

# THE TALE OF A SIMPLE ACCURATE MPEG VIDEO TRAFFIC MODEL

H. Liu, N. Ansari, and Y. Q. Shi

New Jersey Center for Wireless Telecommunications  
Department of Electrical and Computer Engineering  
New Jersey Institute of Technology  
University Heights, Newark, NJ 07102, USA

## ABSTRACT

This paper traces the development/evolution of three of our recently proposed MPEG video traffic models, that can capture the statistical properties of MPEG video data. The basic ideas behind these models are to decompose an MPEG compressed video sequence into several parts according to motion/scene complexity or data structure. Each part is described with a self-similar process. These different self-similar processes are then combined to form the respective models. In addition, Beta distribution is used to characterize the marginal cumulative distribution (CDF) of the self-similar processes. Comparison among the three models shows that the latest model (called the simple models) is the most practical one in terms of accuracy and complexity. Simulations based on a real MPEG compressed movie sequence of *Star Wars* have demonstrated that the simple model can capture the ACF and the marginal CDF very closely.

## 1. INTRODUCTION

The trend to transmit video over network, especially over ATM, is emerging. Traffic models are important to network design, performance evaluation, bandwidth allocation algorithm design, and bit-rate control. It was observed, however, that traditional models fall short in describing the video traffic because video traffic is strongly autocorrelated and bursty [11]. To accurately model video traffic, autocorrelations among data should be taken into consideration. A considerable amount of effort on video modeling has been reported that include: Markov Modulated Rate Process (MMRP) [12], Discrete Auto-Regressive Process (DAR(1)) [3]; Fluid Models [9], Markov-Renewal-Modulated TES Models [10], Long Range Dependency (LRD) models or Self-Similar models [1],  $M/G/\infty$  input process models [5], and GBAR Model [4]. Most of the work dealing with video traffic appeared in the literature is based on a short period of video sequences or video conference sequences, which contain little drastic motions. It is therefore far from enough to model general video traffic, say, movie sequences containing all kinds of motions.

In this paper, we analyze and compare three recently developed MPEG video traffic models that can capture the LRD characteristics of video ACF [6], [7], [8]. The basic

ideas behind these models are to decompose an MPEG compressed video sequence into several parts according to motion/scene complexity or data structure; each part described by a self-similar process. These different self-similar processes are then combined in respective fashion to form the models. In addition, Beta distribution is used to characterize the marginal cumulative distribution (CDF) of the self-similar processes. Comparison among three models leads to the observation that the simple MPEG video traffic model is preferred in terms of accuracy and simplicity. Simulations on a real MPEG compressed movie sequence of *Star Wars* have demonstrated that our new simple model can capture the LRD of ACF and the marginal CDF very well. This paper is organized as follows. In Section 2, empirical data and the ACF are described. Section 3 discusses the decomposition of data according to motion/scene complexity or data structure. Modeling of each part and combination of the decomposed parts are discussed in Section 4. Using Beta distribution to model CDF of the video traffic is presented in Section 5. Comparison and discussion are made in Section 6.

## 2. EMPIRICAL DATA AND ACF

The empirical data used here was MPEG coded data of *Star Wars*<sup>1</sup>. The source contains materials ranging from low complexity/motion scenes to those with high and very high actions. The data file consists of 174,136 integers, whose values are frame sizes (bits per frame). The movie length is approximately 2 hours at 24 frames per second. The original video was captured as 408 lines by 508 pels, and then converted to 240×352 (Luminance - Y), and 120×176 (Chrominance - U and V). Motion estimation techniques were used to compress data volume. The frames were organized as follows: IBBPBBPBBPBB IBBPBB... , i.e., 12 frames in a Group of Pictures (GOP). I frames are those which use intra frame coding method (without motion estimation), P frames are those which use inter frame coding technique (with motion estimation), and B frames are predicted using both forward and backward prediction. Every frame was partitioned into blocks of 8×8 pixels. These data blocks were transformed using DCT. After DCT transformation,

<sup>1</sup>The MPEG coded data were the courtesy of M.W. Garrett of Bellcore and M. Vetterli of UC Berkeley.

coefficients were quantized and Huffman coded. Run length coding was further used to reduce bit rate.

