TES-Based Video Source Modeling For Performance Evaluation of Integrated Networks

B. MELAMED*, D. RAYCHAUDHURI†, B. SENGUPTA* and J. ZDEPSKI†

Abstract

This paper considers modeling methodologies of variable bit-rate (VBR) video sources for performance evaluation of integrated networks. We consider an example in which compressed VBR video is transmitted over a local area network carrying both video and data. We devised a group-of-block (GOB) level simulation of an H.261 algorithm over representative VBR image sequences. This simulation data is processed to obtain a statistical characterization in terms of bit-rate histogram, autocorrelation function, etc. We used a new modeling methodology, called TES (*Transform-Expand-Sample*), which is a general method for generating autocorrelated time series. TES is a non-parametric method that can accurately capture the histogram and approximate the autocorrelation function of any data set, simultaneously. We used the TES model to drive a network simulation for performance evaluation. The simulation results were compared with those obtained earlier with a simpler frame-oriented autoregressive (AR) model. The numerical results show a small, but not negligible, difference in key network performance measures (such as throughput-delay).

*NEC USA Inc., C&C Research Laboratories, Princeton, New Jersey 08540
†David Sarnoff Research Center, Princeton, New Jersey 08543

1 Introduction

This letter describes a new TES (Transform-Expand-Sample) methodology [1, 3, 4, 9] for modeling of variable bit-rate (VBR) video sources [2, 6, 8, 11]. Such VBR video source models are an important prerequisite for performance evaluation of integrated broadband networks such as FDDI and DQDB/IEEE802.6 and, most notably, asynchronous transfer mode (ATM) B-ISDN. TES is a non-parametric technique that enables a user to model stationary stochastic processes with an exact match of the marginal distribution and a good approximation of the autocorrelation function. In our attempts to model typical VBR video source processes, we have found that the TES method provides a more accurate representation (both qualiatively and quantitatively) than alternative conventional approaches while involving comparable computational effort. In this work, we illustrate the proposed TES modeling method through an example packet video scenario, and provide comparative network performance results based on alternative TES and autoregressive (AR) approaches. The numerical results show a small, but not neglegible, difference in key network performance measures such as throughput-delay.

2 Source Modeling Scenario

We consider the VBR video source modeling problem in context of an example near-term scenario in which low bit-rate CCITT H.261 teleconferencing / multimedia video is transmitted over a 10 Mbps IEEE 802.3 CSMA/CD local area network (LAN). The specific packet video system considered employs VBR H.261 codecs connected to the LAN via a transport interface that provides video-specific features such as delay compensation buffering, modified (compatible) CSMA/CD retransmission backoff and priority queueing (for two-layer video). In [13, 14], performance results for this system were presented using a frame-level autoregressive (AR) modeling approach for the VBR video. This AR modeling approach, which has also been used by other authors, e.g., [2, 12] appears to be satisfactory for a first-order evaluation of network capacity/delay. However, the lattice-filter AR formulation suffers from two serious limitations for the VBR H.261 source: (a) The model is only appropriate for H.261 frame level data, rather than the "group-of-blocks" (GOB) level that is the natural transport unit for video data entering the network. The GOB level bit-rate process is characterized by an autocorrelation that is periodic both at the spatial GOB scan rate and at temporal frame rate, effects that cannot easily be incorporated into the AR model. (b) A gaussian approximation used in the AR method implies that the marginal bit-rate distribution cannot be exactly matched to the measured bit-rate data. These two limitations motivate a more accurate source modeling approach capable of capturing the GOB-level marginal bit-rate distribution and periodic autocorrelation properties.

3 Video Source Model

A variable bit-rate (VBR) encoder based on the CCITT H.261 motion compensated (MC) discrete cosine transform (DCT) algorithm [7] is considered in this study (see Fig. 1). The pixels to be encoded are taken from images in the CIF (Common Intermediate Format) frame format shown in Fig. 2. The CIF format consists of 352×288 pixels of luminance (Y) and 176×144 pixels each of filtered chrominance (U and V), at a frame rate of 30 fps. Observe that each frame of CIF video is partitioned into 12 GOB units, which are further subdivided into 33 "macroblocks" (MB) each. A macroblock is the basic unit of video to be processed, and consists of four 8×8 pixel luminance blocks (Y) along with the associated pair of 8×8 chrominance blocks (U,V). Motion vectors, adaptation co-efficients, etc. are computed once per macroblock for transmission in a MB header, which is then followed by six variable length coded (VLC) DCT blocks each containing quantized coefficient values, zero runlengths, etc. Taking into account the typical size of a MB's and GOB's described, the GOB has been identified as the a suitable protocol data unit (PDU) for network transport. Specifically, for a VBR H.261 encoder configured to operate at an average bit-rate of 128 kbps, a GOB will vary in size between .5 and 5 kb (with an average of about 2 kb), and is therefore within the acceptable range of packet sizes for an 802.3 LAN. Occasionally, a GOB will exceed the maximum link-level packet size limit of the LAN, requiring transport level segmentation and reassembly.

