GENERATING ANTIPERSISTENT VBR VIDEO TRAFFIC

[1]Mehdi Rezaei, [2]Imed Bouazizi, [3]Moncef Gabbouj

[1,3]Tampere University of Technology, [2]Nokia Research Center, Tampere, Finland

ABSTRACT

A novel model for antipersistent variable bit rate (VBR) video traffic is proposed. Antipersistent VBR video bit streams are a subset of VBR video bit streams in which the bit rate is controlled by a rate controller when the video is encoded. Statistical properties of antipersistent VBR video bit streams are very different from those of uncontrolled VBR video bit streams that are encoded with a constant quantization parameter. Discriminating controlled and uncontrolled VBR, an accurate model is proposed for controlled VBR video traffic. The proposed model is built based on the interaction of video encoder and video bit stream that is controlled by the rate controller. The model parameters depend on the encoding and rate control parameters and also depend on the properties of video content. The proposed model not only captures the long-range dependent (LRD) and short-range dependent (SRD) properties of video traffic, but also it can include some properties related to the content and the encoding parameters into the synthetic generated video traffic. This is valuable when generated traffics are used for simulations in which the network behavior is studied based on the content properties and encoding parameters. The proposed model has been validated successfully by several measures.

***Index Terms*—** Antipersistent, Buffer, Controlled, Hurst Exponent, Model, Smoothed, Traffic, Variable Bit Rate, Video

1. INTRODUCTION

Video traffics are going to consume the largest part of bandwidth in communication networks. From the video quality points of view VBR video is preferred to constant bit rate in many applications such as video streaming. Accurate modeling of VBR video traffic is important in proper allocating of resources such as bandwidth and delay in communication networks. Moreover, it can be used as a tool in many research fields that need to generate huge synthetic video traffic for computer simulations. For example, when the performance of a statistical multiplexer is studied, a huge volume of traffics with different properties is required to feed into the multiplexer. Using real traffic with such huge volume is almost impossible for a simulation task while synthetic traffic can be easily generated by a traffic model. A good model predicts or provides a desired metric or a set of desired metrics for the modeled data similar to original data. For example if the packet loss probability is the desired metric, then a good model produces traffics that precisely provide this metric in simulations.

Generally, the performance of a communication network in terms of delay, data drop rate, bandwidth usage and number of services is very dependent on the statistical properties of traffics in the network. For example the autocorrelation function of service traffics has a major impact on the performance of communication networks.

VBR video traffic was found to exhibit self-similar characteristics [1]. In mathematics, a self-similar object is exactly or approximately similar to a part of itself, e.g., the whole has the same shape as one or more of the parts. Self-similarity is a typical property of fractals. A fractal is a rough or fragmented geometric shape that can be subdivided in parts, each of which is (at least approximately) a reduced-size copy of the whole.

The main feature of self-similar processes is that they exhibit long range dependence (LRD), that is, their autocorrelation function $r(k)$ decays less than exponentially fast, and is non-summable, i.e. $r(k) \sim k^{-\beta}$, as $k \to \infty$, for $0 < \beta \le 1$. The quantity $H = 1 - \beta / 2$ is called *Hurst parameter* or *Hurst exponent*. The Hurst exponent was originally developed in hydrology [2]. It shows whether the data is a purely *random walk* or has underlying trends. The Hurst exponent is related to the *fractal dimension* and it is a measure of the smoothness of fractal time-series based on the asymptotic behaviour of the *rescaled range* of the process. The Hurst exponent is defined as:

$$H := \frac{\log(R/S)}{\log(T)} \tag{1}$$

where T is the duration of the sample of data, and R/S is the corresponding value of rescaled range. S denotes the standard deviation of the sample data and R stands for the difference between the max and min of accumulated deviation from the mean value during the time period T. If $H = 0.5$, the behavior of the time-series is similar to a random walk and samples are uncorrelated. When $H > 0.5$, the time series covers more distance than a random walk. In this case the process is namely persistent and samples are positive correlated. This means that if the time-series is increasing, it is more probable that it will continue to increase. When $H < 0.5$, the time-series covers less distance than a random walk. In this case the process is namely antipersistent and samples are negative correlated. This means that if the time-series is increasing, it is more probable that then it will decrease, and vice versa.