The ACF of frame size of MPEG coded *Star Wars* is shown in Fig. 1, and it is quite different from the ACF of frame size of JPEG coded movies *Star Wars* (see Fig. 2). The ACF of MPEG coded data fluctuates around an envelope, reflecting the fact that, after the use of motion estimation techniques, the dependency between frames is reduced. This characteristic should be taken into consideration in modeling MPEG coded video sequences.

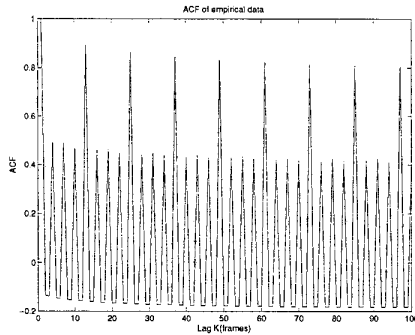


Figure 1: ACF of MPEG compressed video *Star Wars*

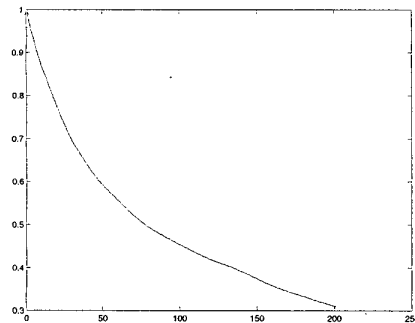


Figure 2: ACF of JPEG compressed video *Star Wars*

### 3. DECOMPOSITION OF MPEG DATA

The fluctuation of the ACF as shown in Fig. 1 has convinced us that such fluctuation can hardly be captured by a single random process, and has further led us to the intuitive belief that the data should be decomposed into several parts, each captured by a random process. This section describes the intuition that has resulted in the three proposed decompositions.

#### 3.1. Decomposition according to Motion/Scene Complexity

With the conjecture that the fluctuation was caused by motion/scene complexity, we proposed to divide the traffic data into three different parts—inactive part, active part, and the most active part (authors in [12] also pointed out that a video bit rate process has three main components: a slowly changing component, a more quickly changing component, and an impulsive component). Suppose  $f(i)$  is the number of bits in the  $i$ th frame. The video traffic can be classified as follows

1. If  $f(i+1)/f(i) > T, i = 2, 3, \dots$ , then  $f(i+1)$  belongs to the non-inactive part; otherwise,  $f(i+1)$  belongs to the inactive part, where  $T$  is a threshold value.
2. Similarly, the non-inactive part can be classified into the active and most active part.

Taking these three data sets as three different random processes, we can calculate their ACFs.

#### 3.2. Decomposition According to MPEG Data Structure (I)

Although the model based on the decomposition introduced above can model each part of the video traffic very well, it cannot capture the ACF of the whole sequence very well. This slight discrepancy highly depends on how one defines motion/scene complexity and has thus inspired us to decompose the MPEG data according to the MPEG data structure.

Specifically, in the second proposed model, we decompose the MPEG traffic into 10 sub-sequences  $X_I, X_P, X_{B_1}, X_{B_2}, \dots$ , and  $X_{B_8}$ .  $X_I$  consists of all I frames,  $X_P$  consists of all P frames, the first B frames in all GOPs constitute  $X_{B_1}$ , the second B frames in all GOPs constitute  $X_{B_2}$ , and so on.

