

Comparing Gamma and Weibull as Frame Size Distributions for High Efficient Video Coding

C. Perra¹

¹*Department of Electric and Electronic Engineering, University of Cagliari,
Piazza D'Armi, IT-09123 Cagliari, Italy
cperra@diee.unica.it*

Abstract—Digital video is one of the major traffic components in communication networks. Modelling the frame size of encoded video data is a preliminary step in the research and development of synthetic video data generators enabling a thorough analysis of video architecture systems that are often difficult to perform with real digital video data. In this paper, a statistical analysis of frame sizes of High Efficient Video Coding (HEVC) video generated at bit rates of interest for high quality Full HD video applications is performed. The selection of potential distributions for modelling the HEVC frame size distribution is based on the results from the modelling of H.264 frame size distribution. Experimental results show that the Gamma distribution has a better fit, to the HEVC frame size distribution, than the Weibull distribution.

Index Terms—Video modelling, video traffic generator, HEVC, Full HD.

I. INTRODUCTION

Video coding standards have mainly gone through the development of ITU-T and ISO/IEC standards. The ITU-T produced H.26x standards [1], ISO/IEC produced MPEG-1 [2] and MPEG-4 Visual [3], and the two organizations jointly produced the H.262/MPEG-2 Video [4], H.264/MPEG-4 Advanced Video Coding (AVC) [5], and High Efficiency Video Coding (HEVC) [6].

A video sequence is composed by group of frames (GOP). Each GOP is a sequence of frames encoded as: intra frame (I frame), if only spatial redundancy is exploited for the frame compression; prediction (P frame) from previous frames, if both temporal and spatial redundancy are exploited for the frame compression; bidirectional (B frame) prediction when the temporal prediction is computed from previous and future frames.

In most video encoders a bit rate control is employed for satisfying bitrate constraints given by the communication channels. Such rate control can target a constant bitrate (CBR) video coding or a variable bit rate (VBR) video coding. VBR is the preferred choice for applications requiring a constant quality service such as digital television broadcasting, DVDs, and Blue-ray disks.

Research and development of services related to digital

video often employ synthetic video traffic generation for analysing the behaviour and performance of a given video service with respect to the coding process, the rate control process, and the multiplexing and transmission process.

A video traffic model should be strictly related to the video coding standard that will be used for coding digital video, capturing the characteristic of the video sequence.

Bitrate, GOP structure (the sequence order of I, B, P frames in a GOP), frame size, frame size distribution, frame sizes dependences (short range dependencies, long range dependencies, autocorrelation), intra-GOP correlation, inter-GOP correlation [7] are some of the characteristics of major interest.

A traffic model for MPEG-4 and H.264/MPEG4 AVC video traces has been proposed in [7]. Inter-GOP and intra-GOP correlation in compressed VBR (variable bit rate) sequences have been addressed by incorporating wavelet-domain analysis into time-domain modelling. The proposed model has been developed at frame-size level, allowing for the analysis of loss ratio for each type of frames.

A statistical analysis of H.264 video frame size distribution has been presented in [8]. Frame sizes of H.264/MPEG4 AVC encoded video have been modelled with well-known statistical distributions. The main results have shown that the Gamma and Weibull distributions give appropriate statistical distributions of video frame sizes.

A video traffic modelling tool for simulation-based performance evaluation studies [9] consider different models for traffic correlations based on the M/G/ process, able to exhibit both short range dependence (SRD) and long range dependence (LRD).

A model for controlled VBR video traffic reflecting different properties of video traffic related to the content, encoder, and rate controller is proposed in [10]. The model is based on the interaction of video encoder and encoded bit stream by the means of a rate controller. The model parameters depend on the encoding parameters, on the rate control parameters, and on the properties of video content. This model exhibits into the synthetic generated video traffic both LRD and SRD dependences and also some properties related to the content and the encoding parameters.

Traffic modelling of VBR video is necessary for performance assessment of a network design and for the creation of synthetic loads to be used for benchmarking a network. A recent classification and survey of VBR video

Manuscript received January 26, 2014; accepted August 25, 2014.

