Adaptive Decomposition OF TV Images

Oleg Gofaizen, Olena Osharovska, Mikola Patlayenko, Volodymyr Pyliavskyi
Dept. Radio and Television Broadcasting
O.S. Popov Odessa National Academy of Telecommunications
Odessa, Ukraine
oleg.gofaizen@gmail.com, osharovskaya@gmail.com, nick_msa@ukr.net, v. pilyavskiy@ukr.net

***Abstract*— The paper presents the results of modeling adaptive decomposition for ultra-high definition television, which provides image compression without subjective deterioration of quality. The decomposition of the images was carried out in the spectral region for the cases of wavelet analysis and Fourier transformation, depending on the signal-to-noise ratio at the image edges. Quantitative estimates of the root-mean-square error in the reconstruction of the image and the dependence of the signal-to-noise ratio on the edges on the compression coefficient are given. The number of non-zero components to be transmitted for the luminance and color difference components at two and three decomposition levels is calculated.**

Keywords*— *Wavelet transform; image compression; compression algorithms; decomposition; signal/noise ratio

I. Introduction

Perfection of light-signal converters based on arrays and charge-coupled devices (CCD) makes it possible to obtain color images of high and ultra-high definition. For viewers, such images create the effect of presence and cause positive emotions, but also raise the requirements for subjective quality. From the point of view of transmission via communication channels, ultra-high definition television sets tasks for further development, algorithms and devices for efficient compression of huge data streams. Existing standards for the compression of TV images MPEG-4 AVC, HEVC, Motion-JPEG-2000 [1, 2] offer a combination of several compression algorithms, which are based on integral transforms and transition to the spectral domain with the subsequent quantization of the conversion coefficients. As an integral transformation is widely used Discrete Cosine Transform (DCT) and, more rarely, wavelet transform.

The aim of the work is to try to increase the image compression ratio, taking into account the correlation between low-frequency and high-frequency parts of the image spectrum and the possibility of predicting individual spectral components at each level of decomposition.

II. Decomposition of the Image Spectrum

A. Selecting a Test Image

As the source was chosen an image of ultra-high definition with dimensions of 3840 elements horizontally and 2160 lines. These dimensions correspond to the image parameters of image definition 4K. The maximum frequency $F_{max} = 124.416$ MHz, the frequency of sampling brightness component $F_s = 248.832$ MHz, $k = 16/9$ aspect ratio. Figure 1 shows the original image.

Fig. 1. Example of an original image

From the point of view of preservation quality assessment, this image contains horizontal, vertical borders (castle), diagonal borders (sculpture in the center of the lake and reflection of the castle in the water), and there are also fuzzy borders on clouds in the sky and reflections of the forest in the water. A large number of small parts and the presence of a relatively large area with a very smooth color change makes it possible to use this image as a test.

B. Ierarchical Decomposition

Any TV image in the spectral region can be considered as a combination of spatial frequencies along the horizontal, vertical and diagonal, with the latter component for two-dimensional images, sufficiently correlated with the other two. Then it is possible to analyze either individual spectral components, as it is done in the DCT, or whole parts of the spectrum of the "sub band", which is typical for the wavelet. In the latter case, such an expansion into separate parts of the spectrum is called a hierarchical decomposition. We attempted to investigate the hierarchical decomposition by the Fourier analysis method, and the number of decomposition levels is determined depending on the energy parameters of the corresponding portions of the spectrum.

For each sub band, the number of elements can be calculated as follows: 3840/2 × 2160/2 = 1920 × 1080 (first decomposition); 1920/2 × 1080/2 = 960 × 540 (second decomposition); 960/2 x 540/2 = 480 × 270 (third decomposition). The transition to the spectral region is accomplished by performing a two-dimensional Fourier transform or wavelet, and the sub-bands of the two-dimensional spectrum at each decomposition level *i* are subdivided as follows:

- LL_i – low frequencies in the horizontal and low frequencies in the vertical direction.
- LH_i – low-frequency vertical and high-frequency horizontal.
- HL_i – high frequency horizontally and low frequency vertically.
- HH_i – high frequencies horizontally, high frequencies vertical.