In communication networks the Horst exponent of traffic reflects the buffering requirements for the transmission of traffic. Considering the definition of R in (1), in fact it is equal to minimum buffering space for perfect transmission of the data during time period T and by a channel with a bandwidth equal to the average of bit rate. Therefore, the overall performance of a communication network in terms of buffering delay, data drop rate and bandwidth usage depends on the statistical properties of traffic such as self similarity and smoothness. Hence fore, the traffic models that are used in design of communication networks should capture these statistical properties.

Traditional stochastic video traffic models that only exhibit short range dependence (SRD), i.e. have an autocorrelation function that decays exponentially or faster, are not accurate to be used in video communication network design. There are a number of stochastic models which exhibit the self-similar property. *Fractional Gaussian Noise* (FGN) and *Fractional Autoregressive Integrated Moving Average* (F-ARIMA) are two used models. A survey of statistical source models for VBR video is presented in [3].

Several approaches have been used for VBR video traffic modelling [3]. A group of proposed models try to capture the frame size distribution [4], [5]. In this group the model is validated by comparing the histogram or distribution of frame sizes. A common method used for comparing distributions is Quantile-Quantile (QQ) plot. In another approach the attempt is to capture the LRD and SRD properties[1], [6], [7]. The autocorrelation function (ACF) is a widely used metric related to LRD for validating of the models. Data loss rate and buffering delay are real practical metrics which have been used for validating of some proposed models [5].

In this paper, a new model for controlled VBR video traffic is presented. The term of VBR is applicable for a wide range of video bit streams. However the controlled VBR video bit streams are a subset of VBR bit streams. A VBR video can be provided by encoding a video sequence using a constant quantization parameter (QP) to provide a high visual quality for encoded video or the QP can be controlled by a rate controller to smooth the variations in the bit rate of bit stream. From the modelling point of view, the statistical properties of a controlled VBR video differ from those of uncontrolled VBR video. While uncontrolled VBR bit streams usually have persistent behaviours with a Hurst exponent $H > 0.5$, the controlled VBR bit streams, depending on the degree of control, move toward the antipersistent with $H < 0.5$. Antipersistent controlled VBR video is preferred in many video communication applications to decrease the end-to-end delay and to increase the bandwidth usage. To the best of our knowledge, all available video traffic models are designed for persistent VBR traffic or at least they have been evaluated by persistent video bit streams. They do not discriminate the persistent and antipersistent video traffics. However, the statistical properties of persistent and antipersistent video traffics are very different. The proposed model in this paper is targeted for antipersistent VBR video traffic. The proposed model is very novel from two aspects: modelling approach and application area.

The proposed model in this paper simulates the interaction between the video encoder and video bit stream for modelling. The interaction of encoder and bit stream is controlled by a rate controller. The rate controller provides some application dependent constraints on the bit stream. The statistical properties of video traffic depend on the video content properties, encoding method, and rate control algorithm. Understanding the relationships between the content, encoding process, and statistical properties of video traffic helps to generate real or synthetic video traffics with various properties proper for a real application or a simulation task. Accordingly, the proposed model can generate various traffics according to the content e.g. sport, movie, news, etc. Also, it can produce video traffics according to the encoding parameters such as bit rate, frame rate, picture size and etc. Moreover it can produce video traffics according to rate control parameters such as buffering delay. These features are beneficial in simulations tasks in which the network behaviours are studied from different points of view.

This paper is organized as follows. The proposed model is presented in Section II. Implementation details of the model for generating video traffic are presented in Section III. Some simulation results are provided in Section IV. The paper is closed with conclusions in Section V.

2. PROPOSED VIDEO TRAFFIC MODEL

A video sequence includes several scenes and each scene includes a number of video frames from different types such as I, P and B frames. Each type of frames in each video scene has a probability distribution function (PDF). Different distributions such as Lognormal and Gamma have been proposed for the size of video frames in uncontrolled VBR in previous research works. Although in controlled VBR the PDF of video frame size can be very different from the uncontrolled case, in the proposed model, a Gamma distribution has been used for each frame type in each video scene. Note that at the sequence level, each frame type can have a PDF very different from the scene level because; the PDF of video scenes are combined together at the sequence level. In the proposed model, the PDF of each frame type is assumed to be a Multi-Gamma distribution at the sequence level. The Gamma distribution has been selected for scene cut level because it is fit enough to the practical results and also it simplifies the implementation of some new concepts into the model. Assuming a Gamma PDF for each frame type at the scene level, the PDF for the whole video frames can be expressed as:

$$P = \sum_{s=1}^{S} \left[A_s . P_{Is}(x, k_{Is}, \theta_{Is}) + B_s . P_{Ps}(x, k_{Ps}, \theta_{Ps}) + \ldots \right.$$

$$\left. \ldots + C_s . P_{Bs}(x, k_{Bs}, \theta_{Bs}) \right], \qquad (2)$$

where A_s, B_s and C_s are constant coefficient as normalization factors. P_{is} is the probability density function of the Gamma distribution for the frame type i in video scene s and it can be expressed in terms of the Gamma function as:

$$P_{is}(x, k_{is}, \theta_{is}) = x^{k_{is}-1} \frac{e^{-x/\theta_{is}}}{\theta_{is}^{\ k_{is}} \Gamma(k_{is})} \qquad for \quad x > 0, \qquad (3)$$

where $k_{is} > 0$ is the shape parameter and $\theta_{is} > 0$ is the scale parameter of Gamma distribution.

The rate of scene change or the number of scene cuts in a video sequence depends on the video content. The video scenes can be very short or very long. The video contents of consequent scenes can be similar or very different from each other. When prototype information about the scene cuts is not available, a uniform PDF for the scene cuts along the sequence is assumed by the proposed model.

From the rate control point of view various frame types have different degrees of coding complexity. This means for a constant level of distortion they need different amounts of bit budget to be encoded. A VBR video rate controller tries to provide a constant quality as much as possible at least over a video scene. Therefore, the following conditions are uses as parts of the model:

$$\mu_{Is} = \mu_{Ps} X_{Ps} = \mu_{Bs} X_{Bs}, \qquad (4)$$

where μ_{is} denotes the mean frame size of type i in video scene s. Parameters X_{Ps} and X_{Bs} show the relative complexity of I-pictures to P and B pictures respectively in video scene s. This is a known concept that is used in some control algorithms. The values of relative complexities not only are affected by the properties of video content such as motion activities but also affected by encoding parameters such as bit rate and frame rate. Moreover, they are affected by the rate control parameters such as smoothing buffer size. For the Gamma PDF in which $Mean = k\theta$, the formula (4) yields

$$\mu_{Is} = k_{Is}\theta_{Is} = k_{Ps}\theta_{Ps} X_{Ps} = k_{Bs}\theta_{Bs} X_{Bs}. \qquad (5)$$

While the rate controller tries to reach an average target bit rate of B for the video sequence in a frame rate of F, for a GOP (Group of Picture) in a video scene the average frame size can be estimated as:

$$\bar{x} = \frac{\mu_{Is} + N_P \mu_{Ps} + N_B \mu_{Bs}}{1 + N_P + N_B} = \frac{B}{F}. \qquad (6)$$

In practice an integer number of GOPs may not fit to the video scenes and video sequence.

A buffer constraint is considered for the controlled VBR video that is implemented by the rate controller. From the buffer constraint point of view, considering a simple buffering model at the decoder side, the following conditions should be met

$$0 \le O_B(n) \le S_B, \quad n = 1, 2, \ldots, N \qquad (7)$$

$$O_B(n) = O_B(n-1) + \frac{B}{F} - x_n, \quad n = 1, 2, \ldots, N \qquad (8)$$

where S_B denotes the size of buffer and $O_B(n)$ shows the buffer occupancy before removing nth picture (x_n) from the buffer.