The author gratefully acknowledges Sardinia Regional Government for the financial support (P.O.R. Sardegna F.S.E. Operational Programme of the Autonomous Region of Sardinia, European Social Fund 2007-2013 - Axis IV Human Resources, Objective I.3, Line of Activity I.3.1).

traffic models [11] provides a review of the state of the art and a classification and comparison among some representative video traffic model.

The objective of this paper is to analyse the statistical behaviour of frame sizes in HEVC video generated at bit rates of interest for high definition video applications to be used as a basis for simulation studies of digital television services.

The paper is organized as follows. Section II briefly reviews the video coding architecture. The statistical analysis is presented in Section III. Section IV concludes the paper.

II. HEVC VIDEO CODING

A digital video sequence is composed of video scenes. A video scene is composed of frames (pictures). Each frame is encoded with different algorithms.

The main frame types are named I, P, and B. I frames are encoded exploiting spatial redundancy for the compression.

The encoding of an I frame is self-contained meaning that no other information is required for the reconstruction (decoding) of that frame. P frames are encoded exploiting both spatial and temporal correlations.

The encoding of P frames requires the knowledge of previously (reference) encoded frames and the decoder must save all reference frames necessary for decoding the target P frame.

The encoding of B frames requires the knowledge of previous and future frames (which must be encoded in advance) and the decoder must save all past and future reference frames necessary for decoding the target B frame.

In general, in MPEG international standards a frame is partitioned into slices and the encoding is performed at slice level, hence the encoding is referred to I-slices, B-slices, P-slices.

Video encoding rate control is designed for satisfying requirement such as encoding at a given bitrate, encoding at given target quality, encoding at the highest possible quality

with a constraint of a target bitrate, encoding at a given bitrate with a constraint of a maximum buffer size.

In constant bit rate encoding, the encoders dynamically adjust encoding parameters in order to minimize the difference between the actual bitrate and the target bitrate. Most of the algorithms are based on the adaptation of the quantization parameters at one or more of the following video levels: scene, GOP, frame, slice, macroblock, block.

The encoded bit streams while satisfying the CBR constraint presents a highly variable video quality. CBR is not the preferred choice in digital video entertainment application.

VBR video encoding can be controlled or uncontrolled. In controlled encoding, the bit rate of the encoded bit stream is controlled by a VBR rate controller according to feedback signals from encoding results and coding complexity of video source.

The rate controller imposes some constraints on the degree of variability allowed in the bit rate. In uncontrolled encoding, the video pictures are encoded with an almost constant quantization (QP) to provide a relative constant quality for encoded video regardless of coding complexity of video source.

Novel service, the increased popularity of high-definition (HD) video, a possible future adoption of video format with resolutions higher than HD have fostered the development of video coding standard improving the capacity of current standards such as H.264/MPEG-4 AVC.

Network traffic caused by video applications for mobile devices, or for video-on-demand services, is very demanding for high quality/high resolution videos.

The recent High Efficiency Video Coding (HEVC) standard is an ITU-T and ISO/IEC joint project activity developed under a partnership named Joint Collaborative Team on Video Coding (JVT-VC).

HEVC is mainly focused into two key points: support higher video resolution and augment the use of parallel processing architectures.