In fig.2 we can see an example of a decomposition the original image into three levels.

Fig. 2. Example of the decompositionn original image into three levels

For subsequent processing, the original UHDTV color image is decomposed into components: luminance Y, reddish C_r and blue difference C_b. Each component was processed separately.

III. Fourier Decomposition by Filtration

The Fourier filtering was performed by a low-pass filter, the slope of which was described by the Blackman window function [3] with seven samples per filtering band. The high-frequency part is obtained by subtracting the low-frequency part from the original image. In the Fourier transformation in the spectral region, the images are represented by two arrays of real and imaginary components.

The greatest interest in the design of filters is the area of transition from the passband to the lane of detention. If the amplitude-frequency characteristic in this region has a sharp transition, then distortions at the edges inevitably arise in the image, accompanied by ejections and oscillatory processes. This phenomenon is called the Gibbs effect, and the only recognized method for standard definition television, the method of reducing this kind of distortion is the formation of the Amplitude frequency response AFR slope of the cosine of a similar shape. Form such a slope is possible by multiplying the two-dimensional frequency response of the filter, which is a two-dimensional Fourier transform of the impulse response of the filter, to one of the window functions [4]. Of the considerable number of window functions, we have studied the functions of Blackman, Kaiser, and Hamming. The Blackman window can be specified in the form (1):

$$W_{Blackman} = a_0 - a_1 cos\,(2\pi/N-1) + a_2 cos\,(4\pi/N-1) \qquad (1)$$

Forms of the slope AFR will vary from the number of *N* windows samples, which, in turn, uniquely determine the order of the filter. To filter images, non-recursive filters with finite impulse response are used. If the AFR slope is stretched significantly, then the Gibbs effect becomes invisible, but the sharpness of the image and the frequency response for different values of α will worsen.

The prediction of the values of high-frequency values by low-frequency requires additionally the inverse Fourier transform. Table 1 shows the EPSNR values at the vertical, horizontal and diagonal edges of the image, as well as the PSNR signal-to-noise ratio for the approximated image as a whole, depending on the size of the selected clusters for prediction.

TABLE I. PREDICTION OF HH IMAGES BY LL FOR FOURIER TRANSFORMATION

Cluster size	**Approximation**	**Horizontal**	**Vertical**	**Diagonal**
512x512	63.2 dB	60.8 dB	60.1 dB	38.7 dB
256x256	55.4 dB	51.7 dB	49.7 dB	39.5 dB
128x128	42.2 dB	39.6 dB	39.2 dB	38,2 dB
64x64	40.6 dB	40,2 dB	39.7 dB	32,6 dB
32x32	36.8 dB	34,5 dB	34.1 dB	30,2 dB

Usually, decimation of matrices with interpolation of new values of element brightness is mathematically described as the application of low frequency and high frequency filtering. Nevertheless, each time in these matrices there is information about the entire image. In the high-frequency matrix, the brightness values of the elements are less than in the low-frequency matrix, and, most often, the high-frequency matrix is not transmitted. In this paper, an attempt is made to estimate the possibility of recovering the values of high-frequency matrices from the values of low-frequency matrices.

As a result of the spectral Fourier transform, part of the values of the components turn out to be very small. We varied the value of the threshold level of spectral components, below which they were equated to zero. The simulation was carried out for three variants of the brightness of the luminance signal, in particular eight, ten and twelve-digit representations. The average percentages of the fullness of the matrixes of the clusters by non-zero values at different decomposition levels and for the three signals are presented in Table 2.

Prediction of the values of high-frequency values by low-frequency is done by the approximation method. If you compare the real values of high-frequency values and predicted, you can calculate the standard deviation, and from it calculate the signal-to-noise ratio.

