

RESEARCH

Open Access

Multifractal analysis of 3D video representation formats

Amela Zeković^{1,2*} and Irini Reljin¹

Abstract

One of the main properties of a three-dimensional (3D) video is the large amount of data, which impose challenges for network transport of videos, in applications such as digital video broadcast (DVB), streaming over IP networks, or for transmission over mobile broadband. Addressing these challenges requires a thorough understanding of the characteristics and traffic properties of 3D video formats.

We analyzed 3D video formats using publicly available long video frame-size traces of videos in full high definition (HD) resolution with two views. Examined 3D video representation formats are the multiview (MV) video format, the frame sequential (FS) format, and the side-by-side (SBS) format. We performed a multifractal analysis through extensive simulation and showed multifractal properties of 3D video representation formats. It was shown that the MV video had the highest multifractal nature, while the FS video had the lowest. Also, a part of the multifractal spectrum connected to the highest changes in the signal (high bitrate variability) is analyzed in detail. Changes in multifractal properties for different streaming approaches of 3D videos with aggregated frames are examined, as well as the influence of frame types and values of quantization parameters. Multifractal analysis was performed by the method of moments and by the histogram method.

Keywords: 3D video; Multifractal spectrum; Transport

Introduction

A three-dimensional (3D) video contains several views of a video scene, which provide depth perception for a viewer. 3D video representation formats with one frame sequence are labeled as the frame compatible format, ones with two frame sequences are the stereoscopic multiview format, while ones with more video sequences as the multiview video format [1-5].

The quantity of data for the multiview video representation format is significantly higher than in the case of the conventional single-view video and presents a restriction on storage and transmission of the video. As the number of applications for this video constantly grows, beyond already steadily present cinema applications, towards home and mobile uses, several important issues should be addressed and some problems resolved.

The move from multiuser applications of 3D video towards single-user applications imposes requests for improvement of coding (dealt with in [6-8]), equipment for production of 3D videos [9], and equipment for displaying [10]. An equally important question is the transmission of 3D video formats.

Previous studies of transport characterization of 3D videos are often dedicated to analysis of protocols for delivering 3D video representation formats [11-13]. Other line of research has quality of video as a central subject, for instance, a video-quality-aware routing algorithm for 3D video transmission in wireless networks is presented in [14], while quality-of-experience aware rate adaptation methods for 3D videos are discussed in [15].

In this paper, we present results of research in 3D videos with two views in the multiview (MV) video format, the frame sequential (FS) format, and in the side-by-side (SBS) representation format. We have used publicly available, long frame-size traces (51,200 frames), in full HD 1,920 × 1,080 pixel resolution [4,16]. For characterization of a video and for the network performance evaluation,

*Correspondence: amelaz@viser.edu.rs

¹University of Belgrade, School of Electrical Engineering, Bulevar kralja Aleksandra 73, Belgrade, Serbia

²School of Electrical and Computer Engineering of Applied Studies, Vojvode Stepe 283, Belgrade, Serbia

video traces are often used [17-19]. In our multifractal analysis of 3D video representation formats, we used publicly available long frame-size traces. Analyzed videos had constant values of quantization parameters, indicating a variable bitrate, which is important in the sense of quality, small delay, and higher multiplexing gain [20,21], and also allow to provide multifractal characteristics of the signal from basic standardized encoders that were selected for 3D videos (JSVM 9.19.10 was used for the FS and the SBS formats [22], while JMVC 8.3.1 for the MV 3D video was chosen [23]).

We used the method of moments and the histogram method to calculate multifractal properties of 3D videos. With multifractal characterization by multifractal spectrum and by generalized dimensions, we found that among the views of the multiview video, the highest burstiness is for the combined view (CV), followed by the left view (LV), being the lowest for the right view (RV). Among different representation formats of 3D videos, MV, FS, and SBS, the MV video has the highest burstiness, followed by the SBS format, while the best results are achieved for the FS format.

Streaming with a merging approach is applied for MV and FS representation formats for aggregation of two consecutive frames and for all frames in one group of pictures (GoP). This streaming approach shows significant improvement in variability characteristics, showed by multifractal spectrum, in the case of the MV video over FS video. The bitrate variability shown in [4], by the means of a coefficient of variation (CoV) and a variability distortion (VD) curve, yields to a very similar conclusion.

Our results, obtained by multifractal analysis, can be helpful for the development and improvement of multifractal network traffic models, [24-27], regardless of the investigated 3D video formats. The exact multifractal model can be derived by the investigation of multifractal spectra of known and easily generated multifractal signals such as binomial and multinomial cascades [28,29], in comparison to the multifractal spectra of different 3D video formats provided in this paper, using two different methods, the method of moments and the histogram method. Also, for appropriate model realization, the values of generalized dimensions provided in the paper can be beneficial.

Management and control of video traffic in current and future applications in a variety of networks [30-32] can be improved having in mind detailed characterization of 3D video representation formats, provided in this paper. Also, given the variability of the examined 3D videos, for real-life applications, some bandwidth management techniques are necessary, such as traffic smoothing [20,33,34] and statistical multiplexing [35-37]. These are the areas where our results can also be beneficial.

3D video representation formats

In this section, an overview of the 3D video representation formats, their coding principles, and streaming approaches are presented, [4,5,38,39]. We analyze and compare three main 3D video representation formats: the MV video, the FS video, and the frame compatible (FC) video formats.

Overview of 3D video formats

The MV video contains several views, where each view ν , $\nu = 1, \dots, V$ is one frame sequence. This video format has full resolution of the underlying spatial format in each view. Also, frame rate f for each view ν of the MV video is the same as in the underlying temporal format. For instance, for full HD $1,920 \times 1,080$ pixel MV video format with a frame rate $f = 24$ frames/s, each view has $1,920 \times 1,080$ pixel frames and frame rate $f = 24$ frames/s. For the coding of the multiview video format, multiview video coding (MVC) is used. This type of coding, in addition to temporal and spatial redundancy, utilizes inter-view redundancy. Thus, ITU reference software, referred to as JMVC, first encodes frames of the left view and then uses these frames as reference frames for encoding frames of the right view [23].

The FS video format has only one frame sequence, where frames from different views are interleaved. The spatial resolution of the FS format is the same as in the underlying spatial format. The frame rate of FS format is Vf , where V is number of frames and f is the frame rate of the underlying temporal format. Coding of the FS video format is done by a conventional single-view video encoder, such as JSVM reference implementation of the scalable video coding (SVC) extension of advanced video coding (AVC) encoder [22].

Frame compatible (FC) formats allow utilization of existing infrastructure and equipment for transmission and services for 3D videos. This format has one video sequence with frame rate f that is the same as in the underlying temporal format. FC formats have lower spatial resolution than the underlying spatial format. For example, for the most widely used FC format, the SBS format, frames are spatially sub-sampled in horizontal direction. For instance, for full HD $1,920 \times 1,080$ resolution, the left and the right views of the SBS format have $960 \times 1,080$ pixel frames. These sub-sampled frames are interleaved into one frame in full HD resolution. As in the case of the FS format, SBS representation also uses conventional single-view video encoder for coding.

3D video traces

For an evaluation of multifractal properties in order to estimate traffic characteristics of 3D videos, long, publicly available, frame-size video traces are used [4,16]. We examined 3D videos with two views ($V = 2$) in the MV,

the FS, and the SBS representation formats. Coding of the multiview video is performed by applying reference software JMVC (version 8.3.1), while for coding of FS and SBS formats, H.264 reference software JSVM (version 9.19.10) in a single-layer encoding mode is used. Each view had 51,200 full HD 1,920 × 1,080 pixel frames and the frame rate $f = 24$ frames/s. We performed evaluations with Tim Burton's movie *Alice in Wonderland*, which is a movie with a combination of live action and computer animation. We analyzed different 3D representation formats of this movie and videos with different quantization parameter settings. The values of quantization parameters in the main analysis were $q_p(I, P, B) = (28, 28, 28)$. Additionally, videos with quantization parameters $q_p(I, P, B) = (24, 24, 24)$, $q_p(I, P, B) = (34, 34, 34)$, where the parameters that are the same among the frame types, and videos with quantization parameters $q_p(I, P, B) = (15, 15, 21)$, $q_p(I, P, B) = (20, 20, 26)$, $q_p(I, P, B) = (24, 24, 30)$, and $q_p(I, P, B) = (30, 30, 36)$, have been analyzed as well. In the main analysis, GoP length for the MV and SBS formats was 16 frames. The FS format had GoP with 32 frames, which means that all encodings have the same playback time between intracoded (I) frames. GoP pattern in the main analysis was B1, which means one bi-directional (B) frame between successive intracoded (I) and predictive encoded (P) frames, while additional consideration was performed on B7 pattern videos.