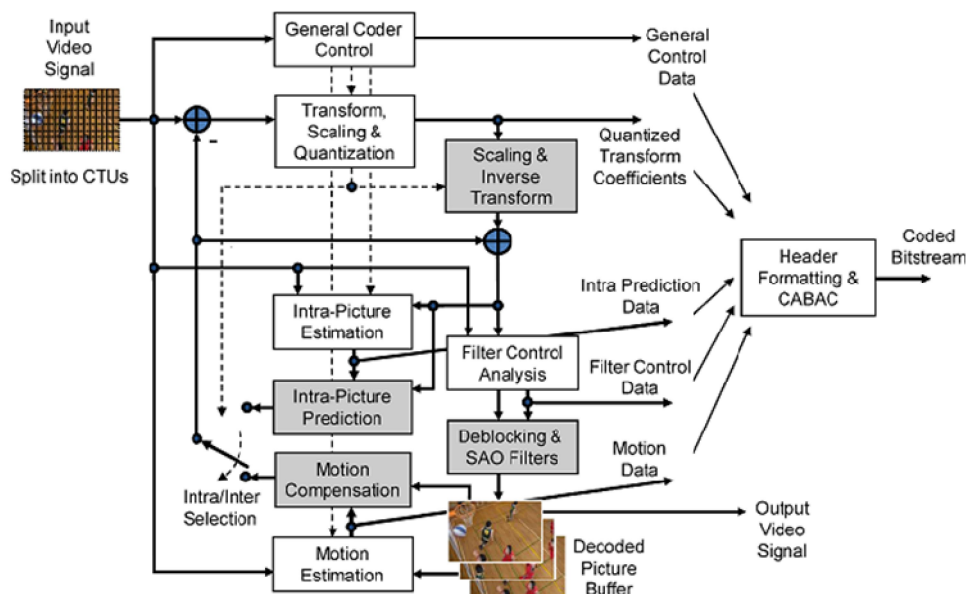


Fig. 1. HEVC Video encoder [12].

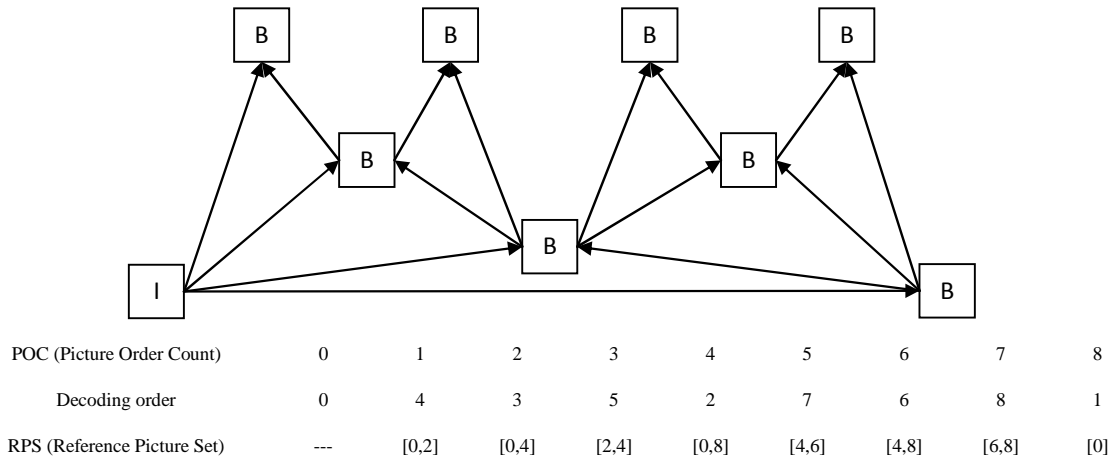


Fig. 2. Temporal prediction structure and the POC values, decoding order, and RPS content for each picture.

As in the previews ITU-T and ISO/IEC, the HEVC standardizes only the bitstream structure and syntax of the decoder. HEVC video coding layer makes use of the same hybrid approach employed by previous video standard since H.261.

Figure 1 shows a diagram of a hybrid encoder for creating a bitstream compliant with HEVC.

A coding algorithm for producing an HEVC bitstream operates as follows. Each image is decomposed into block units by a coding tree unit (CTU) process. Each block is transmitted to the encoder. The first image of a video sequence is encoded exploiting intra-picture prediction. The other images are encoded exploiting inter-picture prediction defining motion vector (MV) highlighting temporal redundancies between frames.

The signal difference between the original picture and the corresponding intra- or inter- prediction is transformed by the mean of linear spatial transform.

Transformed coefficients are then scaled and quantized. The quantized coefficients are then encoded (entropy coding) by a context-adaptive binary arithmetic coding (CABAC) process. Finally, the encoded data is transmitted together with the other prediction information.

