

# Video images compression and restoration methods based on optimal sampling

V.N. Drynkin<sup>1</sup>, S.A. Nabokov<sup>1</sup>, T.I. Tsareva<sup>1</sup>

<sup>1</sup> State Research Institute of Aviation Systems (GosNIIAS), Moscow, Russia

## Abstract

The study proposes video images compression and restoration methods based on multidimensional sampling theory that provide four-fold video compression and subsequent real-time restoration with loss levels below visually perceptible threshold. The proposed methods can be used separately or along with any other video compression techniques, thus providing additional quadruple compression.

**Keywords:** video image compression, image reconstruction-restoration, three-dimensional image processing, quincuncial sampling, spatial filtering, spatial resolution.

**Citation:** Drynkin VN, Nabokov SA, Tsareva TI. Video images compression and restoration methods based on optimal sampling. *Computer Optics* 2019; 43(1): 115-122. DOI: 10.18287/2412-6179-2019-43-1-115-122.

## Introduction

Modern video generation, transmission and reproduction requirements stimulate the development of high-quality digital video systems with image sensors having more than 8 million sensels and operating at high frame frequencies (60–120 Hz or more). This leads to a dramatic increase in video stream data rates, which has significant impact on physical communication channels, tightened by spectrum regulations and information storage costs. Under these conditions, the research of effective video compression methods, despite a rather large number of already existing ones, is still relevant.

Transition from high definition (HD, FHD, 2K) to ultra-high definition (UHD, QFHD, 4K) television with each video frame having up to 3840×2160 pixels [1] has led to the development of H.264/MPEG-4 AVC coding standard. Further adoption of 8K video format (up to 7680×4320 pixels per frame [1]) has initiated the development of H.265/HEVC standard [2], which has roughly doubled the compression ratio of H.264.

Some modern video compression algorithms employ discrete wavelet transform [3, 4]; others use adaptive coding, fractal image compression [5] and alternative techniques. However, in most common codecs, starting from H.261, not one but many compression techniques are used employing the so-called hybrid approach [2], which involves a number of procedures, such as block partitioning, inter-frame differences calculation, intra- and inter-frame prediction, motion compensation, various modifications of discrete sine and/or cosine transforms, quantization, etc.

Higher compression ratios within the hybrid approach are possible by improving and optimizing the algorithms being used, but capabilities of those have almost reached the limits by the already achieved compression ratios. In addition, the implementation of such algorithms in real time requires highly advanced equipment, which is not always acceptable (e.g., for industrial television systems) because of the high cost and demands placed on ease of maintenance and reliability in harsh environments.

In this paper, we propose two relatively simple methods of lossy video images compression and one complementary restoration method that provide quadruple com-

pression of video data with real-time restoration, with information loss levels below visually perceptible threshold. These methods, based on multidimensional sampling theory, can be used standalone or in conjunction with any other compression techniques (like the ones described in H.26x and VP8/9 coding standards), providing additional four-fold compression [6].

## Description

### A. Background

The proposed video images compression and restoration methods are based on video signal frequency multiplexing by resampling in order to achieve such sampling of moving images that would be close to optimal [7–9].

Moving pictures (hereinafter referred to as video images or frames) form a message  $x(n_1, n_2, n_3)$ , which is a function of at least three variables: two spatial coordinates (horizontal  $n_1$  and vertical  $n_2$ ) and time coordinate  $n_3$ . Traditional sampling of such images on a rectangular raster suffers from voids, which widen the image spectrum, requiring a broader pass band of the circuitry, and become the source of unwanted noise. In this sense, such sampling cannot be considered optimal. Thus, packing density of a discrete spectrum, achieved by minimizing the number of samples of a discrete signal, provided that initial video quality is preserved, is usually used as a criterion of optimality [7, 10, 11].

Therefore, the problem of optimal sampling of such messages lies in their resampling in order to obtain the densest possible packing of the three-dimensional (3D) discrete spectrum  $S(v_1, v_2, v_3)$  of the message  $x(n_1, n_2, n_3)$  in the frequency space  $\{v_1, v_2, v_3\}$ , where  $v_1, v_2, v_3$  are the corresponding spatial horizontal, vertical and temporal frequencies normalized with respect to their upper values.

During video images restoration (reconstruction) process, the main spectrum is extracted from the full spectrum of the sampled image and the secondary components are suppressed [10, 12] using a space-time reconstructing 3D low-pass interpolation filter (LPF). Such approach is possible because the anisotropy of the properties of the image source and the image receiver is taken into account, that lets one to conclude that the pass region  $D_0$  of the spatial frequency response (SFR)

of the restoring 3D interpolating LPF must have a form of octahedron [7]:

$$D_0 : |v_1| + |v_2| + |v_3| = a, \quad 0 < a \leq 1, \quad a \in \mathbb{R}. \quad (1)$$

As shown in [7] and [9], in order to achieve an extremely dense packing of the 3D spectrum in the 3D message space  $\{n_1, n_2, n_3\}$  for optimal video image sampling, the sampling points of the message  $x(n_1, n_2, n_3)$  need to be staggered (i.e., placed in a quincunx pattern), as shown in Fig. 1. The vectors  $v_1, v_2, v_3$  form a regular triangular lattice of points at which the message counts are taken. Therefore, we will further call such 3D message sampling triangular (although it could be equally called quincuncial or staggered sampling).