#### 3.3. Decomposition according to MPEG Data Structure (II)

By observing that the B-frames have similar properties in terms of coding mechanism, we combine all the B-frames in the previous decomposition into one subsequence, resulting in three sub-sequences,  $X_I, X_P, X_B$ . As before,  $X_I$  consists of all I frames,  $X_P$  consists of all P frames, but, now,  $X_B$  consists of all B frames.

### 4. MODELING EACH PART AND COMBINING PARTS TO OBTAIN MPEG CODED TRAFFIC MODELS

To obtain a model that can catch the ACFs of MPEG data, we model each part by a self-similar process and then combine these processes in an appropriate fashion. In this section, the utilization of the self-similar processes is justified. Three different ways in combining, leading to three different models, are described.

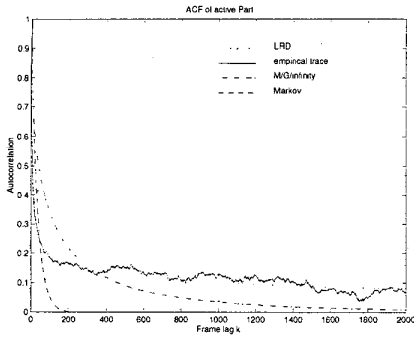


Figure 3: ACF of the active part of *Star Wars* (Self-similar process is denoted by LRD)

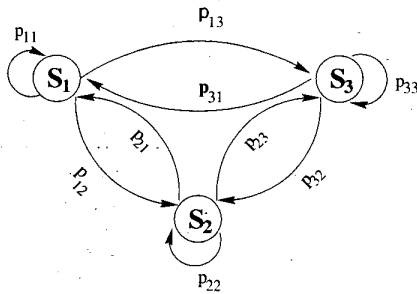


Figure 4: A Markov modulated self-similar process model for MPEG video

#### 4.1. Markov Modulated Self-similar Processes Model

The ACF of each self-similar process is very different from that of the original sequence. For the sake of brevity, only the ACF associated with the active part is shown in Fig. 3. The fluctuation is no longer that big. We have used  $k^{-\beta}$ ,  $e^{-\beta k}$  and  $e^{-\beta\sqrt{k}}$ , corresponding to the ACFs of a self-similar process, a Markov process and an  $M/G/\infty$  input process, respectively, to approximate the ACFs of these three processes. From Fig. 3 it is quite clear that  $k^{-\beta}$  is a better approximation of the ACFs of these classified data, and we therefore use self-similar processes  $s_1$ ,  $s_2$ , and  $s_3$  to model these processes. Using the least square method, we obtained  $\beta = 0.3321$ ,  $0.3069$ , and  $0.4396$  for the active, inactive, and most active part, respectively. The corresponding Hurst parameters for these self-similar processes are  $H = 0.8339$ ,  $0.8465$ , and  $0.7802$ .

To model the whole data set, we need a process to govern the transition among the processes  $s_1$ ,  $s_2$ , and  $s_3$  obtained above. Markov chain is used because of its simplicity.

Using Markov chain as the dominating process, our model for MPEG video traffic can be described by the state diagram shown in Fig. 4, where state  $S_1$ ,  $S_2$ , and  $S_3$  cor-

Table 1: Least square errors obtained by self-similar process, Markov and  $M/G/\infty$  method

	1	P	$B_1$	$B_2$	$B_3$	$B_4$	$B_5$	$B_6$	$B_7$	$B_8$
LRD	0.5	0.4	0.5	0.5	0.4	0.5	0.4	0.5	0.5	0.5
$M/G/\infty$	5.1	12	5.4	4.8	6.2	5.2	5.6	5.0	5.0	5.3
Markov	8.0	21	9.5	7.3	11	8.1	10	8.0	7.2	8.8

respond to the three respective self-similar processes. At state  $S_i$ , bit rates are generated by process  $s_i$ . The transition probability from  $S_i$  to  $S_j$  can be estimated from the empirical data as follows:

$$p_{ij} = \frac{N_{ij}}{N_i}, \quad (1)$$

where,  $N_i$  is the total number of times that the system goes through state  $S_i$ ,  $N_{ij}$  is the number of times that the system make transition to state  $S_j$  from state  $S_i$ . For the *Star Wars* video, the following transition matrix

$$\hat{P} = \begin{bmatrix} 0.0002 & 0.9998 & 0 \\ 0.1174 & 0.5232 & 0.3594 \\ 0.0209 & 0.9791 & 0 \end{bmatrix}$$

is obtained. This matrix is useful for the synthesis of video traffic.