TABLE II. NUMBER OF NON-ZERO ELEMENTS IN DECOMPOSITION CLUSTERS IN FOURE

Level of decomposition	component	Horizontal	Vertical	Diagonal
1	Y	29.6 %	31.4 %	38.5 %
2		31.4 %	22.7 %	30.1 %
3		33.2 %	31.8 %	16.4 %
1	C_b	35.2 %	35.8 %	34.4 %
2		32.7 %	34.6 %	28.6 %
3		28.9 %	31.4 %	30.6 %
1	C_r	34.6 %	36.4 %	33.8 %
2		33.7 %	35.6 %	37.2 %
3		35.5 %	30.8 %	32.1 %

Fig.3 shows the dependence of the RMS on the number of pixels in the Fourier transformation.

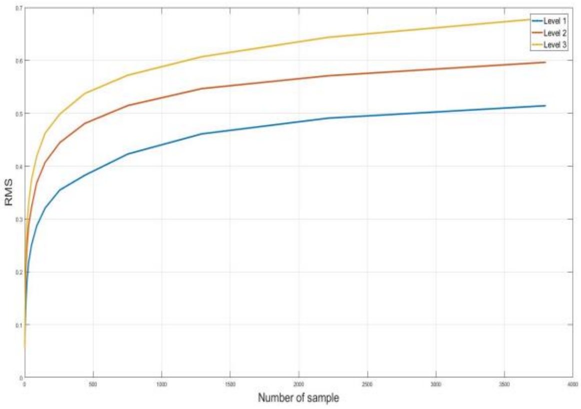

Fig. 3. The standard deviation in the Fourier transformatio

The RMS calculations are performed for each of the three levels of decomposition. The image was restored every time to the full original size. The more decomposition levels, the greater the RMS.

The dependence of the signal-to-noise ratio (PSNR) for a different compression ratio is shown in Fig. 4, where three versions of the digitization capacity of the digital luminance signal are analyzed, namely eight, ten and twelve. In the latter case, at the same value of the compression ratio, the signal-to-noise ratio over the entire image is higher.

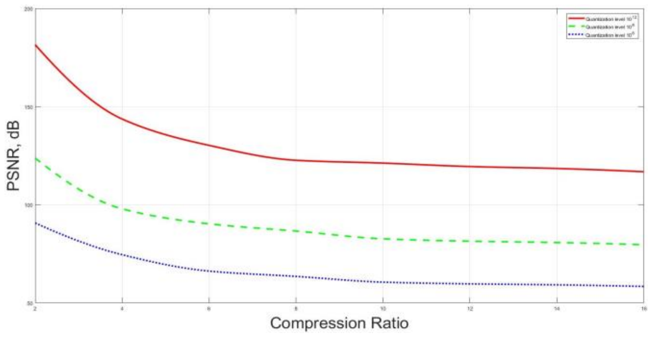

Fig. 4. The PSNR dependence on the compression ratio in the Fourier transformation

Recently, for high-definition and ultra-high definition television, it is increasingly recommended to use such an image quality indicator as the signal-to-noise ratio at the image boundaries (EPSNR). To compute this index, the outlines of objects were selected using the gradient method, and the threshold level is selected according to the semantic content of the image scene. Noise is considered within a certain zone near the contours. The dependence of EPSNR in the Fourier transform for different compression coefficients is shown in Fig. 5.

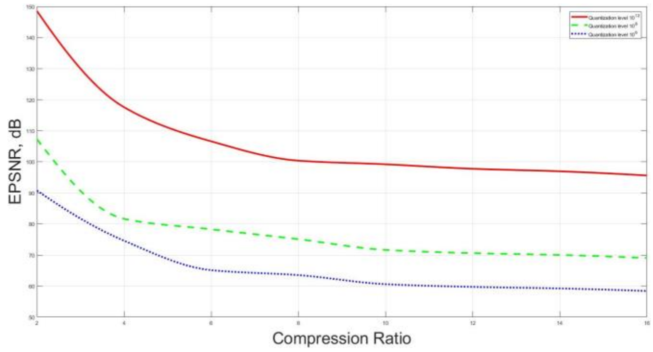

Fig. 5. The EPSNR dependence on the compression ratio in the Fourier transformation

A two-dimensional Fourier transformation was performed on the UHDTV image as a theoretical justification. For practical applications, it is advisable to use the discrete cosine transform or wavelet transform.