III. STATISTICAL ANALYSIS

Test video sequences have been chosen from [13], where several contributions of video content that are most relevant for determining the effectiveness of consumer video processing applications and quality measurement algorithms are provided.

The selection of test video sequences has been oriented by the following constraints: YUV 4:2:0 sampling, full HD 1920 × 1080 resolution, 60 Hz frame rate.

The cyclic GOP structure used repeatedly throughout the test video sequence is shown in Fig. 2. This coding structure is of size 8 (8 B slices).

This prediction structure is characterized by a structural delay of eight pictures and can provide an improved coding efficiency compared to IBBP coding [14].

Video sequences have been encoded at the following bitrates: 1, 2, 4, 6, 8, and 10 Mbps.

Figure 3 shows the average PSNR at different bitrates. PSNR (Y) is the average for all the frames, PSNR (Y, only I frame) is the average for the I frames, PSNR (Y, only B

frame) is the average for the B frames. I frames have a higher PSNR than B frames.

Statistical analysis has been done for each prediction layer. Referring to Fig. 2: layer 0 (L0) is composed of frames with POC 0 and 8; layer 1 (L1) is composed of frames with POC 4; layer 2 (L2) is composed of frames with picture order code (POC) 2 and 6; layer 3 (L3) is composed of frames with POC 1, 3, 5, and 7.

For each layer, frame size histograms have been analysed in order to empirically determine a set of statistical distribution to be used as a starting point for fitting to the data sets.

An example of the frame size, for 60 consecutive HEVC encoded frames, is shown in Fig. 4.

Gamma and Weibull distribution have been selected for the fitting analysis, on the basis of the results obtained when modelling H.264 frame sizes [8].

The parameters of the Gamma and Weibull distributions have been computed using the maximum likelihood estimation (MLE) method.

Figure 5 shows an example of a L3 frame sizes histogram, for a video sequence encoded by HEVC at 6 Mbps.

The goodness of the fit has been tested using the Pearson's chi-square test. The value of the test is

$$t^2 = \sum_{i=1}^n \frac{(o_i - e_i)^2}{e_i}, \quad (1)$$

where t^2 is the Pearson's cumulative test statistic, o_i is an observed frequency, e_i is an expected frequency, n is the size of the histogram domain.

Table I shows the t^2 results for each layer for the dataset encoded at 6 Mbps. Numerical results for the other experimental bitrates are similar to the results in Table I. Gamma distribution has shown better fit than Weibull distribution, for each layer and for each bitrate.

TABLE I. CHI-SQUARE TEST RESULTS FOR EACH LAYER AT 6 MBPS.

	Gamma	Weibull
L0	191	261
L1	943	1400
L2	253	547
L3	40	191

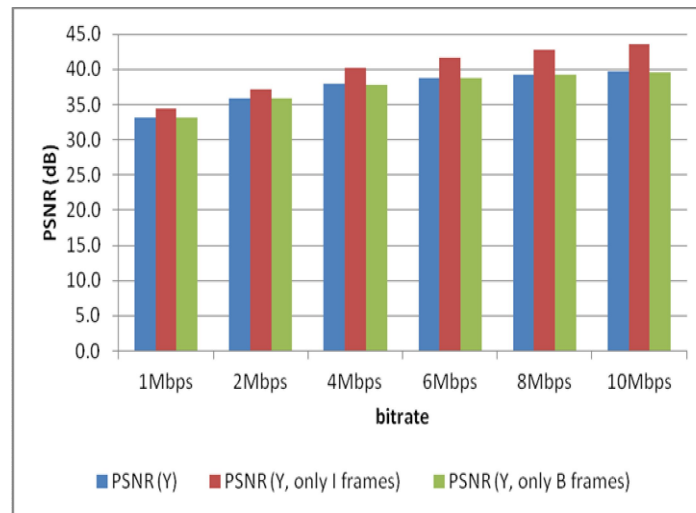


Fig. 3. Average PSNR at different bitrates. PSNR (Y) is the average for all the frames, PSNR (Y, only I frame) is the average for the I frames, PSNR (Y, only B frame) is the average for the B frames. I frames have a higher PSNR than B frames.