#### 4.2. Structurally Modulated Self-similar Processes Model

In the second model, we have also used  $k^{-\beta}$ ,  $e^{-\beta k}$ , and  $e^{-\beta\sqrt{k}}$ , corresponding to the ACFs of a self-similar process, a Markov process, and an  $M/G/\infty$  input process, respectively, to approximate ACFs of these processes. Owing to the space limit, only the approximation for  $X_P$  is shown in Fig. 5. The sums of squares of errors obtained by the three kinds of methods are tabulated in Table 1. Again, it is quite obvious that self-similar processes are better choices, justifying our usage of self-similar processes for modeling these data.

Using the least squares method, we obtained  $\beta = 0.4663$ ,  $0.3546$ ,  $0.4468$ ,  $0.4779$ ,  $0.4294$ ,  $0.4656$ ,  $0.4380$ ,  $0.4682$ ,  $0.4465$ , and  $0.4606$  are derived for  $X_I$ ,  $X_P$ ,  $X_{B_1}$ ,  $X_{B_2}$ ,  $\dots$ , and  $X_{B_8}$ , respectively. The corresponding Hurst parameters for these processes are  $H = 0.7668$ ,  $0.8227$ ,  $0.7766$ ,  $0.7610$ ,  $0.7853$ ,  $0.7672$ ,  $0.7810$ ,  $0.7659$ ,  $0.7768$ ,  $0.7697$ , respectively.

In our second model, we combine  $X_I$ ,  $X_P$ ,  $X_{B_1}$ ,  $X_{B_2}$ ,  $\dots$ , and  $X_{B_8}$  in a manner similar to the GOP pattern to obtain the model for the MPEG coded traffic. This model can be used to more accurately generate traffic data than the first one.

#### 4.3. The Simple MPEG Video Traffic Model

In the third model, again we have used  $k^{-\beta}$ ,  $e^{-\beta k}$ , and  $e^{-\beta\sqrt{k}}$ , corresponding to the ACFs of a self-similar process, a Markov process, and an  $M/G/\infty$  input process, respectively, to approximate ACFs of these processes. Similar to the previous case, the self-similar processes are the better

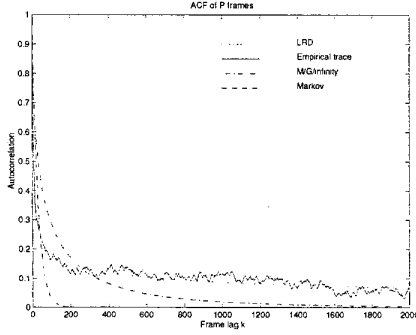


Figure 5: Approximation for ACF of P frames by : Self-similar (LRD), M/G/ $\infty$ , and Markov processes.

choices. Again, using the least squares method,  $\beta = 0.4662$ ,  $0.3404$ ,  $0.3040$  are derived for  $X_I$ ,  $X_P$ ,  $X_B$ , respectively. The corresponding Hurst parameters for these processes are  $H = 0.7669$ ,  $0.8296$ ,  $0.8480$  respectively.

Modulating these three parts in a way similar to the GOP structure leads to our third model, which is simple, yet accurate.