IV. DECOMPOSITION OF THE IMAGE BY WAVELET

The application of a discrete wavelet transformation also leads to a hierarchical decomposition of images, and the number of operations for performing decomposition is several times smaller than for Fourier. From the family wavelet were employed in these paper biorthogonal ones. To obtain the next level of decomposition with fewer decomposition elements, spline interpolation is taken advantage of the opportunities to calculate new values. In Table. III are shown the results of comparing the original image and the wavelet decomposed after the wavelet, both in the whole of the entire image and in separate directions.

TABLE III. PREDICTION OF HH IMAGES BY LL FOR WAVELET TRANSFORMATION

Cluster size	Approximation	Horizontal	Vertical	Diagonal
512x512	67.2 dB	65.8 dB	65.4 dB	48.3 dB
256x256	66.4 dB	53.7 dB	53.3 dB	42.1 dB
128x128	42.2 dB	43.8 dB	44.1 dB	41.4 dB
64x64	46.2 dB	40,2 dB	42.7 dB	40.4 dB
32x32	39.8 dB	38,6 dB	39.2 dB	38.3 dB

In the case of a wavelet decomposition, four new matrices with a dimension that is twice as small as the original matrix appear at each successive level. A feature of the wavelet transformation is a large correlation between low frequency and high frequency components compared to Fourier analysis, which allows predicting values under certain conditions. As it can be seen from Table. III, the prediction of high-frequency components gives good results up to the fifth level of decomposition. Thus, if the prediction provides, when recovering PSNR, for example, greater than 35 dB, then the corresponding coefficient matrices are not transmitted.

If we introduce threshold levels for each of the components of the full color signal, then some of the elements of the decomposition matrices can be equated to zero. In Table. IV it is shown the percentage of remaining non-zero values for the luminance and color difference signals at the three decomposition levels.

TABLE IV. NUMBER OF NON-ZERO ELEMENTS IN DECOMPOSITION CLUSTERS IN WAVELET

Level of decomposition	Component	Horizontal	Vertical	Diagonal
1	Y	26.4 %	30.1 %	34.6 %
2		28.6 %	20.5 %	28.7 %
3		29.5 %	29.9 %	17.6 %
1	C_b	30.8 %	29.2 %	28.6 %
2		30.2 %	33.3 %	26.4 %
3		25.4 %	27.8 %	21.6 %
1	Cr	27.3 %	28.6 %	25.8 %
2		30.8 %	29.6 %	30.3 %
3		28.7 %	27.5 %	24.6 %

We adjusted the threshold level for the conversion coefficients, thereby changing the compression coefficient. Fig. 6 shows the PSNR dependence on the compression ratio.

