A MODLE FOR CONTROLLED 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 controlled variable bit rate (VBR) video traffic is proposed. The proposed model is built based on the interaction of video coder and video bit stream that is controlled by a 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 the synthetic traffic is used for the simulation of network behavior based on content and coder properties. The proposed model has been validated successfully by different methods.

***Index Terms*—** Model, Traffic, Video, Variable Bit Rate

1. INTRODUCTION

Video bit streams 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 (CBR) 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 a huge volume of synthetic video traffic for computer simulations. 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.

Several approaches have been used for video traffic modelling [1]. A group of proposed models try to capture the frame size distribution [2], [3]. In this group, the model is validated by comparing the histograms or distributions 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 [4], [5], [6]. The autocorrelation function (ACF) is a widely used metric related to LRD for validating of the models. It can be argued that these metrics are not the final real metrics which are required in network resource allocation task and related simulations. Data loss rate and delay in leaky bucket (buffer) simulation are real practical metrics which have been used for the validating of some models [3].

In this paper, a new model for controlled VBR video traffic is presented that simulates the interaction between video coder and video bit stream for modelling. Interaction of coder 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 help to generate real or synthetic video traffics with various properties proper for a real application or a simulation task. 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 coder properties 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 to study the network behaviours 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. Section IV includes some simulation results. The paper is closed with conclusions in Section V.

2. PROPOSED 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 frame in each video scene has a probability distribution function (PDF) that can be different from others. 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. The PDFs of scene level are combined together at the sequence level. Therefore, at the sequence level any PDF can be constructed. 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 level because it is fit enough to the practical results and also it simplifies the

1-4244-1236-6/07/$25.00 © 2007 IEEE

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:

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

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:

$$P_{is}(x, k_{is}, \theta_{is}) = x^{k_{is}\;1} \frac{e^{\;x/\theta_{is}}}{\theta_{is}^{\;k_{is}} \Gamma(k_{is})} \quad for \quad x > 0, \quad (2)$$

where $k_{is} > 0$ is the shape parameter and $\theta_{is} > 0$ is the scale parameter of the 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 former 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 a part of the model:

$$\mu_{Is} = \mu_{Ps} X_{Ps} = \mu_{Bs} X_{Bs}, \quad (3)$$

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 by some control algorithms. The values of relative complexities are affected by the properties of video content such as motion activities and also they are 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 (3) yields

$$\mu_{Is} = k_{Is}\theta_{Is} = k_{Ps}\theta_{Ps} X_{Ps} = k_{Bs}\theta_{Bs} X_{Bs}. \quad (4)$$

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. Hence, the average frame size can be estimated as:

$$\bar{x} = \frac{\mu_{Is} + N_P \mu_{Ps} + N_B \mu_{Bs}}{1 + N_P + N_B} = \frac{B}{F}. \quad (5)$$

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

In many applications 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 decoder buffering model the following condition should be met

$$0 \le O_B(n) \le S_B, \quad n = 1, 2, ..., N \quad (6)$$

$$O_B(n) = O_B(n-1) + \frac{B}{F} - x_n, \quad n = 1, 2, ..., N \quad (7)$$

where S_B denotes the size of buffer and $O_B(n)$ shows the buffer occupancy before removing n^{th} 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 bit stream by the proposed model a number of parameters 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 going to be generated, the all parameters above can be extracted easily from the prototype and video traffic can be generated by (1). Here we discussed a more general case in which no prototype bit stream is used. In this case 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 the algorithm. 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. 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.