## 5. MODELING OF CDF USING BETA DISTRIBUTION

As mentioned at the beginning, the CDF is another important statistics to catch. Beta distribution [2] is used to model the marginal distributions of these processes. The marginal distribution of a Beta distribution process has the following form

$$f(x; \gamma, \eta, \mu_0, \mu_1) = \begin{cases} \frac{1}{\mu_1 - \mu_0} \frac{\Gamma(\gamma + \eta)}{\Gamma(\gamma)\Gamma(\eta)} \left(\frac{x - \mu_0}{\mu_1 - \mu_0}\right)^{\gamma-1} \left(1 - \frac{x - \mu_0}{\mu_1 - \mu_0}\right)^{\eta-1} & \mu_0 \leq x \leq \mu_1, 0 < \gamma, 0 < \eta \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where  $\gamma$  and  $\eta$  are shape parameters, and  $[\mu_0, \mu_1]$  is the domain where the distribution is defined. Beta distribution is quite versatile and can be used to model random processes with quite different shapes of marginal distributions. The following formulae are used to derive the parameters of Beta distribution:

$$\hat{\eta} = \frac{1 - \bar{x}}{s^2} [\bar{x}(1 - \bar{x}) - s^2] \quad (3)$$

$$\hat{\gamma} = \frac{\bar{x}\hat{\eta}}{1 - \bar{x}} \quad (4)$$

where

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i, \quad (5)$$

$$s^2 = \frac{N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2}{N(N-1)}, \quad (6)$$

and  $N$  is the number of data in the data set.

For simplicity, only the parameters for the third model are listed here. That is,  $\hat{\gamma} = 4.0605$ ,  $1.6605$ ,  $1.6431$ , and  $\hat{\eta} = 10.4273$ ,  $12.0277$ ,  $14.0742$ , which are derived for  $X_I$ ,  $X_P$ ,  $X_B$ , respectively. The simulations demonstrate that the Beta-distribution follows the CDF very closely in all three models. For illustration purpose, only CDF of I frames and its approximation by Beta distribution of the third model are shown in figure 6.

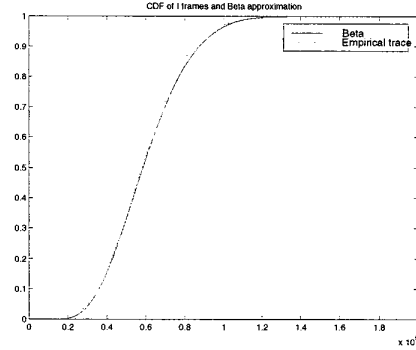


Figure 6: CDF of I frames and its approximation by Beta distribution.

## 6. COMPARISON AND DISCUSSION

The performance of the three models have been presented. First, it is found from our simulations that the Markov modulated self-similar processes model cannot reflect the fluctuating pattern existing in the ACF of the data trace as well as the last two models can. Furthermore, the data generated by using the first model cannot follow the real traffic data very closely. The last two models can however be used to generate traffic data, which are similar to the data trace. A trace of traffic data and the ACF generated by the third model are shown in Fig. 7 and 8, respectively. We therefore conclude that the last two models which are based on the MPEG structure are more accurate than the first model.

Secondly, as shown above, the third model, though simpler than the second model, is almost as accurate in tracking the ACF of MPEG data. It can generate traffic data, which are similar to the data trace, as demonstrated in Fig. 7 and 8.

Since motion/scene complexity is not the focus of this paper, the technique used to classify inactive, active and very active parts of MPEG video traffic adopted in the first model (refer to the formula in Section 3.1) is rather unsophisticated. Therefore it is conceivable that the performance of the first model (decomposing video into three parts based on motion/scene complexity) can be improved significantly with an advanced technique to accurately identify motion/scene complexity. This task is by no means easy, and is a hot research subject in the field of computer vision and video processing. On the other hand, the GOP structure of MPEG data is universal regardless of motion/scene