Streaming of 3D videos

Streaming of the SBS representation format is performed frame by frame, where each frame is integrated from spatially sub-sampled frames from the LV and the RV, and with the same frame rate as with the underlying temporal format. Streaming of the MV representation format can be performed in several different ways. The basic way of streaming the MV video is to stream each view individually. A second streaming option is to perform some kind of merging of views, such as sequential (S) merging or aggregation by combining (C). With sequential merging, frames from different views are used to form one sequence in the following order: first view 1 of frame 1, followed by view 2 of frame 1, ..., followed by view V of frame 1, followed by view 1 of frame 2 ... We also name this signal as the CV. With aggregation streaming approach, multiview frames are formed, where one multiview frame is the sum of all frames with the same frame number from different views. For the FS format, a sequential and aggregation streaming approach can be applied. Aggregation of frames on the level of two frames performs smoothing of the data across $V = 2$ views. This approach can be further extended on the level of 16 frames, one GoP of the encoder. Aggregations of two frames are labeled as CV-C 2 for the multiview representation format and FS-C 2 for the frame sequential format, while aggregations of 16 frames

are labeled as CV-C 16 and FS-C 16 in the following text.

Estimation of multifractal properties

Fractals can be viewed as sets with visual expression of a region drawn in black ink against white paper. Most natural phenomena cannot be expressed in terms of contrast between black and white, and they demand more general mathematical objects that embody the idea of 'shades of gray'. These more general descriptors are called measures. For instance, measures can represent level of ground water, pixel values from pictures, or frame sizes as in our case. When a measure performs high variability at all scales, and when the variability is the same at all scales, or at least statistically the same, one says that the measure is self-similar or that is multifractal. Self-similar sets have a property that each piece (regardless how small) is identical to the whole after some rescaling and translation, [28,29,40,41].

A process that fragments a set into smaller and smaller components according to a rule and at the same time fragments the measure of the components by another rule is called a multiplicative process or cascade. The simplest multiplicative process is a binomial cascade.

For characterization of multifractals, only one number, such as fractal box-counting dimension D , is not sufficient. If a set S supporting a measure μ is covered by boxes of size ϵ and if the number of boxes $N(\epsilon)$ is evaluated, one can determine box-counting dimension as $N(\epsilon) \sim \epsilon^{-D}$. A problem with this characterization is that the value of the measure in each box is disregarded.

Usually, density of probability defined as $\mu(S)/\epsilon^E$, where E is the Euclidian space dimension, would be used to describe properties of the data, but as $\epsilon \rightarrow 0$, this characterization loses its meaning. Instead, the density becomes embodied in a quantity,

$$\alpha = \frac{\log \mu(\text{box})}{\log(\epsilon)} \quad (1)$$

called coarse Hölder exponent. This quantity is the logarithm of measure of the box over the logarithm of the size of the box. Usually, α is restricted to a region $[\alpha_{\min}, \alpha_{\max}]$, where $0 < \alpha_{\min} < \alpha_{\max} < \infty$.

Once α is determined, the next step is to find the frequency distribution of α . For each value of α , number of boxes $N_\epsilon(\alpha)$ that have coarse Hölder exponent equal α is determined. As in the previous case, determining a probability of hitting the value α and further distribution of these probabilities does not have meaning because as $\epsilon \rightarrow 0$, these values no longer converge to a limit. This is why weighted logarithms and a function

$$f_\epsilon(\alpha) = -\frac{\log N_\epsilon(\alpha)}{\log(\epsilon)} \quad (2)$$

are required. This function for $\epsilon \rightarrow 0$ converges to a limit $f(\alpha)$. The definition of $f(\alpha)$ means that for decreasing the box size ϵ , the number of boxes with coarse Hölder exponent equal α , $N_\epsilon(\alpha)$, increases by the scaling relation $N_\epsilon(\alpha) \sim \epsilon^{-f(\alpha)}$. Function $f(\alpha)$ describes distribution of α . Graph $f(\alpha)$, usually called a multifractal spectrum or $f(\alpha)$ curve, has for some simple types of multifractals (such as binomial cascade) shape of mathematical symbol \cap . For some multifractals $f(\alpha)$, the curve can lean to one side.

Alternatively, quantity α is called singularity strength, while $f(\alpha)$ represents singularity or Hausdorff singularity, and $f(\alpha)$ curve is labeled as $f(\alpha)$ singularity spectrum. Singularity α follows local changes in the signal, while $f(\alpha)$ provides global characteristics of data, [28,29,42-44].

An empirical self-similar measure has only one, n th, stage of measure known. So, for the evaluation of multifractal spectrum, previous stages of measure reconstructed by coarse-graining the measure are necessary. Given the discrete data, the smallest measures are the given n th stage, and it is for the size of the boxes $\epsilon = 1$. Sum of all measures in a stage is one, or normalized to one.

In our research, two methods for obtaining an estimate of $f(\alpha)$ curve are used: the method of moments and the histogram method. The method of moments is chosen as a method in which $f_\epsilon(\alpha)$ converges to $f(\alpha)$ the fastest, resulting in a short execution time. The second histogram method has slower convergence of $f_\epsilon(\alpha)$ to $f(\alpha)$ and slower execution, but has a tendency to show additive processes in signal and allows inverse multifractal analysis (determination of the exact part of data with chosen values of pair $(\alpha, f(\alpha))$). Also, these two methods are different in the way they handle the data, where the method of moments tends to smoothen the data, while the histogram method handles raw data and has less approximation.

The method of moments

The first step in evaluating a multifractal spectrum by the method of moments is covering the self-similar measure with non-overlapping boxes of size ϵ_k . We have used values $\epsilon = [1, 2, 4, 8, 16]$ for the size of boxes. Now, partition functions, defined as

$$X_q(\epsilon) = \sum \mu(\epsilon)^q, \quad (3)$$

are calculated, where q is the moment order, $q \in \mathbb{R}$, and $\mu(\epsilon)$ is the total measure in the boxes of size ϵ . Function $\tau(q)$ is estimated as the slopes of plots $\log(X_q(\epsilon))$ versus $\log(\epsilon)$. Coarse Hölder exponent $\alpha(q)$ is now calculated by numerical differentiation of $\tau(q)$ over q values. Finally, minimization over q for equation

$$f(\alpha) = \min_q (\alpha(q)q - \tau(q)) \quad (4)$$

which is known as Legendre transform is performed [28]. Plotting $f(\alpha)$ versus α gives an estimation of the multifractal spectrum.

Using the method of moments, in addition to multifractal spectrum $f(\alpha)$, it is possible to evaluate D_q spectrum, where

$$D_q = \frac{1}{q-1} \tau(q). \quad (5)$$

Values D_q are known as generalized dimensions [29,45]. Especially interesting are dimensions for $q = 0$, $q = 1$, and $q = 2$; D_0 , D_1 , D_2 , respectively. Dimension D_0 is usually called the fractal dimension, dimension D_1 is the information dimension, while dimension D_2 is called the correlation dimension. Dimension D_0 is equal to the maximum of the multifractal spectrum $f(\alpha)$, when the most probable α occurs, labeled as α_0 . Dimension D_1 is called information, because it is proportional to $\mu \log(\mu)$ that scales similarly to the information for probability distribution. The correlation dimension D_2 defines probability that two randomly chosen points are on the distance greater than ϵ . These generalized dimensions and D_q spectra are evaluated for our data using (5).

It has been shown that minimal values of multifractal spectra correspond to $q \rightarrow -\infty$ for α_{\max} and to $q \rightarrow \infty$ for α_{\min} . Also, maximal values of partition function are found for $(\alpha_{\min}, f(\alpha_{\min}))$, while minimal values occur for $(\alpha_{\max}, f(\alpha_{\max}))$ [29].

An algorithm for evaluation of multifractal spectrum by the method of moments is also implemented using sliding boxes, instead of non-overlapping, for covering the measure. The results for these additional tests are the same as in the previous method for $q > 0$, but for $q < 0$, the multifractal spectrum for sliding boxes shows missing undefined values of the spectrum. This is characteristic of some fractals as reported in [28].

The histogram method

The histogram method of determining multifractal spectrum starts with covering the measure with boxes of size ϵ . In the case of this method, $f_\epsilon(\alpha)$ slowly tends to $f(\alpha)$. So, for better estimation, instead of non-overlapping boxes, sliding boxes are used to cover the measure.