To define the shape parameter and the scale parameter of the Gamma distributions $P_{is}(x, k_{is}, \theta_{is})$ in video scenes, we proceed as follow. First, a max and a min value are defined for the relative complexities X_{Ps} and X_{Bs} over the whole video scenes (i.e. $X_P^{\max}, X_P^{\min}, X_B^{\max}, X_B^{\min}$) by the user. Then, two sets with s members corresponding to the relative complexities X_{Ps} and X_{Bs} are produced by a uniform PDF between the max and min values of relative complexities. Solving the three equations in (3) and (5) for each scene gives the numerical values for μ_{Is}, μ_{Ps} and μ_{Bs}. Still, there are three equations with six parameters that should be solved for each scene as:

$$\mu_{Is} = k_{Is}\theta_{Is}, \quad \mu_{Ps} = k_{Ps}\theta_{Ps}, \quad \mu_{Bs} = k_{Bs}\theta_{Bs}. \quad (8)$$

To solve these equations we try to find a relationship between the shape parameters and the relative complexities. Experimental results show that the video scenes with very large relative complexities have small shape parameters and video scenes with large shape parameters have small relative complexities and in the middle range there is no preference. To simplify the problem above, just a max and a min value for the shape parameter of each frame type i.e. $k_{Is}^{\max}, k_{Is}^{\min}, k_{Ps}^{\max}, k_{Ps}^{\min}, k_{Bs}^{\max}, k_{Ms}^{\min}$, are defined by user. Then, using a uniform PDF between the min and max value, a number of shape parameters corresponding to the number of scenes are generated for each frame type. Now, generated shape parameters are allocated to the scenes according to their relative complexities such that a percentage of the large shape parameters are allocated to the scenes that have small X_{Ps}. Also a percentage of small shape parameters are allocated to the scenes that have large X_{Ps}. The remained shape parameters are allocated without any preference. The max and min values of shape parameters depend on the video content and encoding parameters. For example a sport video sequence with high motion activities has smaller relative complexities and thereafter smaller shape parameters than a video sequence with low motion activities. The scale parameters of the Gamma distribution for each scene can be computed by

$$\theta_{ls} = \mu_{ls} / k_{ls} \tag{9}$$

Now the whole required parameters for producing video frames are available. Provided parameters are used by (1) to generate the video frame sizes. To impose the buffer constraint (6), considering an initial buffering period the buffer occupancy for the provided video sequence is computed as (7). The condition (6) should be met for any interval, including L frames, over the bit stream. Considering a typical value $L = S_B F / 2B$, if the constraint (6) is not met for an interval, the size of video frames in the interval are scaled to prevent buffer overflow and underflow. Now the synthetic video frames are generated.

4. SIMULATION RESULTS

To evaluate the proposed video traffic model we selected a set of known video sequences including *Foreman, Carphone, Silent, New York* and *Football*. 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 Nokia H.264 encoder 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. The frame size histograms of the bit streams are depicted in Fig. 2. The ACF of the two sequences are plotted in Fig. 3. 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 Pox diagrams plotted in Fig. 6 have been used to compute the Hurst-Exponent of the two bit streams. Two values of 0.221 and 0.218 were obtained for the Hurst-Exponent of original and synthetic bit streams respectively. Note that the Hurst-Exponent is smaller than 0.5 in both streams. This means that the bit streams are Anti-Persistent and this is result of control on the bit streams. The Hurst-Exponent of synthetic traffic can be controlled by the buffer size in the model. The model has provided buffering properties for the