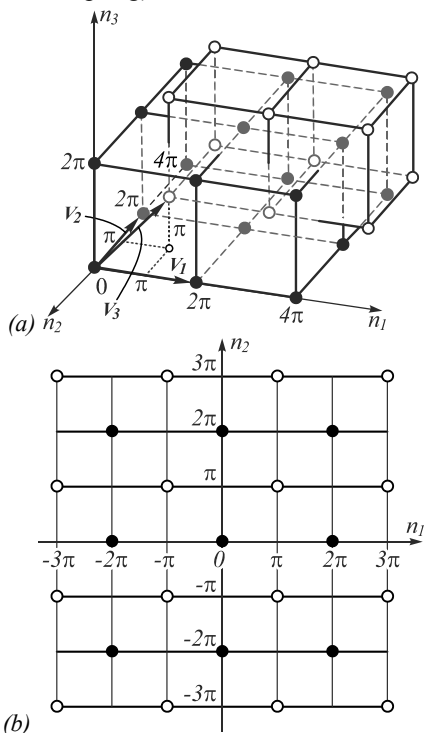


Fig. 1. Sampling points lattice (a) in 3D message space  $\{n_1, n_2, n_3\}$  and their projection (b) onto the spatial frequency plane  $\{n_1, n_2\}$

It should be noted that for the densest packing of the 3D spectrum in message space  $\{n_1, n_2, n_3\}$ , the optimal shape of pass region  $D_0$  would be rhombododecahedron, which is the first Brillouin zone of a body-centered cubic lattice [13]. The octahedral shape is chosen as an approximation, which gives an acceptable SFR, sufficient for practical use.

From Fig. 1 it can be seen that the optimal sampling of video images allows them to be compressed by reducing the number of samples in the original sequence of video frames by resampling them, resulting in a spatio-temporal triangular arrangement of samples.

**B. Video images compression**

Video images resampling for compression can be performed by decimating original video frames through row and column exclusion, e.g., odd columns and rows can be excluded from odd frames, even columns and rows –

from even frames, or vice versa. The remaining samples form a space-time triangular lattice of image samples as shown by white squares in Fig. 2. Such resampling gives four-fold compression of video sequence due to a bifold decrease in video frame sample count and spatial resolution horizontally and vertically.

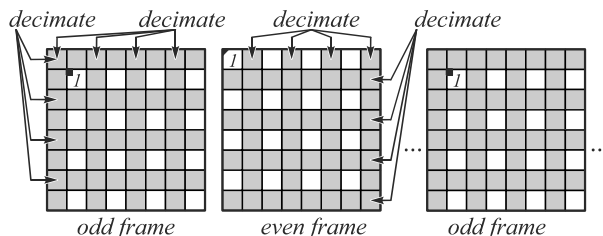


Fig. 2. Video frames compression by sample decimation

Resampling can also be performed using bilinear filtering, i.e., by averaging pixel intensity values over  $2 \times 2$  sample regions (shown in gray in Fig. 3). These regions in neighboring frames should be selected with a one-pixel shift diagonally, as shown in Fig. 3. For example, if in odd frames averaging starts with even rows and columns, and in even frames – with odd rows and columns, then we also get a four-fold compression of the video image size with the space-time triangular sampling structure of the sample intensity values, shown in white in Fig. 3. In this case, an edge effect occurs in some frames, when there are not enough samples to form  $2 \times 2$  regions. Such samples should either be replaced with zero intensities or averaged over the  $2 \times 1$  and  $1 \times 1$  regions, which leads to some complication of the averaging algorithm.

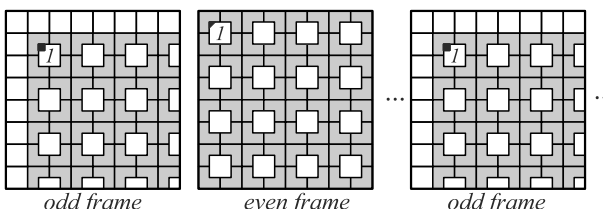


Fig. 3. Video frames compression by sample averaging

Alongside the bilinear interpolation, other well-known traditional non-adaptive methods include cubic interpolation and spline interpolation; all of them are of relatively low complexity. Ones that are slightly more complex use weighted averaging techniques based on different square and non-square window functions, e.g., Lanczos or Fejer. Adaptive methods include the ones that interpolate a missing sample in multiple directions, and then fuse the directional interpolation results by minimum mean square-error estimation [14]. There is also a method of spline domain interpolation of a non-uniformly sampled image with an adaptive smoothness regularization term [15]. Possibly, one of the most complex approaches relies on adaptive two-dimensional (2D) autoregressive modeling and soft-decision estimation [16], which gives promising results in terms of visual quality and peak signal-to-noise ratio (PSNR) values. Applicability of the mentioned interpolation techniques to the methods proposed in this paper is of future concern.