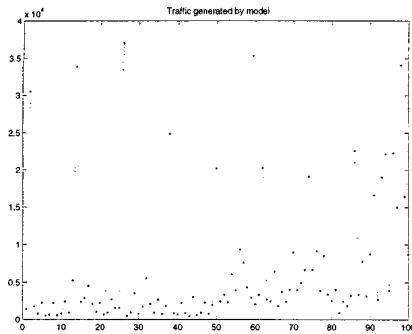


Figure 7: Traffic data generated by our model

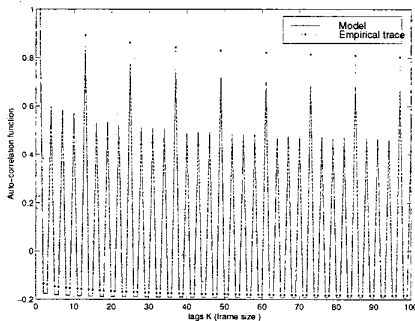


Figure 8: ACFs of empirical trace and our traffic model

complexity. Therefore, the decomposition according to the MPEG data structure is practical and straightforward. The simple model can be readily used to accurately model as well as generate MPEG video data.

## 7. REFERENCES

- [1] M.W.Garrett and W.Willinger, "Analysis, Modeling and Generation of Self-Similar VBR Video Traffic," *Proc. ACM SIGCOMM'94*, London, U.K., Aug. 1994.
- [2] G.J. Hahn, and S.S.Shapiro, *Statistical Models in Engineering*, John Wiley & Sons, Inc., New York 1967.
- [3] D.P.Heyman, and T.V.Lakshman, "Source Models for VBR Broadcast-Video Traffic," *IEEE/ACM Trans. on Networking*, Vol. 4, No.1, Feb. 1996.
- [4] D.P.Heyman, "The GBAR Source Model for VBR Videoconferences," *IEEE/ACM Trans. on Networking*, Vol.5, No.4, Aug. 1997.
- [5] M.M.Krunz, and A.M. Makowski, "Modeling Video Traffic Using  $M/G/\infty$  Input Processes: A Compromise between Markovian and LRD Models," *IEEE Journal on Selected Areas in Comm.*, Vol.16, No.5, June 1998.
- [6] H. Liu, N. Ansari and Y. Q. Shi, "Models of MPEG coded video traffic: Sequentially modulated self-similar process," presented in IFIP Broadband Communications'99, Hong Kong, November 10-12, 1999, included in D. H. K. Tsang and P. J. Kuhn Ed., *Broadband Communications*, pp. 63-72, Norwell, MA, Kluwer Academic Publishers, 1999.
- [7] H. Liu, N. Ansari and Y. Q. Shi, "Markov-modulated self-similar processes: MPEG coded video traffic modeler and synthesizer," *Proceedings of Globcom 1999*, pp. 1184-1188, Rio de Janeiro, Brazil, December 1999.
- [8] H. Liu, N. Ansari and Y. Q. Shi, "A simple model for MPEG video traffic," *Proceedings of IEEE International Conference of Multimedia and Expo*, New York City, July 31 to August 2, 2000.
- [9] B.Maglaris et al., "Performance Models of Statistical Multiplexing in Packet Video Communications," *IEEE Journal on Selected Areas in Communications*, Vol.36, No. 7, July 1988.
- [10] B.Melamed, and D.E. Pendarakis, "Modeling Full-Length VBR Video Using Markov-Renewal-Modulated TES Models," *IEEE/ACM Trans. on Networking*, Vol.5, No.1, Feb. 1997.
- [11] V.Paxson and S.Floyd, "Wide-Area Traffic: The Failure of Poisson Modeling," *Proc. ACM SIGCOMM'94*, London, U.K., Aug. 1994.
- [12] P.Skelly, M.Schwartz, and S.Dixit, "A Histogram-Based Model for Video Traffic Behavior in an ATM Multiplexer," *IEEE/ACM Trans. on Networking*, Vol.1, No.4, Aug. 1993.