We used $n_\epsilon = 8$ different sizes of boxes, with the following values $\epsilon = [1, 3, 5, 9, 13, 21, 29, 37]$, and therefore ϵ_k is indexed with $k = 1, 2, \dots, n_\epsilon$. For each ϵ_k , total measure of the boxes is determined, $\mu_{i,k}$, where $i = 1, 2, \dots, n$. Length of the data and ϵ_k values determine the value of n , where for the smaller box size, the larger number of total measures exists. For easier calculation, all measures are stored in matrix M , where the size of the matrix is $[n_\epsilon \times n]$, for the largest possible n (the smallest ϵ_k). Coarse Hölder

exponents $\alpha_i, i = 1, 2, \dots, n$ are now determined as slopes of plots $\log(\mu_i, k)$ versus $\log(\epsilon_k/L)$, where L is the length of one-dimensional data.

A range of α values, $[\alpha_{\min}, \alpha_{\max}]$ is discretized in $D = 100$ pieces of equal length $\Delta\alpha$ and values α_d are formed as centers of the intervals. In the domain of α values α_i , for different values of $j, j = 1, 3, 5, \dots, 199$ number $N_j(\alpha_d)$ is determined, as a number of boxes of size j that have α_i value in the region of $\Delta\alpha$ around α_d . This procedure is conducted for all values of $\alpha_d, d = 1, 2, \dots, 100$. Finally, $f(\alpha_d)$ values are calculated as slopes of plots $-\log(j/n)$ versus $\log(N_j(\alpha_d))$. Graph $f(\alpha_d)$ versus α_d is an estimation of $f(\alpha)$ curve.

For discussion of multifractal properties, it is important to know that α_{\min} corresponds to the highest value of the measure, while α_{\max} is related to the smallest and the smoothest data.

Simulation results

We examined 3D video representation formats: the MV video and its different views (the LV, the RV, and the CV), the FS format, and SBS format, discussed in the previous section. Multifractal properties of 3D video representation formats are calculated using the method of moments and the histogram method. First, we present the results obtained by the method of moments, that has a higher level of approximation, and later, the results by the histogram method that handles raw data.

Multifractal analysis by the method of moments

Multifractal spectra

In this section, calculated multifractal spectra of examined 3D video representation formats by the method of moments are presented. We first analyze spectra of the views of the multiview video and spectra of different 3D

video formats; then, we proceeded to examine multifractal spectra of different streaming approaches of the videos, influence of quantization parameters values, and frame types on multifractal properties.

Multifractal spectra of different views of the multiview video and spectra of different 3D video representation formats are presented in Figure 1 a,b. It can be seen that among the views of the multiview 3D video, the highest variability, that is the smallest values of α , is present in the case of CV, followed by LV, while RV, the video with the smallest frame sizes, has the lowest burstiness. A comparison of multifractal properties by multifractal spectra for different 3D video representation formats shows the highest burstiness in the case of CV of the multiview video, followed by the SBS 3D video representation format, while the FS format has the lowest level of burstiness. It should be noted that although the SBS format falls between the other two 3D formats in the sense of smallest values of α , the values of $f(\alpha)$ in this part of the spectrum for SBS are higher than those for the other two formats. This means that the highest values of the frames for this video are more frequently present than for the other two formats.

In the sense of variability, as it was presented, the FS 3D video format shows better properties than the CV 3D video format, but an aggregated streaming approach can make an advantage for the CV 3D video. We analyzed multifractal properties by spectra for videos with aggregated frames (aggregation levels 2 and 16) for FS and CV, and in Figure 2, these results are presented. It can be seen that the smallest values of α are in all cases higher with than without aggregation (comparison with multifractal spectra in Figure 1). A difference in burstiness between the FS and CV format is lower for aggregation level 2, and approximately the same for level 16. Also, a closer look to the right side of the spectra ($\alpha > 1$) in Figure 2 indicates

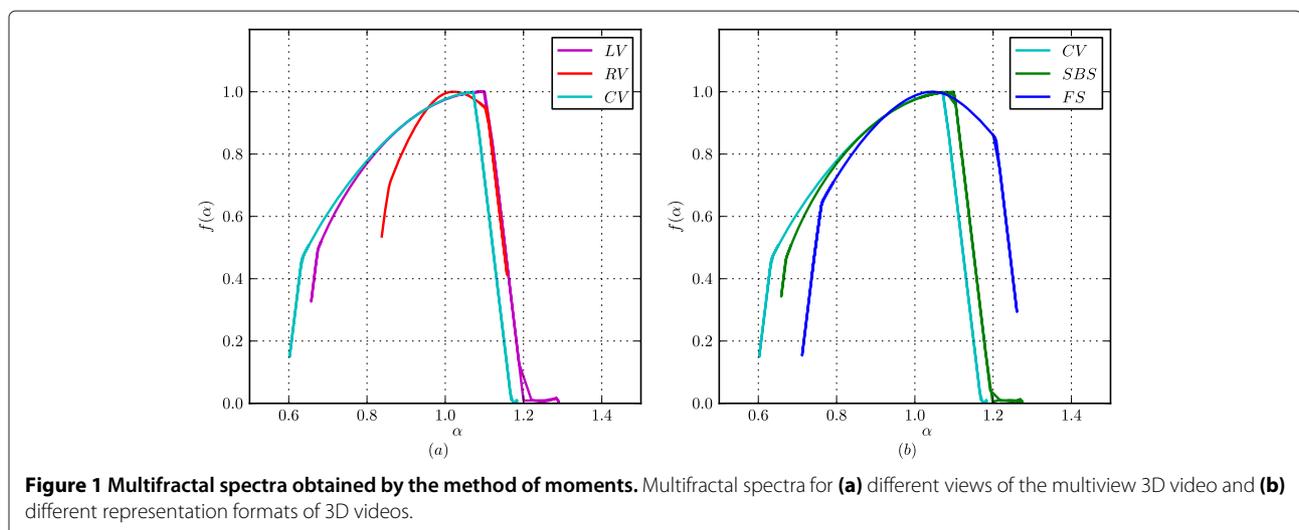
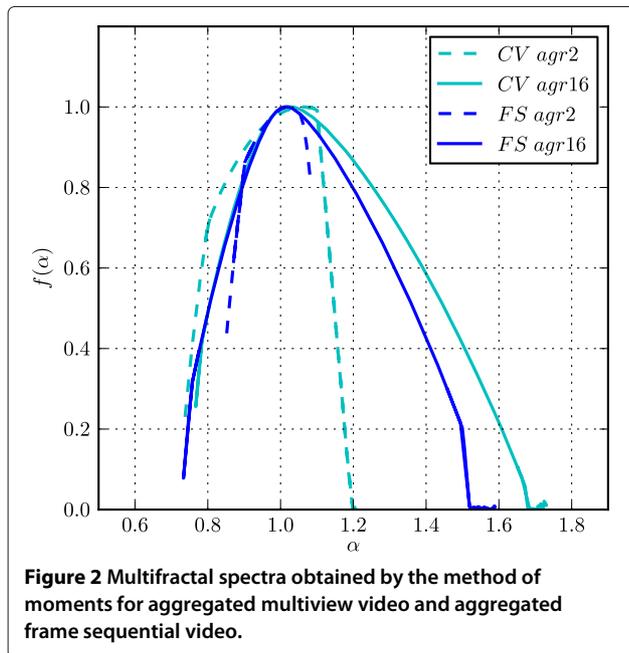


Figure 1 Multifractal spectra obtained by the method of moments. Multifractal spectra for (a) different views of the multiview 3D video and (b) different representation formats of 3D videos.



a significant improvement in the sense of small values in new aggregated sequences of frames for both types of 3D representation formats, especially for CV.

Multifractal spectra were examined for videos with different quantization parameters q_p . Regardless of the 3D video representation format, very similar changes occur in the spectra with the change of quantization parameters. As an example, results for CV of multiview 3D video are presented in Figure 3a. These spectra show that the video with the smallest value of quantization parameters has the lowest burstiness. The other side of the spectrum (the

right-hand side) shows that the video with the smallest value of q_p has the highest level of small values, and with a high level of $f(\alpha)$, these events are frequent in the video. The experiment is repeated in the case of the multiview 3D video with quantization parameters different than those used in the main analysis ($q_p(I, P, B) = (24, 24, 24)$, $q_p(I, P, B) = (28, 28, 28)$, and $q_p(I, P, B) = (34, 34, 34)$) on the video with GoP G16B7 and values of quantization parameters $q_p(I, P, B) = (15, 15, 21)$, $q_p(I, P, B) = (20, 20, 26)$, $q_p(I, P, B) = (24, 24, 30)$, and $q_p(I, P, B) = (30, 30, 36)$. Results from this analysis are presented in Figure 3b, and as in the case of the previous set of quantization parameters, they show wider multifractal spectra for higher values of quantization parameters, meaning higher variability. Also, the values of α_{\min} have lower values, connected to the highest burstiness, in the new set of quantization parameters.