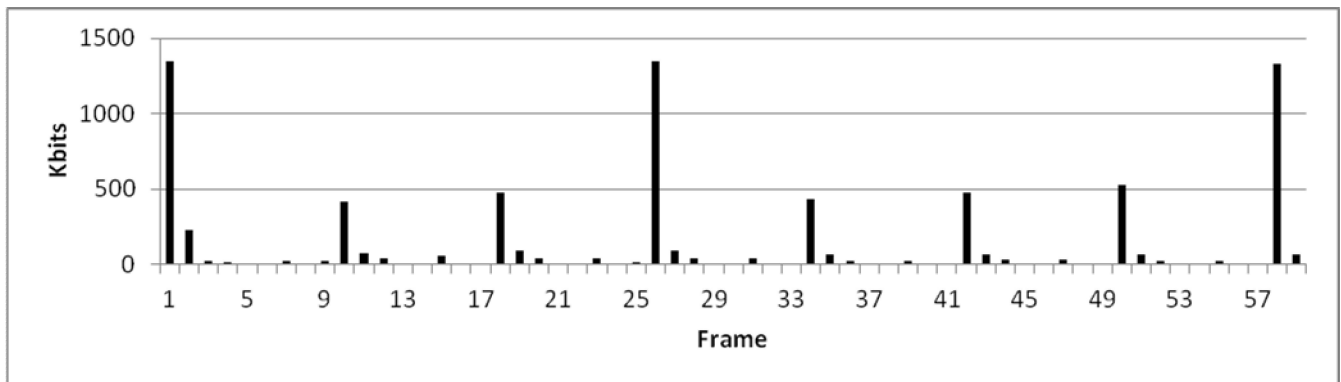


Fig. 4. Example of one second of HEVC encoded frame size at 6 Mbps of a Full HD video (1920 × 1028), 60 Hz, YUV 4:2:0.

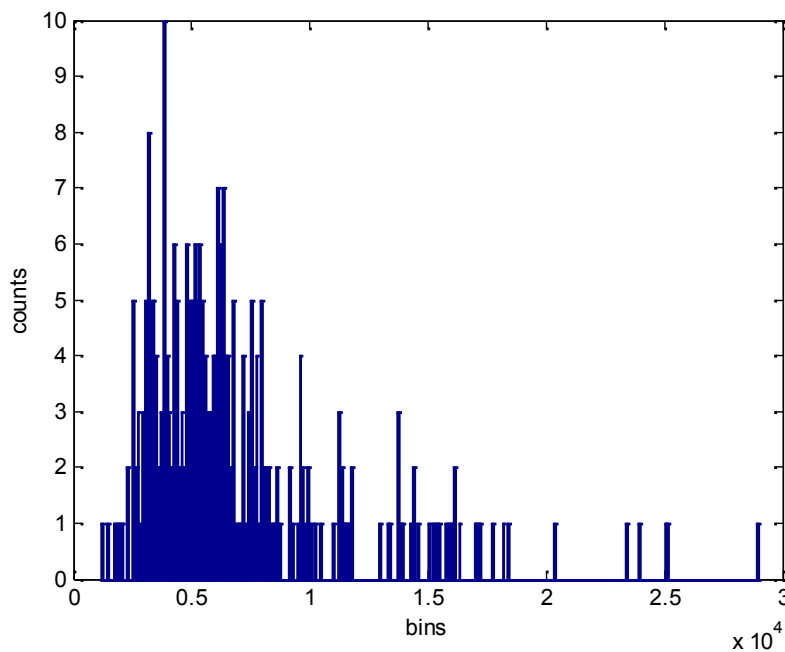


Fig. 5. Example of histogram for L3 frame sizes of test video sequences encoded by HEVC at 6 Mbps.