### C. Video images restoration

Reconstruction of video sequence frames compressed by one of the above methods is performed by upsampling and subsequent interpolation.

During upsampling, the size of odd and even frames of the compressed video sequence is restored by interleaving their structure with zero intensity rows and columns corresponding to previously decimated ones. Thus, in any two adjacent frames a space-time lattice with triangular sampling is formed.

During the interpolation process, each upsampled frame is sequentially read and processed using the spatio-temporal 3D reconstructing LPF with the pass region (1).

Harmonized with the SFR of human visual system (HVS) and the spectra of real video images, the octahedral form of the reconstructing LPF transmission region allows for the best extraction of the main image spectrum from the discrete spectrum while also suppressing the side components and high-frequency noise during video images reconstruction from discrete samples. In this case, almost complete restoration of the initial video sequence is provided due to a bifold increase in sample count and a nearly bifold increase in horizontal and vertical spatial resolution of the compressed video frames [17]. Further it will be shown that information loss levels after the reconstruction are kept below the threshold of visual perception.

The proposed approach explicitly determines restoration algorithm of a continuous video signal from its samples, unlike various de-interlacing methods (e.g., Bob, EEDI2, Yadif, MCBob, etc.), based on some heuristic procedures and designed to improve visual quality of standard television signal (PAL, SECAM, NTSC) when reproduced by digital receivers [7].

#### Implementation

The proposed video images compression and restoration methods can be implemented using hardware-based approach (e.g., by using field-programmable gate arrays or application specific integrated circuits) or software-based one with hardware support. Below are the results of software-based implementation with hardware support from general purpose central and graphics processing units (CPUs and GPUs) capable of real-time video processing.

The main element of the compression part of the software is the resampling module that performs resampling either by sample decimation or by sample averaging with a one-sample diagonal shift in adjacent frames, according to previous description. When a video sequence is being input to the resampling module, each frame is compressed according to one of the two methods described above, after which the processed video information is stored on the drive in a pre-selected format for further processing and/or restoration.

The main element of the restoration part of the software is the reconstruction module including a submodule responsible for upsampling the frames of compressed video sequence, and a submodule implementing the 3D

interpolation LPF with a 3D octahedral pass region that restores samples in the reconstructed frames.

The interpolation filter is implemented through a cascade structure of 3D, 2D and one-dimensional (1D) recursively-non-recursive (RNR) blocks with each block consisting of combination of infinite and finite impulse response filters. Such structure makes it possible to form the required octahedral pass region (1) of the 3D LPF  $K(v_1, v_2, v_3)$  with sufficient accuracy for practical use [7], [18]:

$$K(v_1, v_2, v_3) = K[v_3, \phi_3(v_1, v_2)] \times \times K[v_2, \phi_2(v_1)] K(v_1), \quad (2)$$

where  $K[v_3, \phi_3(v_1, v_2)]$ ,  $K[v_2, \phi_2(v_1)]$  and  $K(v_1)$  are SFRs of 3D, 2D and 1D filter blocks, respectively.

In the direction of time frequencies  $v_3$  SFR (2) pass region configuration is formed by 3D RNR block with a frame delay chain  $\exp(-j\pi v_3)$ :

$$K[v_3, \phi_3(v_1, v_2)] = \frac{0.5(1 + e^{-j\pi v_3})(1 - \beta(v_1, v_2))}{1 - \beta(v_1, v_2)e^{-j\pi v_3}}, \quad (3)$$

$$\beta(v_1, v_2) = \frac{\text{ctg } 0.5\pi(a - v_1 - v_2) + w_p}{\text{ctg } 0.5\pi(a - v_1 - v_2) - w_p}, \quad (4)$$

where  $\beta(v_1, v_2)$  is an SFR of 2D non-recursive feedback loop,  $w_p$  is an analog prototype filter pole.

In spatial frequencies plane  $\{v_1, v_2\}$  of the image SFR (2) pass region configuration is formed by 2D RNR block with a row delay chain  $\exp(-j\pi v_2)$ :

$$K[v_2, \phi_2(v_1)] = \frac{0.5(1 + e^{-j\pi v_2})(1 - \beta(v_1))}{1 - \beta(v_1)e^{-j\pi v_2}}, \quad (5)$$

$$\beta(v_1) = \frac{\text{ctg } 0.5\pi(a - v_1) + w_p}{\text{ctg } 0.5\pi(a - v_1) - w_p}, \quad (6)$$

where  $\beta(v_1)$  is an SFR of 1D non-recursive feedback loop.