Multifractal properties given by the spectrum are examined for different frame types, I, P, and B, for examined 3D video formats. These results are shown in Figure 4. Multifractal spectra for I frames show that this frame type has very similar multifractal properties, regardless of representation format, which is a consequence of a very similar way of coding these frame types for different 3D videos. The left-hand sides of these multifractal spectra ($\alpha < 1$), where high values and big changes in the signal are present, show small values of α for P frame types, then very close are I frames, and the highest values in the left-hand side of spectra for B frames. Although, P frames have the smallest values of α , I frames have higher value of $f(\alpha)$ which means more frequent higher values.

Multifractal properties are usually presented using multifractal spectra, as it was done previously, but by extracting characteristic points of the spectrum, it is possible to

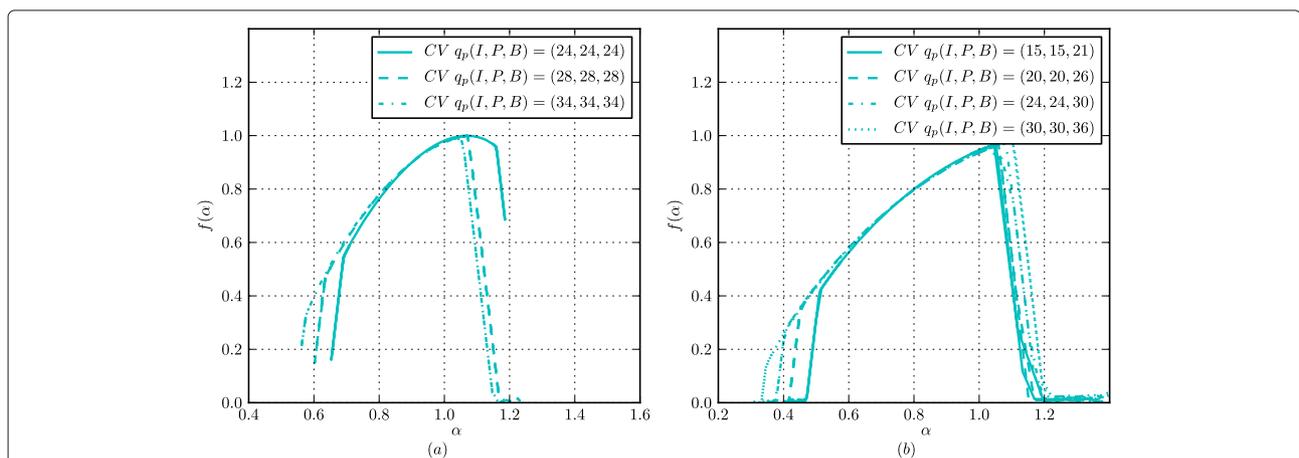


Figure 3 Multifractal spectra obtained by the method of moments for the different values of quantization parameters. Multifractal spectra obtained by the method of moments for the multiview video with sequentially merged views (the combined view), for different values of quantization parameters: **(a)** $q_p(I, P, B) = (24, 24, 24)$, $q_p(I, P, B) = (28, 28, 28)$, and $q_p(I, P, B) = (34, 34, 34)$; and **(b)** $q_p(I, P, B) = (15, 15, 21)$, $q_p(I, P, B) = (20, 20, 26)$, $q_p(I, P, B) = (24, 24, 30)$, and $q_p(I, P, B) = (30, 30, 36)$.

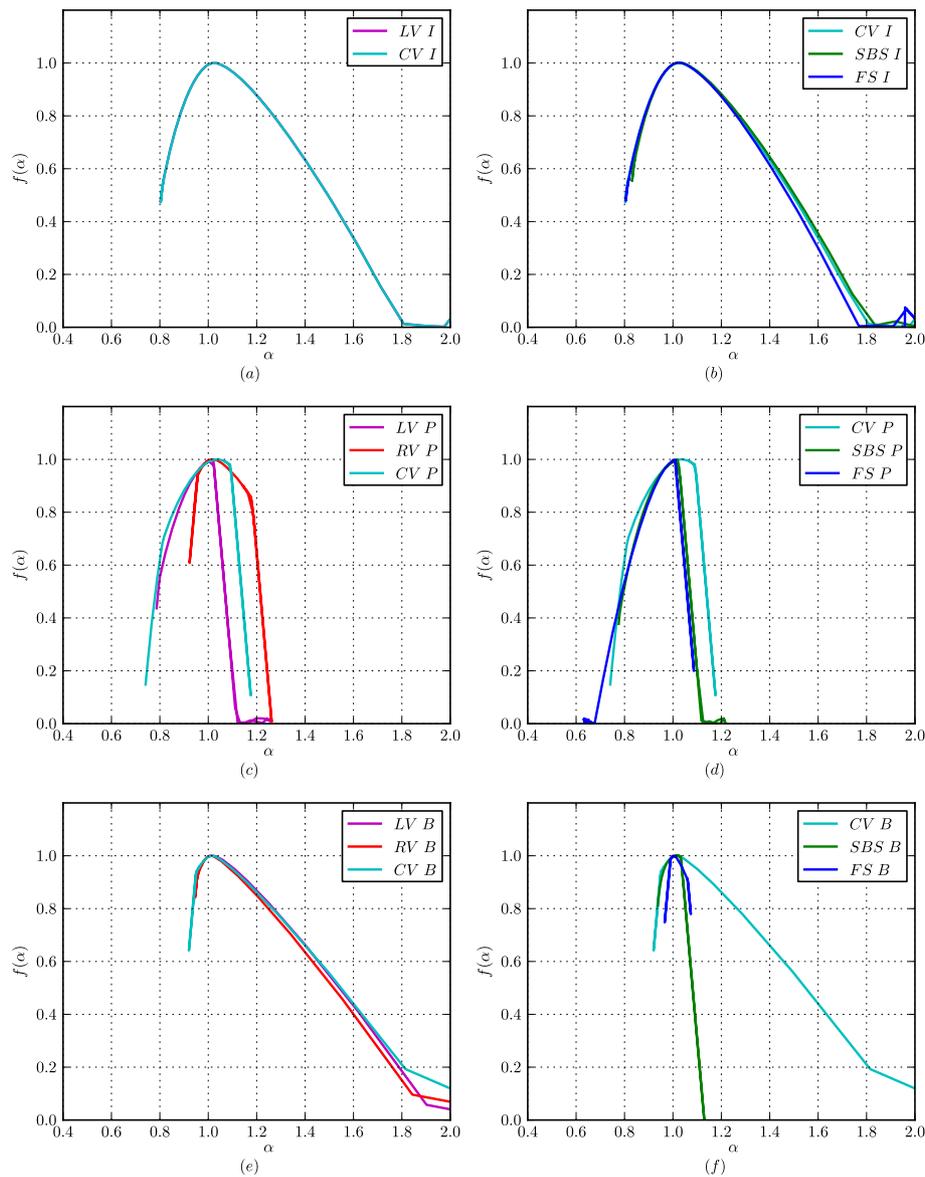


Figure 4 Multifractal spectra obtained by the method of moments for only one frame type isolated from 3D videos. Multifractal spectra for (a) only I frame types from different views of the multiview 3D video, (b) only I frame types from 3D videos in different representation formats, (c) only P frame types from different views of the multiview 3D video, (d) only P frame types from 3D videos in different representation formats, (e) only B frame types from different views of the multiview 3D video, and (f) only B frame types from 3D videos in different representation formats.

conduct a more accurate, but narrower, quantitative comparison. Numerical results for characteristic points of the spectrum - f_{\max} , $\alpha(f_{\max})$, α_{\min} , and α_{\max} , by the method of moments, for 3D video representation formats are given in Table 1. It can be seen that in all of the cases, values for f_{\max} and $\alpha(f_{\max})$ are very close to 1. Also, the values of $\alpha(f_{\max})$ given in Table 1 show that RV has the smallest value $\alpha(f_{\max})$ of 1.0216, that is the simplest structure in the most frequent case, while LV has the highest value of $\alpha(f_{\max})$ of 1.1007. The most prominent burstiness (the smallest value α_{\min}) from the results of Table 1

is present in the case of CV of the multiview 3D video of $\alpha_{\min} = 0.6023$, while the lowest level of burstiness is present in the case of RV of $\alpha_{\min} = 0.8377$. In the cases with aggregated frames, burstiness is lower (higher values of α_{\min}), so for CV, α_{\min} grows from 0.6023 without aggregation to 0.7357 for aggregation of two frames and to 0.7668 for aggregation on the GoP level, while for FS format it goes from 0.7119 without aggregation to 0.8508 (aggregation of 2 frames) and 0.7331 (aggregation of 16 frames). Aggregation of frames to the great extent improves the level of burstiness of CV of the multiview

Table 1 Comparison of multifractal properties of 3D video formats obtained by the method of moments

3D video	f_{max}	$\alpha(f_{max})$	α_{min}	α_{max}	D_0	D_1	D_2
LV	1.0003	1.1007	0.6574	1.2898	1.0000	0.9013	0.8195
RV	1.0001	1.0216	0.8377	1.1577	1.0000	0.9799	0.9585
CV	0.9994	1.0717	0.6023	1.1844	1.0000	0.8990	0.7893
SBS	1.0002	1.1003	0.6585	1.2746	1.0000	0.9016	0.8231
FS	1.0001	1.0445	0.7119	1.2608	1.0000	0.9456	0.8736
CV-C 2	1.0002	1.0641	0.7354	1.2182	1.0000	0.9375	0.8831
CV-C 16	1.0007	1.0234	0.7668	1.7303	1.0006	0.9800	0.9593
FS-C 2	1.0000	1.0222	0.8508	1.0804	1.0000	0.9735	0.9406
FS-C 16	1.0006	1.0144	0.7331	1.5892	1.0005	0.9857	0.9672

3D video format and puts it in this sense closer to FS 3D format.