IV. CONCLUSIONS

Defining the statistical properties of a VBR video traffic and relying upon a video traffic model able to produce simulated traffic data having the desired statistical properties

enables for an accurate study of the performance of communication networks.

In this paper, the statistical properties of frame sizes of HD video sequences encoded by the HEVC standard have been analysed and the Gamma and Weibull distributions

have been evaluated and compared as fitting distributions. HEVC videos, selected from a dataset of sports video content-types, have been generated at bit rates of interest for high definition video applications. For a GOP structure of interest in digital video applications the analysis has been conducted at different prediction layers. Experimental results have shown that the Gamma distribution is more appropriate than the Weibull distribution for modelling the statistical distribution of HEVC encoded frame.

REFERENCES

- [1] Video Coding for Low Bit Rate Communication, ITU-T Rec. H.263, 1995.
- [2] Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to About 1.5 Mbit/s—Part 2: Video, ISO/IEC 11172-2 (MPEG-1), ISO/IEC JTC 1, 1993.
- [3] Coding of Audio-Visual Objects—Part 2: Visual, ISO/IEC 14496-2 (MPEG-4 Visual version 1), ISO/IEC JTC 1, 1999.
- [4] Generic Coding of Moving Pictures and Associated Audio Information— Part 2: Video, ITU-T Rec. H.262 and ISO/IEC 13818-2 (MPEG 2 Video), ITU-T and ISO/IEC JTC 1, 1994.
- [5] Advanced Video Coding for Generic Audio-Visual Services, ITU-T Rec. H.264 and ISO/IEC 14496-10 (AVC), ITU-T and ISO/IEC JTC 1, 2003.
- [6] G. J. Sullivan, J. Ohm, W. Han, T. Wiegand, A. High, E. Video, C. Hevc, “Overview of the high efficiency video coding (HEVC) standard,” *IEEE Trans. Circuits And Systems For Video Technology*, vol. 22, no. 12, pp. 1649–1668, 2012. [Online]. Available: <http://dx.doi.org/10.1109/TCSVT.2012.2221191>
- [7] M. D. M. Dai, Y. Z. Y. Zhang, D. Loguinov, “A unified traffic model for MPEG-4 and H. 264 video traces”, *IEEE Trans. Multimedia*, vol. 11, no. 5, pp. 1010–1023, 2009. [Online]. Available: <http://dx.doi.org/10.1109/TMM.2009.2021802>
- [8] K. Salah, F. Al-Haidari, M. H. Omar, A. Chaudhry, “Statistical analysis of H.264 video frame size distribution”, *IET Communications*, vol. 5, no. 14, pp. 1978–1986, 2011. [Online]. Available: <http://dx.doi.org/10.1049/iet-com.2010.0868>
- [9] M. E. Sousa-Vieira, “A video traffic modelling tool for simulation-based performance evaluation studies,” in *4th Int. Conf. Signal Processing and Communication Systems*, 2010.
- [10] M. Rezaei, I. Bouazizi, M. Gabbouj, “A model for controlled VBR video traffic”, in *IEEE Int. Conf. Signal Processing and Communications (ICSPC)*, 2007.
- [11] S. Tanwir, H. Perros, “A survey of VBR video traffic models,” *IEEE Communications Surveys & Tutorials*, no. 99, pp. 1–25, 2013.
- [12] G. J. Sullivan, J. R. Ohm, Woo-Jin Han, T. Wiegand, “Overview of the high efficiency video coding (HEVC) standard”, 2012.
- [13] The Consumer Digital Library, 2010. [Online]. Available: <http://www.cdvl.org>
- [14] H. Schwarz, G. Sullivan, T. Tan, T. Wiegand, “Comparison of the coding efficiency of video coding standards—including high efficiency video coding (HEVC),” *IEEE Trans. Circuits And Systems For Video Technology*, vol. 22, no. 12, pp. 1669–1684, 2012. [Online]. Available: <http://dx.doi.org/10.1109/TCSVT.2012.2221192>