETERS

Parameters	Value
Compression Ratio *CF*	6.91
Signal to Noise Ratio *SNR*	38.8 dB

For comparison, the *CF* of the base method with equal *SNR* is 6.04. That is, the proposed method with equal quality of the decoded image improves the compression ratio by 14.4%. This effect is slightly enhanced with increasing coefficient *k* and weakens at his small values.

A small disadvantage of the method is the increased time spent on image compression process. This is because numerical solutions of systems of differential equations are used. However, in general, the possibility of using the method in real time is preserved.

Moreover, increasing the compression ratio is not the only use of the proposed method. The gain in compression ratio can be converted into an image quality gain. Then, with equalized compression ratios, the proposed method gives improvement of about 6 dB in *SNR*. This is significantly for compressed video.

VI. CONCLUSIONS

Thus, we get an improved method of digital video compression and the ability to control the priorities between the compression ratio and the quality of the decoded image. Optimization was applied to the high-frequency component of the image spectrum.

Due to the gain in compression ratio with saving the image quality, the proposed method is recommended for wide-format video images in video communication systems. It can be local connection systems with a video channel, as well as television and mass communication systems. Video codec based on the method is equally well suited for both low-frequency images and high-frequency images.

In addition, thanks to the interpretation of compression as a task of the optimization and variation calculus, we get a convenient way to adjust the efficiency parameters of a video codec, including in real time. This is very useful when in the process of transmitting a television report it is necessary to change priorities in favor of the compression ratio or image quality with switching the optical system from panoramic landscapes to important details in the foreground.

Regarding further research, we note that the smoothing of the high-frequency component of the Haar wavelet transform is not the only direction for optimization. Similarly, the opposite task of optimal changing for the low-frequency component of the Haar wavelet transform can be posed. Correspondently, to do this, we need a low-pass filter on the decoding side.

Moreover, the design of the functional can include not only low-pass and high-pass filters of the first order, but also band-pass filters of the second order. The mathematical models will be different, but the general concept remains the same - to change the signal in the encoder so as to provide the best compression factor or the best image quality on the decoding side.

For further development, it is also very interesting to explore in detail the method not separately, but in combination with other known compression methods, for example, with a group of Run Length Encoding methods. Theoretically, such a combination can significantly enhance the compression effect of the proposed method.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

The research and its results analysis were realized by Dmitry Kalistratov, including the statement of the variation calculus task, mathematical modeling and preparation of the final version of the article.

REFERENCES

[1] L. Hanzo, P. Cherriman, J. Streit, *Video Compression and Communications*, Wiley & Sons, 2007.

[2] C. Zhu and Y. Li, *Advanced Video Communications over Wireless Networks*, CRC Press, 2017.

[3] J. R. W. Heath, *Introduction to Wireless Digital Communication*, Prentice Hall, 2017.

[4] I. B. Djordjevic, *Advanced Optical and Wireless Communications Systems*, Springer, 2018.

[5] M. Parker, S. Dhanani, *Digital Video Processing for Engineers. A Foundation for Embedded Systems Design*, Elsevier, 2013.

[6] A. Beach and A. Owen, *Video Compression Handbook*, Peachpit Press, 2018.

[7] D. S. Kalistratov, E. I. Minakov, and V. A. Barhotkin, *Video Coding. Optimization of Static Video Image Compression Methods*, Tula: Publishing House of TSU, 2016.

[8] I. E. G. Richardson, *H.264 and MPEG-4 Video Compression*, Wiley & Sons, 2003.

[9] Jae-Beom Lee, Hari Calva, *The VC-1 and H.264 Video Compression Standards for Broadband Video Services*, New York: Springer, 2008.

[10] D. Sundararajan, *Discrete Wavelet Transform. A Signal Processing Approach*, Wiley & Sons, 2015.

[11] F. Rindler, *Calculus of Variations*; Springer, 2018.

[12] V. I. Vanko, O. V. Ermoshina, and G. N. Kuvirkin, *Variation Calculus and Optimal Control*, Moscow: Publishing House of MSTU by N.E. Bauman, 2006.

Dmitry Kalistratov was born in Tula, Russian Federation, in 1986. He received the B.S. degree and the M.S. degree from the Tula State University (TSU), Tula, in 2009 and 2011 correspondently, both in robotics. He also received the Ph.D. degree from the Vladimir State University (VSU), Vladimir, in 2015, in radio engineering. Now he teaches at the Department of radio electronics in TSU. His research interests include wireless communication systems, remote connection, digital signal processing, and mathematical modeling.