In the direction of frequencies  $v_1$  SFR (2) cutoff frequency is formed by 1D RNR block with a row element delay chain  $\exp(-j\pi v_1)$ :

$$K(v_1) = \frac{0.5(1 + e^{-j\pi v_1})(1 - \beta)}{1 - \beta e^{-j\pi v_1}}, \quad (7)$$

where feedback circuit coefficient  $\beta$  is calculated according to the following formula:

$$\beta = \frac{\text{ctg } 0.5\pi a + w_p}{\text{ctg } 0.5\pi a - w_p}. \quad (8)$$

To obtain the practically usable structure of the restoration LPF, let us take 1D Chebyshev Type I analog prototype having one real pole  $w_p = -1.9652267$  with pass-band ripple  $\delta = 1$  dB, make  $a = 0.8$  and approximate expressions (4) and (6) with the corresponding trigonometric series [7]:

$$\hat{\beta}(v_1, v_2) = 0.656\gamma - 0.312(\cos \pi v_1 + \cos \pi v_2) - 0.436 \cos \pi v_1 \cos \pi v_2, \quad (9)$$

$$\hat{\beta}(v_1) = 0.114 - 0.778 \cos \pi v_1 + 0.052 \cos 2\pi v_1 + 0.002 \cos 3\pi v_1 - 0.08 \cos 4\pi v_1. \quad (10)$$

In accordance with (8), we obtain  $\beta = -0.716$ . To ensure the stability of a 3D RNR block, the coefficient  $\gamma$  is chosen equal to 0.81.

In order to obtain the transfer function of the restoring LPF (2) in a form suitable for implementation let us using the Euler's formula  $\exp(j\pi v) = \cos \pi v + j \sin \pi v$  make a substitution  $\cos \pi v = 0.5(z + z^{-1})$ , where  $z = \exp(j\pi v)$  is the  $z$ -transform on unit circle [19]. Then the transfer function of the restoring 3D interpolation LPF will have the form:

$$H(z_1, z_2, z_3) = H[z_3, \phi(z_1, z_2)] H[z_2, \phi(z_1)] H(z_1) = \frac{(1+z_3^{-1})[1-\hat{\beta}(z_1, z_2)]}{1-\hat{\beta}(z_1, z_2)z_3^{-1}} \times \frac{(1+z_2^{-1})[1-\hat{\beta}(z_1)]}{1-\hat{\beta}(z_1)z_2^{-1}} \times \frac{1+z_1^{-1}}{1+0.716z_1^{-1}}, \quad (11)$$

$$\hat{\beta}(z_1, z_2) = 0.531 - 0.156(z_1 + z_1^{-1} + z_2 + z_2^{-1}) - 0.109(z_1 + z_1^{-1})(z_2 + z_2^{-1}), \quad (12)$$

$$\hat{\beta}(z_1) = 0.114 - 0.389(z_1 + z_1^{-1}) + 0.026(z_1^2 + z_1^{-2}) + 0.001(z_1^3 + z_1^{-3}) - 0.04(z_1^4 + z_1^{-4}), \quad (13)$$

where  $z_3^{-1}$  represents video image delay,  $z_2^{-1}$  and  $z_2$  represent video image row delay,  $z_1^{-1}$  and  $z_1$  represent video image row element delay.

Block diagram of the restoring 3D interpolation LPF (11) is shown in Fig. 4.

From the upsampling submodule, the sequence of compressed and upsampled odd  $x_1(n_1, n_2)$  and even  $x_2(n_1, n_2)$  frames is being output to the LPF. The restoration of samples is carried out using a combination of 3D RNR block  $H[z_3, \phi(z_1, z_2)]$ , comprising frame delay  $z_3^{-1}$ , 2D RNR block  $H[z_2, \phi(z_1)]$  comprising row delay  $z_2^{-1}$ , and 1D RNR block  $H(z_1)$  comprising row element delay  $z_1^{-1}$ .

After being processed by the LPF, video signal  $x(n_1, n_2, n_3)$  is passed through dynamic range correction and adaptive sharpening submodules. These submodules are implemented in the form of four consecutive window filters performing non-linear processing of the signal in order to reconstruct the original one.