Generalized dimensions

In addition to multifractal spectra, the method of moments allows calculation of generalized dimensions not only from a multifractal spectrum (as other methods), but also directly calculating from values of τ and q . We have chosen the latter approach.

In Table 1, generalized dimensions D_0 (the fractal dimension), D_1 (the information dimension), and D_2 (the correlation dimension) for 3D video representation formats are given. Dimensions D_0 are approximately in all cases equal to 1. Values of f_{max} , given in Table 1, are actually values of D_0 directly from the multifractal spectrum, and they are all also very close to 1. These values mean that the mostly present fractal dimension is approximately equally probable for all 3D video representation formats. If we order 3D video formats by their values of information and correlation dimensions, as presented in Table 1, a very similar regularity would be observed. The highest correlation dimension is in the case of the CV video, followed by LV, SBS, FS, and RV videos. In the research about traffic characteristics of 3D video formats [4], where the same 3D videos are used, the order of the videos by CoV criteria is exactly the same as in our results for the order of the video by values of the correlation dimension D_2 . Aggregation on the level of 2 frames of the videos leads to lower values of CoV and closer characteristics of the CV and FS formats (CV CoV moved from 1.3334 to 1.0731 and for FS moved from 1.0338 to 0.8108), which is consistent with higher and closer values of the correlation dimensions for these formats. Values for CoV for aggregation of 16 frames (on the level of GoP) for CV and FS for the movie *Alice in Wonderland* are not given in [4], but by repeating and extending their research, we found CoV for CV-C 16 to be 0.7416 and for FS-C 16 0.6507, which means that CV and FS videos in the sense of coefficient of variation are even closer together. The same regularity

can be observed by looking in correlation dimensions D_2 , given in Table 1, where the dimensions are getting higher and closer. According to the values of CoV and D_2 , the FS 3D format has slightly smoother traffic than the CV multiview 3D format, even with aggregation, but based on the burstiness, that is, the highest for the smallest value of α_{min} , CV-C 16 has better properties than FS-C 16, as presented in our results of the multifractal analysis given in Table 1.

The method of moments allows simple calculation of a generalized dimension spectrum, using function $\tau(q)$. These spectra and the function $\tau(q)$ for different views of the multiview 3D video are shown in Figure 5. It can be seen that RV, the view that utilizes inter-view prediction and has the smallest frames, has the narrowest range of generalized dimensions D_q , while LV and CV have similar but higher range of values.

Multifractal analysis by the histogram method

Multifractal spectra

3D video representation formats are examined in a multifractal sense using the histogram method. Multifractal spectra are provided having in mind different views of the multiview video, different 3D video formats, different streaming approaches, quantization parameter values, and frame types.

Multifractal spectra obtained by the histogram method for different views of multiview 3D video and for different 3D representation formats are shown in Figure 6a,b, respectively. Views of the multiview 3D video have different complexity of data (frame size sequences), where RV has the lowest complexity, while CV and LV have similar higher complexity, given the diversity of the dimensions in each spectrum. Also, the maximum of multifractal spectrum for RV is the highest (most frequent case), while singularity $\alpha(f_{max})$ is the lowest in comparison to LV and CV. Multifractal spectra for CV of multiview 3D video, FS and SBS 3D representation formats show that the highest width of the spectrum (the most complex data) is present

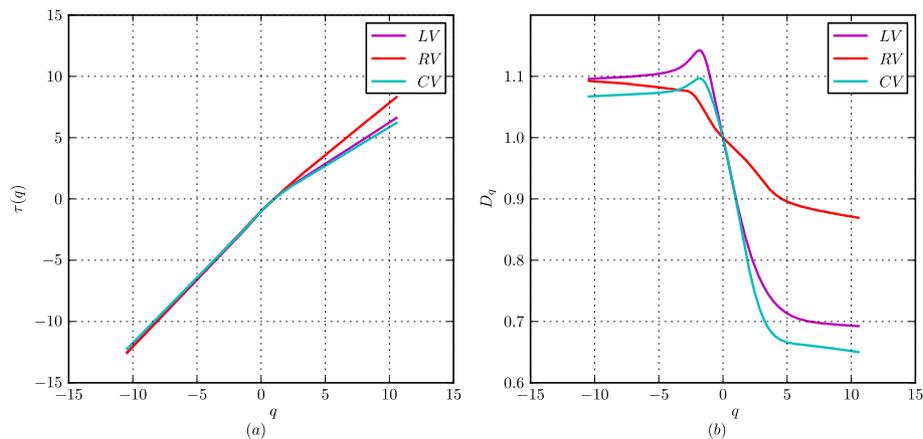


Figure 5 Multifractal analysis applying the method of moments. **(a)** Function $\tau(q)$ obtained by the method of moments. **(b)** Generalized dimensions for views of the multiview 3D video.

for CV, followed by SBS, while the least complex structure is found in the case of the FS format.

Multifractal spectrum of RV has two dominant bumps in the top of the spectrum, as a consequence of the two processes present in the data - P and B frame types in the signal that are formed using LV frames as a reference. Similar, but a less distinctive process, is present in the case of CV spectrum. Additive processes could not be observed in multifractal spectra by the method of moments, because that method has a higher level of approximation. An advantage of the histogram method for multifractal spectra is the ability to show these processes. The method is used for the reason that it is interesting for examining influences of the system on data, such as the network parameters influence on the data traffic.

As stated in [4], the FS 3D format has better properties in the sense of variability than the MV 3D format, but

with frame aggregation (for a pair of consecutive frames, or for 16 consecutive frames - one GoP) video sequences show better performance in the variability for both of the formats, especially for MV 3D that becomes similar to FS. These results are confirmed in our study using multifractal analysis by the method of moments and further analyzed for burstiness that is particularly important for the data traffic. Video sequences with aggregated frames were analyzed by the multifractal spectrum using the histogram method. These results are shown in Figure 7. The dominant change in the spectra, for videos with aggregation, is in the location of the spectrum maximum $\alpha(f_{max})$ that represents complexity of the structure in the most frequent case. Both 3D formats, FS and CV, have a less complex structure, lower α in the most frequent case, for sequences with aggregation. A slight improvement is observed for aggregation of 2 frames, and higher for

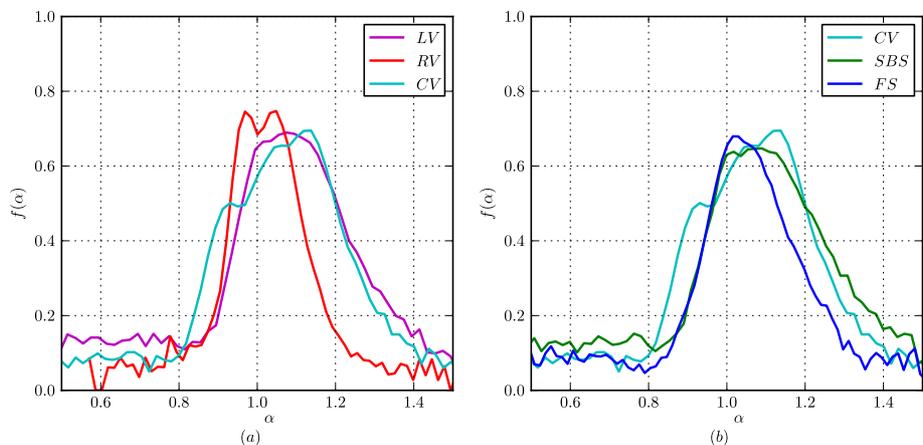
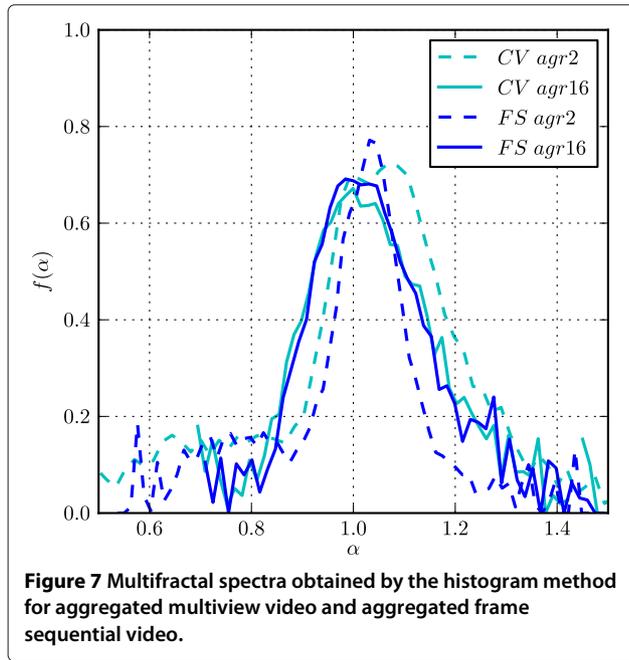