3. GENERATING VIDEO TRAFFIC

In this section we explain how the proposed model can be used to generate synthetic video traffic similar to a real prototype or just based on description of some properties for the video content and the rate control parameters. To generate a synthetic video traffic by the proposed model a number of parameter should be determined. The main parameters include total number of frames in the video sequence (N), structure of GOP (N_P, N_B), video scenes and their parameters (s, k_{ij}, θ_{ij}), average bit rate (B), frame rate (F) and buffer size (S_B). When a prototypal video traffic is generated, the all parameters above can be extracted easily from the prototype and video traffic can be generated by (3). Here we discussed a more general case in which no prototype bit stream is used. In our previous paper we used some assumption to simplify the algorithm and to minimize the user defined parameters in this case [8]. In this paper some modifications to the algorithm is proposed that is extracted based on new simulation results. To produce video traffic without a prototype, the main parameters such as $N, N_P, N_B, B, F,$ and S_B are set directly by the user and the scene parameters (s, k_{ij}, θ_{ij}) are defined by an algorithm as follows. A uniform PDF is used to define whether a picture is a scene cut or not. A sequence of N numbers corresponding to the video frames is produced by a uniform PDF. If the number is greater than a threshold then the corresponding frame in the sequence is a scene cut frame. The threshold can be determined by the user or it can be read from a look up table based on a description of content by the user. The look up table is built based on experimental results on a number of real video sequences. For example a sport video sequence has more scene cuts than a news video sequence and therefore has a lower threshold. As another option a PDF can be used for the length of video scene to determine the scene cuts. Note that some encoders utilizes a scene cut detector and inserts an I-picture at scene boundaries and some encoders do not care about the picture type at scene cuts. In the both case, similar statistics in terms of size can be assumed for the scene cut pictures. Therefore, in this algorithm the scene cut pictures are treated as I-pictures.

The values of relative complexities X_{Ps}, X_{Bs} are the main parameters that should be defined for each scene cut. Collected statistics from different video sequences which are encoded with similar encoding parameters show that $X_{Ps}-1$ and $X_{Bs}-1$ have distributions close to Gamma PDF over the scenes. Therefore, the user is required to define only 4 parameters for 2 Gamma PDFs $P_{X_P}(X_P-1,k_{X_P},\theta_{X_P})$ and $P_{X_B}(X_B-1,k_{X_B},\theta_{X_B})$ that are: $k_{X_P},\theta_{X_P},k_{X_B},\theta_{X_B}$. Using P_{X_P} and P_{X_B} two sets of numbers, with S members in each, corresponding to the relative complexities X_{Ps} and X_{Bs} are generated.

To define the shape parameter and the scale parameter of the Gamma distributions for the frames in video scenes we proceed as follow. Solving the equations in (4) and (6) for each scene gives the numerical values for μ_{Is}, μ_{Ps}, and μ_{Bs}. Still, there are three equations with six unknown parameters that should be solved for each scene as:

$$\mu_{Is}=k_{Is}\theta_{Is}, \quad \mu_{Ps}=k_{Ps}\theta_{Ps}, \quad \mu_{Bs}=k_{Bs}\theta_{Bs}. \tag{9}$$

To solve these equations we try to find a relationship between the shape parameters and the relative complexities in the scenes. A video scene with low relative complexities usually has more motion activities and thereafter a larger variance in the picture sizes. Experimental results also show that the video scenes with smaller relative complexities usually have larger shape parameters. Moreover, collected statistics show that a Gamma PDF also can be assumed for the distribution of shape parameters over the scenes. However, the Gamma PDFs of shape parameters differ from the Gamma PDFs of relative complexities. The user defines the shape and scale parameters of 3 Gamma PDFs $P_{k_I}(k_I,k_{k_I},\theta_{k_I})$, $P_{k_P}(k_P,k_{k_P},\theta_{k_P})$, and $P_{k_B}(k_B,k_{k_B},\theta_{k_B})$ corresponding to the three picture types. These parameters depend on the content and encoding parameters. They can be defined by the user or it can be read from a look up that is built based on experimental results. Using P_{k_I},P_{k_P}, and P_{k_B}, 3 sets of numbers, with S members in each, are generated. Each member represents the shape parameter of a Gamma PDF $P_{is}(x,k_{is},\theta_{is})$ in (2). The generated shape parameters are allocated to the scenes according to their relative complexities. Simply, that the larger shape parameters are allocated to the scene with smaller relative complexities and vice versa. Only, the relative complexity of P-picture can be used for allocation of whole shape parameters. Now, the scale parameters of the Gamma distributions of different frames in each scene can be computed by

$$\theta_{is}=\mu_{is}/k_{is} \tag{10}$$

Now, the whole required parameters for generating video frames are available. Using provided parameters for the Gamma PDFs of frames in scenes (P_{is}), a synthetic video bit stream is generated according to (2).