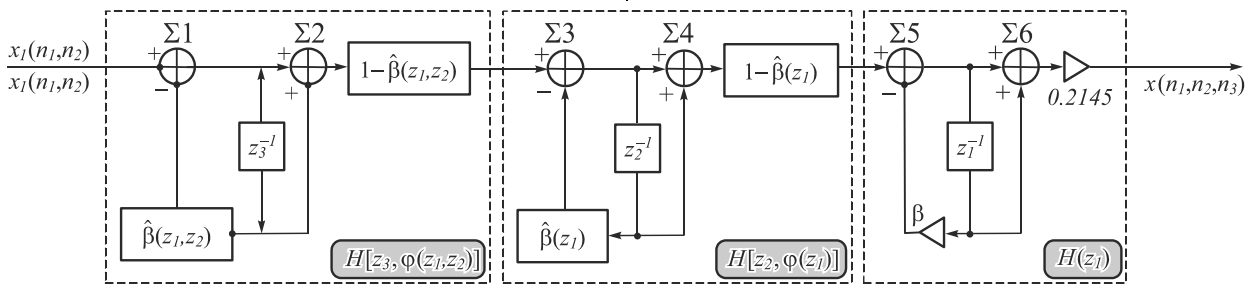


Fig. 4. Block diagram of the restoring 3D interpolation LPF

Software-based implementation of the proposed methods was carried out as a multithreaded application with hardware support of CPUs and GPUs. The source code was written in high-level programming languages C++ and HLSL and optimized for execution on multiprocessor (multicore) systems with shared memory using OpenMP standard. Superscalar architecture of modern CPUs and GPUs have made it possible to organize compression and restoration of 4K 60 Hz video signal in real time by multipass shader processing on a computer with aggregate CPU and GPU single precision performance of just under 1.5 teraFLOPS.

### Experiments and results

Simulation using real-world video images and developed software has been carried out in order to demonstrate the possibility of quadruple video compression with subsequent restoration in real time.

The considered compression and restoration methods are applicable to video images of any resolution and frame rate, but are especially relevant for video streams with high spatial and/or temporal resolution that generate significant amounts of data. Therefore, for testing and evaluation purposes of the proposed methods 4K video

sequences have been selected (see Tables 1 and 2). Another reason for such choice comes from the fact that 4K format has been adopted as a de facto standard for digital cinema, and is becoming the near future broadcasting standard for digital television and streaming multimedia.

Table 1. Test video sequences

Sequence	Codec	Format	Bitrate, Mbit/s
WindAndNature [20]	YUV4MPEG2	2160p, 60 Hz, 4:2:0, 10 bit	15925
TunnelFlag [20]			
Jockey [21]	ProRes 422 HQ	2160p, 30 Hz, 4:2:0, 10 bit	7465
Raptors 60p [22]			
Air Acrobatics [22]		2160p, 59.94 Hz, 4:2:2, 10 bit	2570
			1690

The selected test video sequences have different bitrates, frame rates and use different codecs, which allows to study the interaction of the proposed compression and

restoration methods together with other known methods (codecs) on a wide variety of video content.

Table 2. Test video sequences scene features

Sequence	Features
WindAndNature	Low-detailed small-sized fast moving objects, static footage
TunnelFlag	Medium-detailed small-sized fast moving objects, dynamic footage
Jockey	Medium-detailed large-sized slowly moving objects, dynamic footage
Raptors 60p	Highly-detailed large-sized slowly moving objects, static footage
Air Acrobatics	Low-detailed middle-sized slowly moving objects, dynamic footage

Examples of video images compression and reconstruction according to the proposed methods are shown in Fig. 5 and 6.

A fragment of the original “Raptors 60p” 4K video sequence frame is shown in Fig. 5a. The same fragment quadruply compressed down to 2K format by row and column decimation is shown in Fig. 5b, and by averaging – in Fig. 5c. The same fragment restored back to original 4K format via upsampling and LPF (11) after decimation is shown in Fig. 5d, and after averaging – in Fig. 5e.

For comparison, the same fragment compressed via traditional bilinear and bicubic averaging and further re-

stored by bilinear and bicubic interpolation is given in Fig. 5f and 5g respectively.

As it follows from Figure 5, the frames reconstructed by the proposed methods and the original one are virtually identical. Conversely, traditional interpolation techniques show greater quality reduction.

The proposed methods can be used together with any other video codecs (e.g., H.264, H.265, VP9) to provide additional four-fold video image compression, as shown in Fig. 6.

A fragment of the original “Jockey” 4K video sequence frame is shown in Fig. 6a. Fig. 6b shows the same fragment quadruply compressed down to 2K format by decimation. The sequence of compressed 2K fragments was then coded in accordance with the H.265/HEVC standard using FFmpeg [23] with the following settings: the frame rate was left untouched; target bitrate was set to 4 Mbit/s; color subsampling of the output file was set to 4:2:0 with 8-bit quantization; coding preset was set to “ultrafast”.

The overall video compression ratio achieved here was more than 1800:1. Comparison of compression ratios (taking video quality into account) achievable by common coding standards in conjunction with and separately from the proposed methods is subject of future work.