Figure 6 Multifractal spectra obtained by the histogram method. Multifractal spectra for **(a)** different views of multiview 3D video and **(b)** 3D video representation formats.



aggregation on the GoP level. Parameter f_{\max} of the FS and CV spectra is very similar.

Due to the complexity of the multifractal spectra obtained by the histogram method, for an easier comparison, some important points from the spectra are numerically represented in Table 2. From Table 2, it can be seen that a band of multifractal spectra B_m , defined as $B_m = \alpha_{\max} - \alpha_{\min}$, is the lowest for RV of the multiview 3D video, the view that utilizes inter-view prediction. Between 3D video representation formats, the lowest range of singularities is for the FS 3D video format. A band of multifractal spectra is wider for CV than for FS, but aggregation streaming approach gets CV multifractal characteristics closer to the values for the FS video.

Table 2 Comparison of multifractal spectra properties obtained by the histogram method;

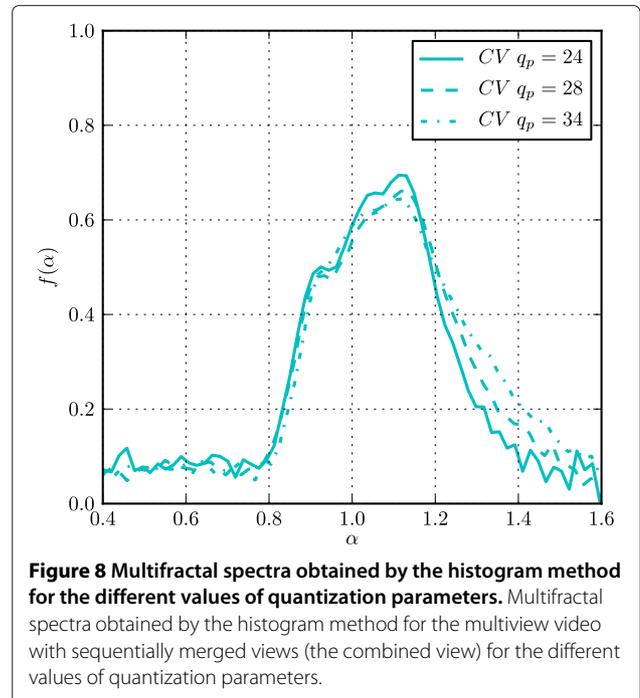
$$S_1 = \sum_{d=1}^{n_d} f(\alpha_{\min} + (d-1)\Delta\alpha), n_d = \frac{1-\alpha_{\min}}{\Delta\alpha};$$

$$S_2 = \sum_{d=1}^{k_d} f(\alpha_{\min} + (d-1)\Delta\alpha), k_d = \frac{\alpha_{\max}-\alpha_{\min}}{\Delta\alpha}$$

3D video	α_{\min}	α_{\max}	B_m	S_1	S_2
LV	0.2552	2.2338	1.9798	5.7506	16.7234
RV	0.5073	2.0830	1.5757	5.1741	13.2493
CV	0.1786	2.0400	1.8614	6.2659	17.2827
SBS	0.2532	2.0540	1.8007	6.1907	18.1982
FS	0.3536	1.7477	1.3941	6.0029	17.0272
CV-C 2	0.3491	2.1756	1.8265	5.9194	15.4685
CV-C 16	0.6945	2.1328	1.4382	6.5116	15.0897
FS-C 2	0.5372	1.8256	1.2884	5.7611	13.3861
FS-C 16	0.7091	2.2251	1.5159	5.1663	13.8045

Influence of the quantization parameter q_p is examined by the multifractal spectrum using the histogram method. It is concluded that a higher value of q_p leads to higher $\alpha(f_{\max})$ - a more complex structure, but also lower f_{\max} which means that other dimensions are more significant in the signal. As an example, the results for CV of the multiview 3D video format and $q_p = 28$ are shown in Figure 8.

Multifractal spectra obtained by the histogram method for isolated frame types: only I frames, only P frames, and only B frames for different 3D video representation formats are presented in Figure 9. The smallest changes in the multifractal spectra were found for the I frame type, given the similarities of coding for this frame type, regardless of the 3D video format. A maximum of multifractal spectra is the lowest for I frame types, followed by the maximum for B frame types and for the P frame types. A slowest decline of the multifractal spectrum around its maximum is for the I frame type that is a consequence of the structure with a large number of different dimensions with significant presence in the signal. Multifractal spectra for P and B frames have faster fall and fewer dimensions with large presence in the signal. For an isolated frame type, a sequence with only P frames has multifractal spectrum with smallest α_{\min} , which means highest burstiness. Similar conclusions were drawn from multifractal spectra by the method of moments. The connections between the multifractal spectra for CV, SBS, and FS formats for only I, only P, and only B frames are the same as in the case of sequences with all frame types.



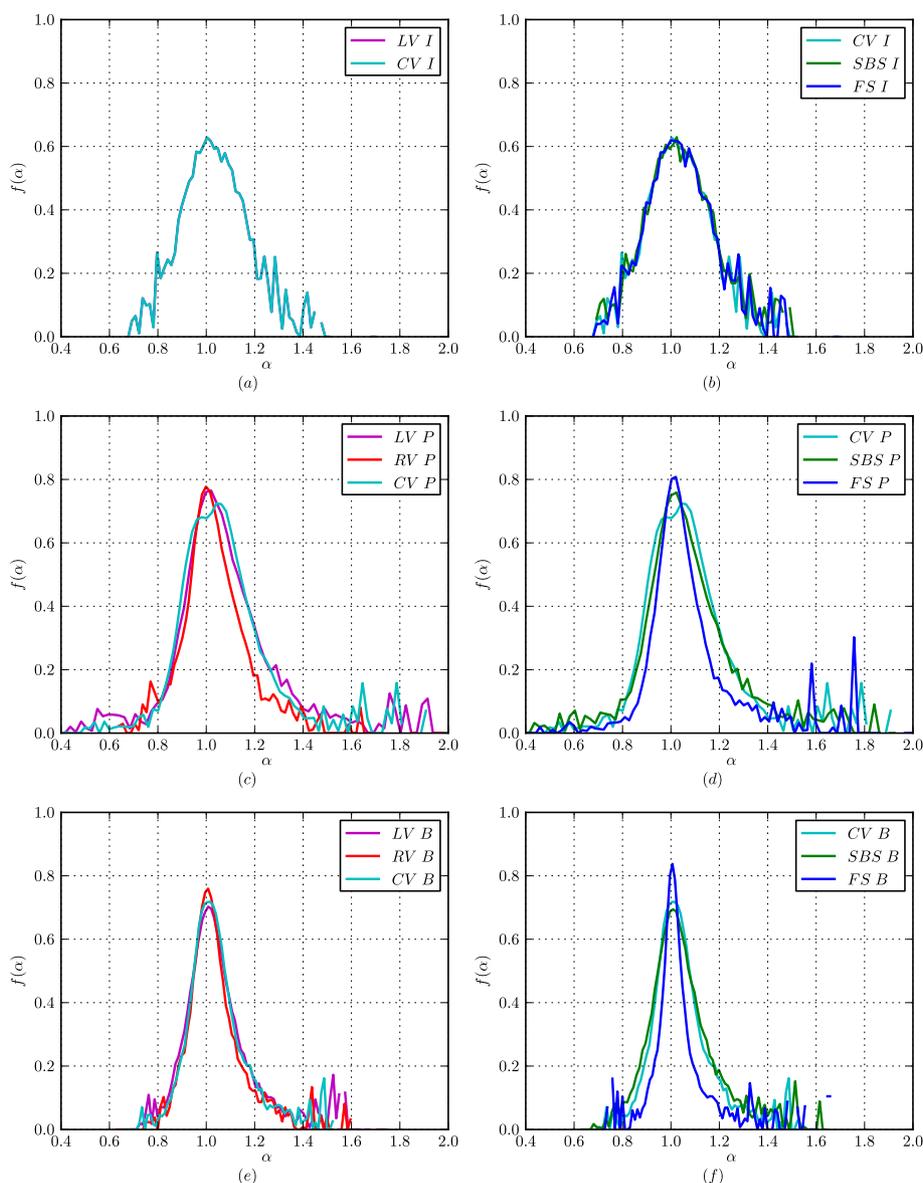


Figure 9 Multifractal spectra obtained by the histogram method for only one frame type isolated from 3D videos. Multifractal spectra for **(a)** only I frames of different views of the multiview 3D video, **(b)** only I frames from 3D videos in different representation formats, **(c)** only P frames of different views of the multiview 3D video, **(d)** only P frames from 3D videos in different representation formats, **(e)** only B frames of different views of the multiview 3D video, and **(f)** only B frames from 3D videos in different representation formats.