The bit rate of generated bit stream is almost controlled. To confirm the buffering constraint on the bit stream, a buffer simulation is run. Considering a virtual buffer with a desired size and an initial buffering period, the buffer occupancy for the synthetic video bit stream is computed as (8). If the buffer has overflow or underflow in some points, the size of video frames in an interval close to the points are scaled to prevent buffer overflow and underflow. Now, a synthetic VBR video bit stream with a buffering constraint has been generated. Generated bit stream is packetized based on video frame. Depending on the desired transmission protocol, the video frames can be splitted and packetized to smaller units.

The whole traffic generating algorithm is summarized below:

1. Define the desired encoding parameters including: number of frames, bit rate, frame rate, GOP structure.
2. Define the parameters of Gamma PDFs P_{X_P} and P_{X_B}. Typical values for k_{X_P},θ_{X_P} are 7 and 0.5, respectively.
3. Define the parameters of Gamma PDFs P_{k_I}, P_{k_P}, and P_{k_B}. Typical values for $k_{k_I},\theta_{k_I},k_{k_P},\theta_{k_P}$ are 3, 2.8, 2, 8 respectively.
4. Determine the scene boundaries using a uniform PDF and a threshold to control the number of scene cuts. As another option, a PDF for the length of scenes can be used.
5. Using parameters defined in part 2, generate the relative complexities X_{Ps}, X_{Bs} for the scenes.
6. Using (4), (6) and parameters defined in part 1, calculate the values of $\mu_{Is},\mu_{Ps},\mu_{Bs}$ for each scene.
7. Using PDFs P_{k_I}, P_{k_P}, and P_{k_B} defined in part 3, generate 3 sets corresponding to the shape parameters of PDF P in (2) i.e. k_{Is},k_{Ps},k_{Bs}.
8. Allocate the shape parameters generated in part 7 to the scenes according to their relative complexities. The smaller shape parameter is allocated to the scenes with larger X_{Ps} and vice versa.
9. Using (10) compute the values of $\theta_{Is},\theta_{Ps},\theta_{Bs}$ for each scene.
10. For each scene generate the video frames using corresponding Gamma PDF P_{is} with parameters provided in parts 7, 9.
11. Using (8) and considering a desired buffer size, perform a buffer simulation on the provided bit stream. If there is an under flow or over flow in the buffer, scale the size of a number

(typically $L = S_B F / 2B$) of frames in an interval close to underflow or overflow point. The size of buffer determines the degree of allowed variations in the bit stream that is relevant to the Hurst exponent.

4. SIMULATION RESULTS

To evaluate the proposed video traffic model experimentally, we selected a set of known video sequences including *Foreman*, *Carphone, Silent, New York,* and *Football* sequences. We repeated and concatenated each of these sequences to provide longer sequences (900 frames) and then the provided sequences were concatenated again to make a longer video sequence. Provided video sequence has several different scenes that is suitable for evaluating of the model. The video sequence was encoded by the Nokia H.264 encoder published at [9] for the bit rate of 300 kb/s, frame rate of 30 f/s, and the buffering delay of 0.4 s to produce a prototype video bit stream. The model parameters were extracted based on the prototype bit stream and a synthetic sequence was generated by the proposed model. The prototype and synthetic generated traffics were compared by several measures including histogram, autocorrelation function (ACF) and Hurst exponent. Simulation results are shown in Fig. 1 to Fig. 9. The size of frames in the original and modeled bit streams are shown in Fig. 1. As shown the synthetic bit stream is very similar to the prototype bit stream. The frame size histograms of the bit streams are depicted in Fig. 2. Small difference between two histograms can be seen especially for the large size pictures because; the number of samples is relatively small. The autocorrelation functions of the two sequences are plotted in Fig. 3. Also the ACF of P-frames in two bit streams are depicted in Fig. 4. The QQ plots for all frames and for P-frames are shown in Fig. 5. The two Pox diagrams plotted in Fig. 6 have been used to compute the Hurst exponent for the two bit streams. Two values of 0.243 and 0.250 were obtained for the Hurst exponent of original and synthetic bit streams respectively. The Hurst exponent of synthetic traffic can be controlled by the buffer size in the model. The model has provided buffering properties for the modeled bit stream exactly similar to the prototype (0.4 seconds). As results show the synthetic video traffic provided by the proposed model is very similar to the prototype traffic from different aspects. The proposed model is applicable for a wide range of VBR video communication applications in which the bit rate of encoded bit streams are controlled. The model parameters should be adjusted according to the application.