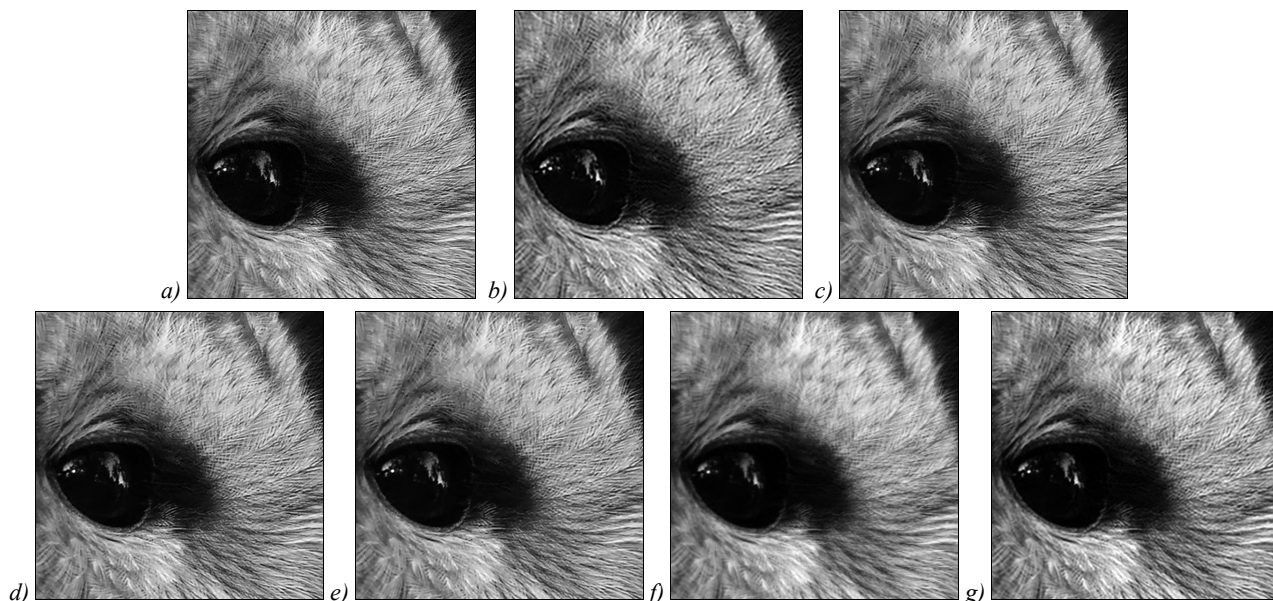


Fig. 5. “Raptors 60p” video sequence frame compression and restoration



Fig. 6. “Jockey” video sequence frame compression and restoration

Video sequence restoration to the original 4K format was carried out in the reverse order: firstly, video was decoded via H.265/HEVC decoder (Fig. 6c), and afterwards it was upsampled and reconstructed by the 3D interpolation LPF (Fig. 6d).

As it follows from Fig. 6, the reconstructed and the original frames look identical. Moreover, because of the LPF interpolation, H.265/HEVC high-frequency coding artifacts that can be seen in Fig. 6c, get significantly reduced in the final image (Fig. 6d). Also, because of the feedback loop in the 3D RNR block of the LPF and moderate frame frequency, the inter-frame restoration noise is present in the final frame. This noise is indistinguishable to HVS during video playback.

### **Quality assessment of video images restored after compression**

When encoding images for the purpose of efficient storage or transmission, it is required to preserve the quality of the reproduced image within the permissible limits [24].

There are two main approaches to static and moving images quality assessment: subjective qualitative assessment based on experts' opinion score, and objective quantitative assessment based on mathematical methods.

Subjective measurement is considered a reliable way of determining video quality and is still widely used in compressive digital television for assessing the quality of video images reconstructed after compression and transmission. Procedures for subjective video quality measurements are described in International Telecommunication Union Recommendations ITU-T P.910 and more recent ITU-R BT.500 [25]. However, subjective assessment has its drawbacks: it is often a rather slow process that requires a group of at least 15 observers [25], each of them having his or her sociocultural or economic background. Therefore, subjective metrics do not always give accurate and robust results.

Quantitative video quality measures are a good alternative to subjective assessment, but that is true only when they correlate with each other. To date, a large number of objective image quality measures have been proposed, for instance, mean absolute difference (MAD), image sharpness measure, mean squared error, Minkowski distance and its variations (e.g., Lebesgue norm, PSNR). However, in a number of cases, namely, when assessing images restored after coding (compression), many of the aforementioned measures do not always correctly reflect structural distortions and correlate badly with the visual ratings [26]. There are a number of video quality metrics that are more consistent with the human perception of image quality. These include structural similarity index, as well as visual information fidelity model [27], the latter employed in the core of the Video Multimethod Assessment Fusion (VMAF) quality metric developed by Netflix [28]. It is important noting that the problem of universal objective quantitative measure of video quality after compression and restoration is not yet fully addressed and requires further research.

Choosing the "right" quantitative video quality metric based on comparative performance analysis, or even developing a new one, requires a separate study and was not the goal of this work. In this paper, it was important only to estimate the quality of the restored videos after compression, in comparison with their original counterparts at least in terms of individual video frames (although this would not be entirely correct for moving images, since the movement itself would not be taken into account). In this sense, the criteria based on the difference between original and restored video images are of interest. It is intuitive that since the difference is zero when the compared images completely coincide, the more the reconstructed image differs from the original, the more non-zero pixels appear in the difference image.

Thus, the following quantitative quality indicators of video images were chosen: MAD  $B_{\text{dif}}$  between the original and reconstructed images, relative number of non-zero pixels (NNZP)  $N_{p \neq 0}$  in the difference image, the width of the difference image histogram  $L_w$ . The quality of restored video images was also controlled visually during comparison. Such approach allowed for quantitative estimations at which image restoration artifacts were visually indistinguishable, i.e., remained below visually perceptible threshold.