Inverse multifractal analysis

An advantage of the histogram method for obtaining multifractal spectrum, in addition to visibility of different processes in signal (mentioned earlier for multifractal spectrum of RV and CV), is the possibility of an inverse multifractal analysis. This means that it is possible to observe a spectrum, and for particular points of the spectrum (values of α and $f(\alpha)$), it is possible to find the exact data in the signal that correspond to these values. Inverse multifractal analysis is illustrated in Figure 10. It can be seen that for small values of α , that correspond to burstier

part of the signal, isolated data are the highest frame sizes, while for the high values of α , some of the smallest frames are extracted from the video.

Conclusions

We analyzed properties of 3D video representation formats: the MV video representation format with multiview video coding and the FS and the SBS formats coded with a conventional single-view video encoder. We determined multifractal properties by the method of moments and by the histogram method for three main 3D video

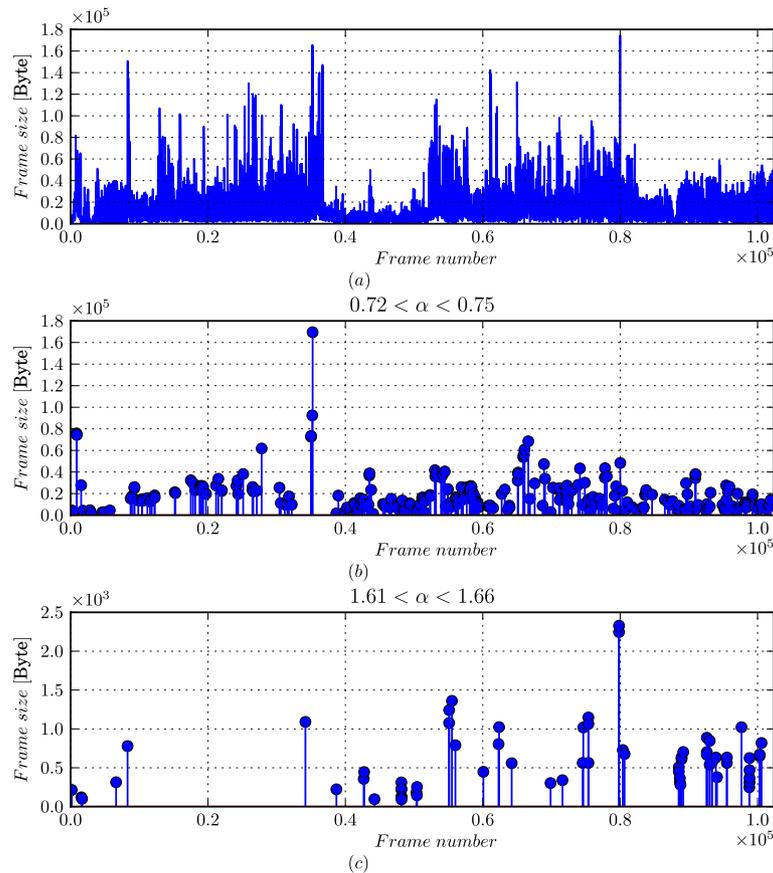


Figure 10 Illustration of the inverse multifractal analysis of the multiview 3D video. **(a)** Frame sizes (video traces) for the combined view of the multiview 3D video. **(b)** Extracted frames that have α in the range $0.72 < \alpha < 0.75$ and **(c)** that have α in the range $1.61 < \alpha < 1.66$.

representation formats with two views using long publicly available HD $1,920 \times 1,080$ resolution video frame-size traces.

We showed that 3D video formats are multifractal and can be modeled as such in traffic network models. In the paper, we present and compare the obtained multifractal spectra as a whole and isolate and compare important points of the spectrum - the most probable singularity of the spectrum that describes multifractal nature of the structure in the most probable case and the smallest value of singularity that is related with the highest burstiness (high bit variability) of the videos. Our results show that MV has the highest bitrate variability and the highest multifractal nature (for the most probable singularity in the spectrum), while the FS video format has the smallest values of these parameters. Obtained results for bitrate variability of 3D videos are compared with the values of a traditionally used statistical parameter for this purpose, the CoV, and showed good agreement.

Our analysis shows and compares multifractal properties of 3D videos with different quantization parameters. It was found that a video with a higher value of

quantization parameters (higher compression ratios) shows higher multifractal nature, as well as higher burstiness. Isolated intracoded (I) frames, predictive encoded (P) frames, and bi-directional (B) frames are analyzed. It is shown that I frames have very similar multifractal spectra regardless of the 3D video representation format. It was also shown that B frames, the smallest frames, have the narrowest multifractal spectrum. With the highest bitrate variability (the smallest value α_{\min}), P frames show rare prominent parts of the signal, while in comparison, I frames have higher α_{\min} but with the higher probability of this singularity.

Results presented in this paper can be beneficial for the traffic smoothing improvement and for the design of more efficient statistical multiplexing. Elementary smoothing techniques over video frames assume their aggregation [20]. The results of the multifractal analysis on the signals with performed smoothing technique based on aggregation of adjacent frames, for all types of the examined 3D video formats, are presented. The smoothing approach is applied for 2 frames (the number of views in the examined 3D video in all formats), as well as for 16 frames (the

number of frames in the group of pictures), showing that generally there is lower variability in the signal by using this approach. Particularly significant improvement in the values of multifractal characteristics was noticed for the MV video. These results can further be used for improving the smoothing techniques of 3D videos, in applications such as smoothing with prefetching [33,46], more precisely in the sense of estimating the bursty traffic, in the process of its management and control, in order to handle it first.

Multifractal parameters calculated in the analysis (multifractal spectra and generalized dimensions) can be used for creating improved multiplexing methods. In [47], fractal properties are used for creating an efficient multiplexing method. Statistical multiplexing methods that pay special attention on the type of frames for their improvement [36] can also potentially improve the performances, having in mind multifractal properties of different frame types that we provided in the paper.

Determined multifractal properties of 3D representation formats have possible application in statistical multiplexing, to develop methods for selection of optimal multiplexer parameters and/or better utilization of available network capacities. Also, the results can be used to analyze how introduction of 3D formats in the same multiplexer with 2D formats affects characteristics of the channel. A complete understanding of the multifractal properties will contribute to the analysis of the behavior of a 3D video signal in a statistical multiplexer, which is a subject of current research.

Competing interests

The authors declare that they have no competing interests.