5. CONCLUSION

Discriminating controlled and uncontrolled variable bit rate, a novel model for controlled variable bit rate video traffic is proposed. The proposed model not only captures the long-range dependent and short-range dependent properties of video traffic, but also it can include different properties related to the content, encoder, and rate controller in the video traffic. The model parameters can be extracted from a prototype bit stream or they can be defined directly by user. Only few parameters are defined by the user and other parameters used in the model are generated by an algorithm. The proposed model was validated by several methods. Simulation results show a very accurate performance for the model.

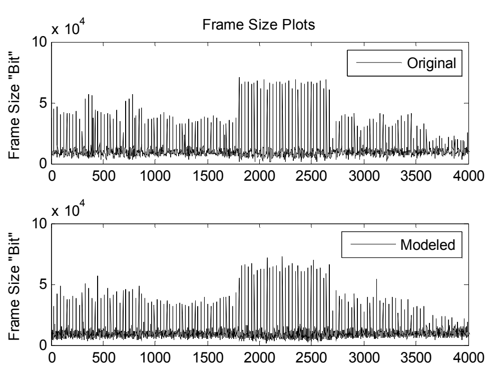

Fig. 1. Frame size of original and modeled bit streams

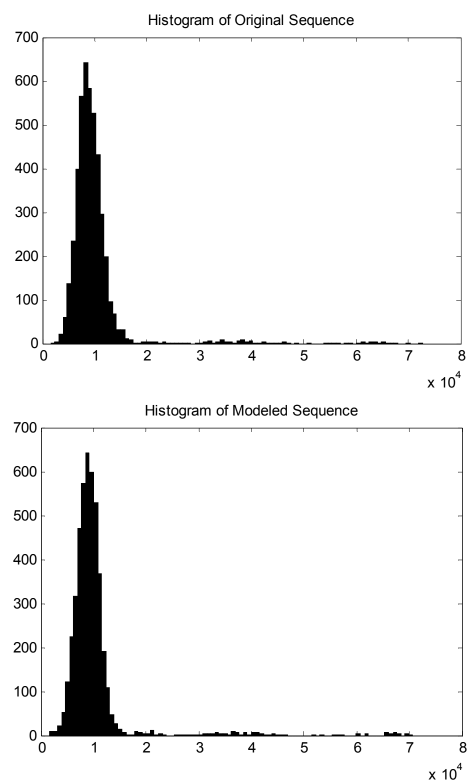

Fig. 2. Histograms of original and modeled bit streams

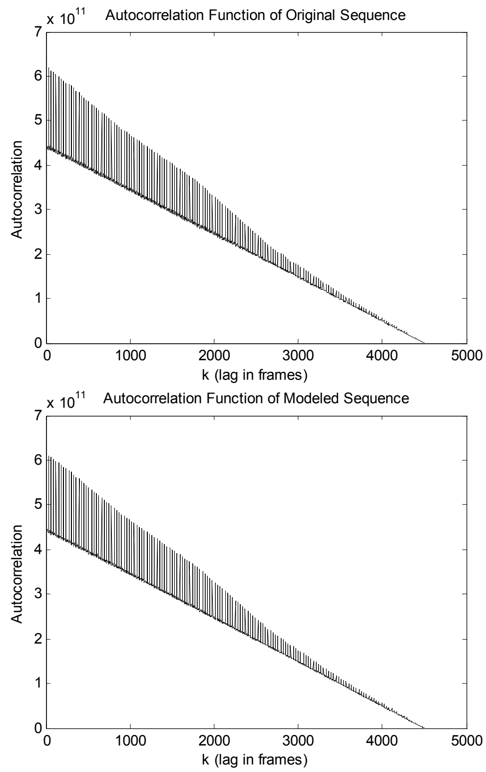

Fig. 3. ACF of original and modeled bit streams for all frames

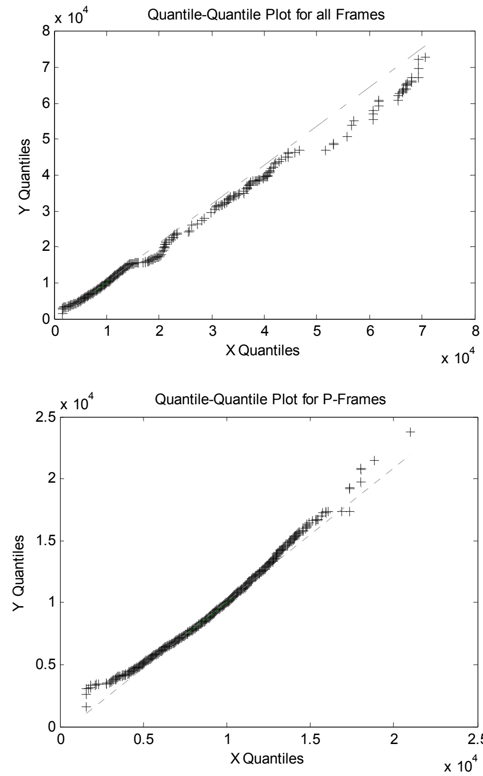

Fig. 5. QQ plots for all frames and for P-frames

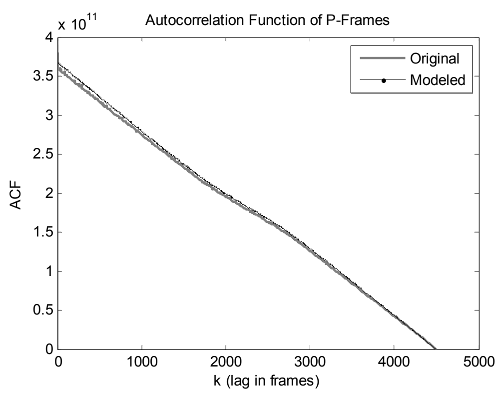

Fig. 4. ACF of original and modeled bit streams for P-Frames

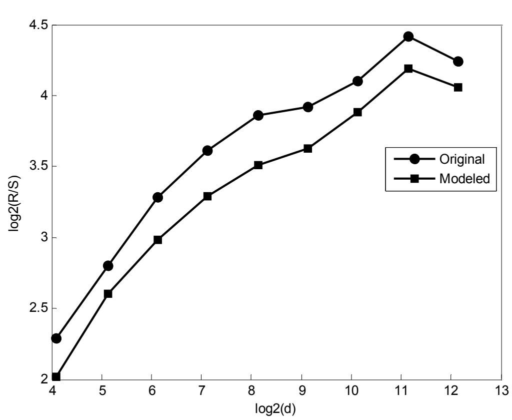

Fig. 6. Pox diagram of R/S for original and modeled data

6. REFERENCES