According to the chosen metrics, the quality of test video sequences presented in Tables 1 and 2 was estimated after their four-fold compression and restoration using the proposed methods.

Firstly, absolute difference video images were obtained. Then NNZP and MAD values were calculated from difference images pixels whose intensity levels exceeded the threshold value of 10 to cut off the non-essential for HVS changes of the black point in the difference image.

NNZP value was calculated as a percentage of total pixel number in each image frame.

MAD value was evaluated by the following formula:

$$B_{\text{dif}} = 1 \frac{1}{k} \sum_{i=1}^k |b_{1i} - b_{2i}|, \quad (14)$$

where  $b_{1i}$  is pixel intensity value of the original image of size  $m \times n$ ;  $b_{2i}$  is pixel intensity value of the restored image of size  $m \times n$ ;  $k$  is the total number of pixel in the image,  $k = m \cdot n$ .

Histogram width  $L_w$  was calculated at 99th percentile level to cut off the histogram "tail" consisting of bins with insignificant number of pixels (less than 1% of their total number):

$$\sum_{L_i=0}^{L_w} \frac{100N(L_i)}{k} \leq 99, \quad (15)$$

where  $N(L_i)$  is the difference image histogram.

Integral values of the quality metrics of full video sequences were evaluated using the arithmetic mean across all frames of the local metrics of each frame. The results of this calculation are given in Table 3.

Table 3 shows that the largest values (i.e., worst restoration quality) are typical for dynamic footage with fast-

moving objects, while static scenes are restored more accurately, irrespective of object size in both cases. However, as noted above, these metrics do not take image movement into account, so the results are to be revised using other metrics that are more consistent with visual perception of video quality (e.g., VMAF).

Table 3. Quality assessment metrics values

Sequence	$B_{diff}$ , levels	$N_{p=0}$ , %	$L_w$ , levels
WindAndNature	0.15	1.8	9
Raptors 60p	0.48	6.9	13
Air Acrobatics	0.65	6.0	20
Jockey	3.02	21.9	37
TunnelFlag	6.82	37.0	49

Comparative subjective analysis of reconstructed and original video images in motion (during playback) showed that they virtually do not differ, i.e., with the obtained values of the proposed indicators, the restoration artifacts remain visually negligible. The obtained results are a consequence of the fact that the proposed methods of video images compression and restoration, as mentioned before, are developed with due regard to the properties of the source and the human receiver (viewer) of video images, and consistent with multidimensional sampling theory.

### Conclusion

The proposed video compression and restoration methods provide for four-fold compression and virtually lossless for human observer reconstruction of video images in real time that can find application in various areas of image processing, including video encoding and compression systems, television broadcasting, machine vision, video transmittance and storage in computer networks. The proposed methods can be used independently from or together with any other compression techniques, providing additional quadruple compression.