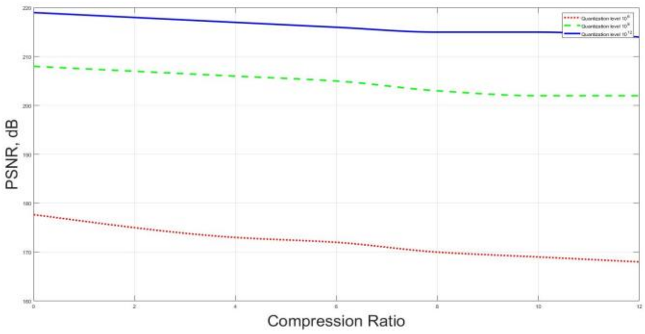

Fig. 6. PSNR with Wavelet transform

Despite the fact that PSNR is widely used as an objective criterion for video signal quality, it is also noted that this criterion does not provide an accurate representation of perceived image quality. According to the results of the analysis of human perception of image quality, it was revealed that human vision is sensitive to deterioration in quality at borders. In other words, if the pixels on the image contours are blurred, the experts will probably give a low rating to the image, even at a high PSNR. Based on this observation, models with a degraded reference signal were created, with the help of which the degradation of quality on the contours is mainly measured. A comparison of the input and output signals may require time matching or spatial matching, the latter to compensate for any vertical or horizontal displacement or cropping of the image. It may also require correction of the shift or gain difference in the luminance and chroma channels. Further, objective evaluation of image quality is calculated, usually by applying the human eye perception model. [8]. Dependence of EPSNR on the desired compression ratio is shown in Fig.7.

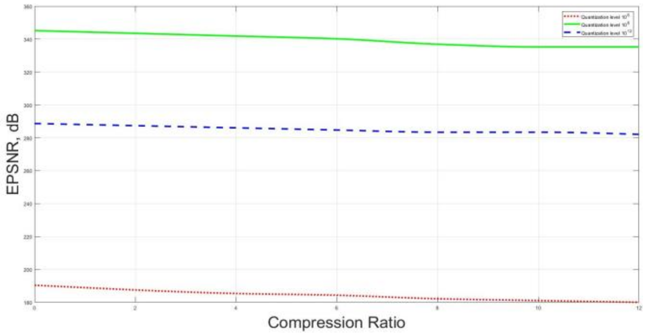

Fig. 7. EPSNR with Wavelet transform

With the help of RR models, the deterioration of quality on the contours is mainly measured. In these models, the contour selection algorithm is primarily applied to the video sequence of the source, in order to determine the contour pixels. The degradation of the quality of these contour pixels is then measured by calculating the root-mean-square error. From the root-mean-square error, a contour PSNR (EPSNR) is calculated.

CONCLUSION

In the simulation it has been tasked to compare compression ratios at about the same ratio of signal to noise. From an analysis of Tables 1 and 3, it can be assumed that the image quality of "excellent", corresponding PSNR> 48dB possible to compress Wavelet-transform image of up to six levels of decomposition that will yield a compression ratio of 67. For the Fourier transform already for 256x256 size clusters after decomposing the image quality closer to the good, and the aspect ratio of 49 times, higher than the standard block by block DCT encoding with no analysis of the contents of high and low frequency components in images.

REFERENCES

[1] Recommendation ITU-R BT.2020-2:2015 “Parameter Values for Ultra-high Definition Television Systems for Production and International Programme Exchange”.

[2] Recommendation ITU-R BT.709-6:2012, “Parameter values for the HDTV standards for production and international programme exchange”.

[3] Recommendation ITU-R BT.1691-1:2009 Adaptive image quality control in digital television systems.

[4] Report ITU-R BT.2245-2:2016 HDTV and UHDTV test materials for assessment of picture quality.

[5] O. Gofaizen, O. Osharovska, M. Patlayenko, V. Pyliavskyi, “Test signals for assessment image quality in HD and UHD TV video path”, in 8-th International conference on Ultrawideband and Ultrashort impulse signals. Odessa: IEEE Conference publications, 2016. – P. 42–46. DOI: 10.1109/UWBUSIS.2016.7724147.

[6] SMPTE ST 2048-1:2011 2048×1080 and 4096×2160 Digital Cinematography Production Image Formats FS/709.

[7] M. Patlaenko, O. Osharovska, “Prefiltration TV high definition images” in Proceedings of the O.S. Popov ONAT, Odessa , vol.2, pp. 144-152 2016.

[8] Recommendation ITU-R BT.1885-0-201103 “Methods of objective measurement of perceived image quality, intended for digital broadcast television of standard definition, in the presence of a reference signal with a reduced frequency band”