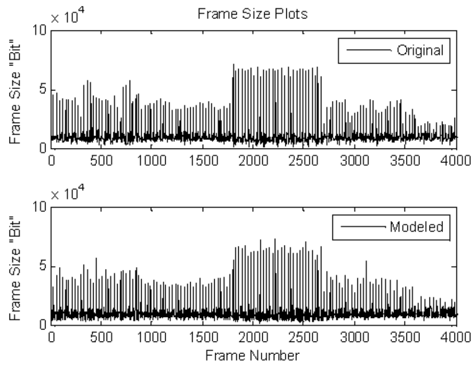

Fig. 1. Frame size of original and modeled bit streams

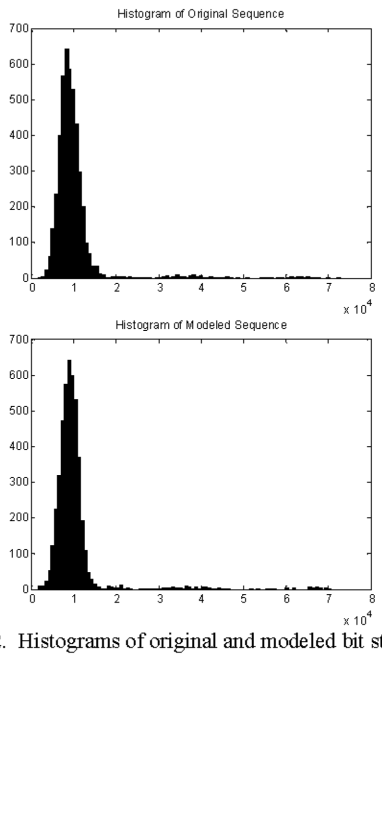

Fig. 2. Histograms of original and modeled bit streams

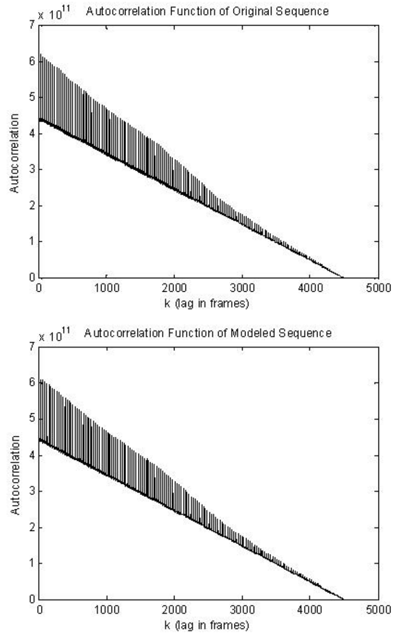

Fig. 3. ACF of original and modeled bit streams for all frames

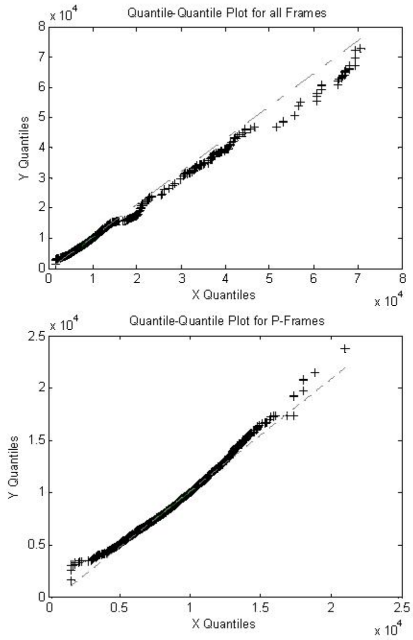

Fig. 5. QQ plots for all frames and for P-frames

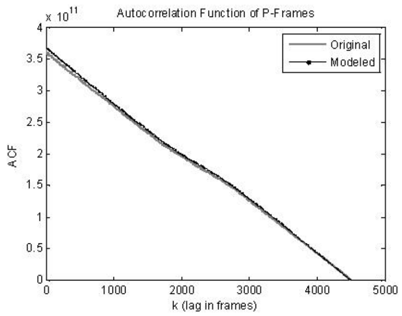

Fig. 4. ACF of original and modeled bit streams for P-Frames

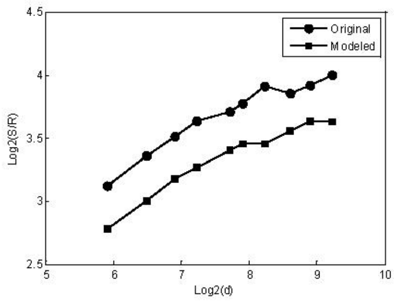

Fig. 6. Pox diagrams for original and modeled bit streams

modeled bit stream exactly similar to the prototype (0.4 seconds). As results show the 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 applications in terms of bit rate and delay. The model parameters should be adjusted according to the application.

5. CONCLUSION

A model for controlled variable bit rate video traffic is proposed. The proposed model can reflect different properties of video traffic related to the content, encoder, and rate controller. The model parameters can be extracted from a prototype or they can be defined directly by user. When user defines the parameters, only few parameters are set by user and many other parameters are generated by an algorithm. The proposed model was validated by several methods. Simulation results show a very accurate performance for the model.

6. REFERENCES

[1] 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.

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

[3] 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.

[4] 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.

[5] 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.

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