Received: 1 March 2014 Accepted: 21 October 2014

Published: 3 November 2014

References

1. P Merkle, K Müller, T Wiegand, 3D video: acquisition, coding, and display. *IEEE Trans. Consum. Electron.* **56**(2), 946–950 (2010)
2. Y Chen, Y-K Wang, K Ugur, MM Hannuksela, J Lainema, M Gabbouj, The emerging MVC standard for 3D video services. *EURASIP J. Adv. Signal Process.* **2009**(786015), 1–13 (2009)
3. A Smolic, K Mueller, P Merkle, C Fehn, P Käuff, P Eisert, T Wiegand, in *Proceedings of the IEEE International Conference on Multimedia and Expo: 9-12 July 2006*. 3D video and free viewpoint video - technologies, applications and, MPEG standards (Toronto, 2006), pp. 2161–2164
4. A Pulipaka, P Seeling, M Reisslein, L Karam, Traffic and statistical multiplexing characterization of 3D video representation formats. *IEEE Trans. Broadcasting.* **59**(2), 382–389 (2013)
5. A Fernando, S Worrall, E Ekmekcioglu, *3DTV: Processing and Transmission of 3D Video Signals*. (John Wiley & Sons, Inc., UK, 2013)
6. A Vetro, W Matusik, H Pfister, J Xin, in *Proceedings of the Picture Coding Symposium, 15 Dec. 2004*. Coding approaches for end-to-end 3D TV systems, (2004)
7. A Vetro, T Wiegand, GJ Sullivan, Overview of the stereo and multiview video coding extensions of the H.264/MPEG-4 AVC standard. *Proc. IEEE.* **99**(4), 626–642 (2011)
8. K Müller, P Merkle, G Tech, T Wiegand, in *Proceedings of the 17th IEEE International Conference on Image Processing (ICIP): 26–29 Sept. 2010*. 3D video formats and coding methods (Hong Kong, 2010), pp. 2389–2392
9. E Stoykova, A Alatan, P Benzie, N Grammalidis, S Malassiotis, J Ostermann, S Piekh, V Sainov, C Theobalt, T Thevar, X Zabulis, 3D time-varying scene capture technologies - a survey. *IEEE Trans. Circuits Syst. Video Technol.* **17**(11), 1568–1586 (2007)
10. P Benzie, J Watson, P Surman, I Rakkolainen, K Hopf, H Urey, V Sainov, C von Kopylow, A survey of 3DTV displays: techniques and technologies. *IEEE Trans. Circuits Syst. Video Technol.* **17**(11), 1647–1658 (2007)
11. GB Akar, AM Tekalp, C Fehn, MR Civanlar, Transport methods in 3DTV - a survey. *IEEE Trans. Circuits Syst. Video Technol.* **17**(11), 1622–1630 (2007)
12. H Mohib, MR Swash, AH Sadka, in *Proceedings of International Conference on Communications, Signal Processing, and Their Applications: 12-14 Feb. 2013*. Multi-view video delivery over wireless networks using HTTP (Sharjah, 2013), pp. 1–5
13. T Schierl, S Narasimhan, Transport and storage systems for 3-D video using MPEG-2 systems, RTP, and ISO file format. *Proc. IEEE.* **99**(4), 671–683 (2011)
14. HH Yen, in *Proceedings of IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PacRim): 23-26 Aug. 2011*. Power-aware, bandwidth-aware and video-quality-aware cooperative routing algorithm for 3D video transmission in wireless networks (Victoria, BC, 2011), pp. 470–475
15. CG Gürlér, AM Tekalp, Peer-to-peer system design for adaptive 3D video streaming. *IEEE Commun. Mag.* **51**(5), 108–114 (2013)
16. Video Trace Library. <http://trace.eas.asu.edu>, Access date: 20 July 2013
17. P Seeling, M Reisslein, B Kulapala, Network performance evaluation using frame size and quality traces of single-layer and two-layer video: a tutorial. *IEEE Commun. Surv. Tutorials.* **6**(3), 58–78 (2004)
18. P Seeling, FHP Fitzek, M Reisslein, *Video Traces for Network Performance Evaluation*. (Springer, Dordrecht, 2007)
19. P Seeling, M Reisslein, Video transport evaluation with H.264 video traces. *IEEE Commun. Surv. Tutorials.* **14**(4), 1142–1165 (2012)
20. G Van Der Auwera, PT David, M Reisslein, *Video traffic analysis of H.264/AVC and extensions: single-layer statistics*. (Arizona State University, Technical report, 2007)
21. TV Lakshman, A Ortega, AR Reibman, VBR video: tradeoffs and potentials. *Proc. IEEE.* **86**(5), 952–973 (1998)
22. JSVM Reference Software. Obtained at cvs-dpserver:jvtuser@garcon.iert.rwth-aachen.de/cvs/jvtcheckoutsjvm, Access date: 8 Aug. 2014
23. JMVC Reference Software. Obtained at cvs-dpserver:jvtuser@garcon.iert.rwth-aachen.de/cvs/jvtcheckoutjmvc, Access date: 8 Aug. 2014
24. O Sheluhin, S Smolskiy, A Osin, *Self-Similar Processes in Telecommunications*. (New York, 2007)
25. RH Riedi, J Lévy Véhel, TCP traffic is multifractal: a numerical study, Research report 3129, Inria Rocquencourt (1997)
26. JW de Godoy Stênio, LL Ling, in *Proceedings of IEEE 27th International Conference on Advanced Information Networking and Applications (AINA): 25-28 March 2013*. A new binomial conservative multiplicative cascade approach for network traffic modeling (Barcelona, 2013), pp. 794–801
27. TD Dang, S Molnár, I Maricza, in *Global Telecommunications Conference, GLOBECOM IEEE: 17-21 Nov. 2002*. Capturing the complete multifractal characteristics of network traffic, (2002), pp. 2355–2359
28. A Evertsz, B Mandelbrot, in *Chaos and Fractals*, ed. by H Peitgen, H Jürgens, and P Andrews. Multifractal measures (Springer, New York, 1992), pp. 849–881
29. J Feder, *Fractals*. (Springer, New York, 1988)
30. P Murali, VMG Krishna, UB Desai, Modelling and control of broad band traffic using multiplicative multifractal cascades. *Sadhana, J. Indian Acad. Sci.* **27**(6), 699–723 (2002)
31. J Cosmas, J Loo, A Aggoun, E Tsekleves, in *Proceedings of IEEE Int. Symp. on Broadband Multimedia Systems and Broadcasting: 24-26 March 2010*. Matlab traffic and network flow model for planning impact of 3D applications on networks (Shanghai, 2010), pp. 1–7
32. N Manap, G Di Caterina, J Soraghan, in *Proceedings of IEEE 3DTV Conference: 4-6 May 2009*. Low cost multi-view video system for wireless channel (Potsdam, 2009), pp. 1–4
33. UC Devi, RK Kalle, S Kalyanaraman, Multi-tiered, burstiness-aware bandwidth estimation and scheduling for VBR video flows. *IEEE Trans. Netw. Serv. Manag.* **10**(1), 29–42 (2013)
34. W Feng, J Rexford, Performance evaluation of smoothing algorithms for transmitting prerecorded variable-bit-rate video. *IEEE Trans. Multimedia.* **1**(3), 302–313 (1999)

35. C-H Hsu, M Hefeeda, in *Proceedings of the 17th ACM International Conference on Multimedia*. On statistical multiplexing of variable-bit-rate video streams in mobile systems, (2009), pp. 411–420
36. G Van Der Auwera, M Reisslein, Implications of smoothing on statistical multiplexing of H.264/AVC and SVC video streams. *IEEE Trans. Broadcasting*. **55**(3), 541–558 (2009)
37. T Raghuvveera, K Easwarakumar, An efficient statistical multiplexing method for H.264 VBR video sources for improved traffic smoothing. *Int. J. Comput. Sci. Inf. Technol.* **2**(2), 51–62 (2010)
38. G Gürler, B Görkemli, G Saygili, AM Tekalp, Flexible transport of 3-D video over networks. *Proc. IEEE*. **99**(4), 694–707 (2011)
39. A Vetro, AM Tourapis, K Müller, T Chen, 3D-TV content storage and transmission. *IEEE Trans. Broadcasting*. **57**(2), 384–394 (2011)
40. H Peitgen, H Jürgens, D Saupe, *Chaos and Fractals*. (Springer, New York, 1992)
41. P Legrand, JL Véhel, in *Thinking in Patterns: Fractals and Related Phenomena in Nature*, ed. by M Novak. Signal and image processing with Fraclab (World Scientific, Singapore, 2003), pp. 321–322
42. JL Véhel, C Tricot, On various multifractal spectra. *Prog. Probability*. **57**(2004), 23–42 (2004)
43. I Reljin, A Samcovic, B Reljin, H.264/AVC video compressed traces: multifractal and fractal analysis. *EURASIP J. Adv. Signal Process.* **2006**(75217), 1–13 (2006)
44. A Chhabra, C Meneveau, V Jensen, K Sreenivasan, Direct determination of the $f(\alpha)$ singularity spectrum and its application to fully developed turbulence. *Phys. Rev. A*. **40**(9), 5284–5294 (1989)
45. SH Strogatz, *Nonlinear Dynamics and Chaos*. (Westview Press, Cambridge, Massachusetts, 2001)
46. S Oh, B Kulapala, AW Richa, M Reisslein, Continuous-time collaborative prefetching of continuous media. *IEEE Trans. Broadcasting*. **54**(1), 36–52 (2008)
47. Linawati, NP Sastra, in *IFIP International Conference on Wireless and Optical Communications Networks: 5-7 May 2008*. Statistical multiplexing strategies for self-similar traffic (Surabaya, 2008), pp. 1–5

doi:10.1186/1687-1499-2014-181

Cite this article as: Zekovic and Reljin: Multifractal analysis of 3D video representation formats. *EURASIP Journal on Wireless Communications and Networking* 2014 **2014**:181.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