### References

- [1] International Telecommunication Union. Image parameter values for high dynamic range television for use in the production and international program exchange, International Telecommunication Union, ITU-R BT.2100-1, 2017. Source: [http://www.itu.int/dms\\_pubrec/itu-r/rec/bt/R-REC-BT.2100-1-201706-I!!PDF-E.pdf](http://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.2100-1-201706-I!!PDF-E.pdf).
- [2] Dvorkovich VP, Dvorkovich AV, Gryzov YuG. New possibilities of video encoding standard HEVC [in Russian]. Tsifrovaya obrabotka signalov 2013; 3: 2-8. Source: [http://www.dsps.ru/articles/year2013/jour13\\_3/art13\\_3\\_1.pdf](http://www.dsps.ru/articles/year2013/jour13_3/art13_3_1.pdf).
- [3] Filippov AK, Ruffitskiy VA. Method of encoding digital images using discrete wavelet transformation of adaptively defined basis [in Russian], Pat RF of Invent N2429541 of September 20, 2011, Russian Bull of Inventions N26, 2011.
- [4] Shoberg AG, Shoberg KA. Method of direct and inverse fast two-dimensional wavelet-transform [in Russian], Pat RF of Invent N2540781 of February 20, 2015, Russian Bull of Inventions N4, 2015.
- [5] Andreyko DN, Komarov PYu, Ignatov FM. Basic methods of data compression in the transmission of digital videos [in Russian]. T-Comm – Telekommunikatsii i Transport 2013; 9: 12-15.
- [6] Drynkin VN, Tsareva TI. Image compression methods and apparatus. Image restoration method and apparatus [in Russian], Pat RF of Invent N2669874 of September 15, 2017, Russian Bull of Inventions N29, 2018.
- [7] Drynkin VN. Development and application of multidimensional digital filters [in Russian]. Moscow: “GosNIAS” Publisher; 2016.
- [8] Borodyanskiy AA. Hypertriangular sampling of n-dimensional messages [in Russian]. Radiotekhnika 1985; 4: 49-52.
- [9] Borodyanskiy AA. Optimal sampling of moving images, [in Russian]. Elektrosvyaz’ 1983; 3: 35-39.
- [10] Tsukkerman II (eds.). Digital Coding of Television Images [in Russian]. Moscow: “Radio i svyaz” Publisher; 1981.
- [11] Dudgeon DE, Mersereau RM. Multidimensional Digital Signal Processing. Englewood Cliffs, N.J.: Prentice-Hall, 1984.
- [12] Yaroslavskiy LP. Introduction to digital image processing [in Russian]. Moscow: “Sovetskoye radio” Publisher; 1979.
- [13] Entezari A. Optimal sampling lattices and trivariate box splines. Ph.D. Dissertation. Simon Fraser University; 2007.
- [14] Zhang L, Wu X. Image interpolation via directional filtering and data fusion. IEEE Trans. Image Process. 2006; 15(8): 2226-2238.
- [15] Vazquez C, Dubois E, Konrad J. Reconstruction of non-uniformly sampled images in spline spaces. IEEE Trans. Image Process. 2005; 14(6): 713-725.
- [16] Zhang L, Wu X. Image interpolation by adaptive 2-D autoregressive modeling and soft-decision estimation. IEEE Trans. Image Process. 2008; 17(6): 887-896.
- [17] Drynkin VN, Tsareva TI. Videosystem resolution increase method [in Russian]. Pat RF of Invent N2549353 of April 27, 2015, Russian Bull of Inventions N12, 2015.
- [18] Drynkin VN, Tsareva TI. Image resolution increasing method [in Russian]. Tsifrovaya obrabotka signalov 2014; 3: 9-14.
- [19] Rabiner LR, Gold B. Theory and application of digital signal processing. Prentice-Hall, Inc, Englewood Cliffs, New Jersey; 1975.
- [20] Xiph.org. Video test media [derf’s collection]. Source: <https://media.xiph.org/video/derf>.
- [21] Ultra Video Group. Test sequences. Source: <http://ultravideo.cs.tut.fi/#testsequences>.
- [22] Harmonic Inc. Free 4K demo footage. Source: <https://www.harmonicinc.com/4k-demo-footage-download>.
- [23] FFmpeg group. FFmpeg 3.4 2017. Source: <http://ffmpeg.org>.
- [24] Pratt WK. Digital image processing. – N.Y./Chichester/Brisbane/Toronto: John Wiley and Sons, Inc; 1978.
- [25] International Telecommunication Union. Methodology for the subjective assessment of the quality of television pictures, International Telecommunication Union, ITU-R BT.500-13, 2012. Source: [http://www.itu.int/dms\\_pubrec/itu-r/rec/bt/R-REC-BT.500-13-201201-I!!PDF-E.pdf](http://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.500-13-201201-I!!PDF-E.pdf).
- [26] Monich YuI, Starovoytov VV. Image quality evaluation for image analysis [in Russian]. Iskusstvenniy intellekt 2008; 4: 376-386. Source: <http://dspace.nbu.gov.ua/bitstream/handle/123456789/7481/046-Monich.pdf>.
- [27] Sheikh HR, Bovik AC. Image Information and Visual Quality. IEEE Trans. Image Process 2006; 15(2): 430-444.

- [28] Li Z, Aaron A, Katsavounidis I, Moorthy A, Manohara M. Toward a practical perceptual video quality metric. Netflix Technology Blog Jun 5, 2016. Source: <https://medium.com/netflix-techblog/toward-a-practical-perceptual-video-quality-metric-653f208b9652>.

---

*Author's information*

**Vladimir Nikolaevich Drynkin** (b. 1957) graduated from Ryazan Radio Engineering Institute (presently, Ryazan State Radio Engineering University, short – RSREU) in 1981, majoring in Radio Engineering. Currently he works as a Head of Sector at GosNIAS; his interests are image processing, 3D graphics, digital photography. E-mail: [drynkinv@gosnias.ru](mailto:drynkinv@gosnias.ru).

**Sergey Alexeyevich Nabokov** (b. 1986) graduated from Moscow Aviation Institute (MAI) in 2009, majoring in Automated Control Systems of Combat Aviation Complexes, Candidate of Technical Sciences. Currently he works as a Senior Researcher at GosNIAS; his interests are image processing, programming, 3D graphics. E-mail: [nabokov@gosnias.ru](mailto:nabokov@gosnias.ru).

**Tatiana Igorevna Tsareva** (b. 1963) graduated from Lomonosov Moscow State University (MSU) in 1986, majoring in Soil science, Candidate of Biological Sciences. Currently she works as a Senior Researcher at GosNIAS; her interests are image processing, mathematical modeling of processes, 3D graphics. E-mail: [tsareva@gosnias.ru](mailto:tsareva@gosnias.ru).

---

*GRNTI: 47.41.29; 47.51.39*

*Received June 20, 2018. The final version – October 25, 2018.*

---