[1] J. Beran, R. Sherman, M.S. Taqqu, W. Willinger, "Long-range dependence in variable-bit-rate video traffic," *IEEE Trans. Communications*, Vol. 43, No. 234, pp. 1566-1579, 1995.

[2] H. E. Hurst, R. P. Black, Y. M. Simaika, Long-term storage: an experimental study. Constable, London, 1965.

[3] M. R. Izquierdo and D. S. Reeves, "A survey of statistical source models for variable-bit-rate compressed video," *Multimedia Systems*, Vol. 7, No. 3, pp. 199-213, July 1999.

[4] M. Kruns, S. K. Tripathi, "On the characterization of VBR MPEG streams," *ACM SIGMETRICS*, Vol. 25, 1997.

[5] U. K. Sarkar, S. Ramakrishnan, and D. Sarkar, "Modeling full-length video using Markov-modulated Gamma-based framework," *IEEE/ACM Trans. Networking*, Vol. 11, No. 4, Aug. 2003.

[6] D. Liu, E.I. Sara, W. Sun, "Nested auto-regressive processes for MPEG-encoded video traffic modeling," *IEEE Trans. Circuits and Systems for Video Technology*, Vol. 11, No. 2, pp. 169-183, Feb. 2001.

[7] M. Dai, D. Loguinov, H. Radha, "A hybrid wavelet framework for modeling VBR video traffic," *IEEE Int. Conf. Image Processing (ICIP)*, 2004.

[8] M. Rezaei, I. Bouazizi, M. Gabbouj, "A model for controlled VBR video traffic," *IEEE Int. Con. on Signal Processing and Communications (ICSPC 2007)*, Dubai, United Arab Emirates (UAE), Nov. 2007.

[9] ftp://standards.polycom.com/IMTC_Media_Coding_AG/, Sep. 2005.