3.1 The TES Modeling Methodology

The limitations of the simple AR model used earlier motivated us to consider alternative approaches to simultaneously capture the empirical GOB-level marginal bit-rate distribution and the associated autocorrelation function. To this end, we used the recently proposed TES methodology [9, 3, 4]. TES is a class of methods for generating autocorrelated random sequences via autoregressive modulo-1 schemes. They enjoy several important advantages. First, a TES method is non-parametric; that is, TES can generate any marginal distribution or an arbitrarily close approximation. In particular, TES can match any empirical histogram. Ordinary first order autoregressive processes do not have this property. Second, the associated autocorrelation function can be rapidly computed numerically [3, 4], thereby facilitating a search for an adequate fit. Third, TES methods are computationally fast; in fact, uniform TES sequences have computational complexity only slightly higher than that of the underlying pseudo-random number generator. And fourth, TES sequence behavior is versatile, ranging from driftless random walks to cyclic ones. The attendant autocorrelation functions can have a variety of functional forms, including monotone, oscillatory and others. Refer to [9, 3, 4] for more details.

The TES methodology models empirical densities as histograms, that is, as probabilistic mixtures of uniform variates of the form

$$h(y) \;=\; \sum_{n=1}^{N} 1_{[l_n, r_n)}(y)\,\frac{p_n}{w_n}, \tag{1}$$

where N is the number of histogram cells, cell n has the form $[l_n, r_n)$, 1_A is the indicator function of set A, $w_n = r_n - l_n$ is the width of cell n, and p_n is the probability of cell n. To model temporal dependences in empirical random sequences, we made use of a TES variant called TES$^+$. Let $\lfloor x \rfloor = \max\{n \text{ integer} : n \le x\}$ be the integral part of a real number x, and denote by $\langle x \rangle = x - \lfloor x \rfloor$ the fractional part of x. Let $\{V_n\}$ denote a sequence of independent identically distributed random variables, called *innovations*, with a common, though arbitrary, density f_V. A TES$^+$ time series $\{U_n^+\}$ is defined recursively by

$$U_n^+ \;=\; \begin{cases} U_0, & \text{if } n = 0 \\ \langle U_{n-1}^+ + V_n \rangle, & \text{if } n > 0 \end{cases} \tag{2}$$

It was shown in [3] that $\{U_n^+\}$ is a random walk on a circle with circumference one (unit circle), and its marginal distribution is uniform on $[0, 1)$.

Recall that any uniform variate U on $[0, 1)$ can be transformed to a desired marginal distribution F through the *inversion method* by applying $F^{-1}(U)$. Transformations of uniform variates on $[0, 1)$ are called *distortions*. TES models routinely use compound distortions, composed of two successive transformations. The first, called a *stitching transformation*, is selected from a family, parameterized by $0 \leq \xi \leq 1$, and defined by

$$S_\xi(y) \; = \; \begin{cases} y/\xi, & 0 \leq y < \xi \\ (1-y)/(1-\xi), & \xi \leq y < 1 \end{cases} \tag{3}$$

For $0 < \xi < 1$, this constitutes a "smoothing" operation, since the associated S_ξ is continuous on the unit circle, and consequently, the corresponding stitched sequence $\{S_\xi(U_n^+)\}$ is "smoother" than the underlying $\{U_n^+\}$. Moreover, it can be shown that all S_ξ, $0 \leq \xi \leq 1$, preserve uniformity [9]. It follows that stitched TES processes can still be transformed to arbitrary marginals by the inversion method. The second transformation, called *histogram distortion*, is $D_H = H^{-1}$, where H is the cumulative distribution function corresponding to the density $h(y)$ of Eq. (1). More explicitly,

$$D_H(x) \; = \; \sum_{\substack{k=1 \\ p_k > 0}}^{N} 1_{[C_{k-1}, C_k)}(x) \left[l_k + (x - C_{k-1}) \, \frac{w_k}{p_k} \right], \tag{4}$$

where $C_n = \sum_{j=1}^{n} p_n$, $1 \leq n \leq N$ is the cumulative distribution of $\{p_n\}$ ($C_0 = 0$ and $C_N = 1$). The outcome of the TES modeling methodology is a sequence $\{X_n^+\}$ of the form $X_n^+ = D_H(S_\xi(U_n^+))$. Here the distortion D_H is determined from the empirical data, while ξ and f_V are selected by the modeler through a heuristic search for a model whose autocorrelation function approximates its empirical counterpart and whose Monte Carlo sample paths bear "adequate resemblance" to the empirical data. The search can be efficiently carried out, using the visual interactive software environment described in [1]. It employs general formulas TES autocorrelations developed in [3], and explicit ones for stitched histogram distortions derived in [4].

3.2 TES-Based GOB Source Model for VBR Compressed Video

Fig. 3 displays the raw empirical data $\{\hat{X}_n\}$, and Fig. 4 shows the empirical autocorrelation function $\hat{\rho}_X(\tau)$ for $1 \leq \tau \leq 100$ (bullets) against its TES-based model counterpart (bars). The cyclical nature of the data is clearly evident; each cycle corresponds to a run of twelve observations comprising the coded blocks of a single frame. Since every twelfth block codes the same screen portion of a frame, we expect to see a strong deterministic periodicity of twelve in the data. This is indeed the case, as evidenced by the same periodicity of the empirical autocorrelation function and its sustained amplitude. Since TES autocorrelation functions decay much faster, it became clear that a TES model could not be properly fitted to the raw empirical data. Rather, the raw data would have to be transformed first, so as to remove the deterministic periodic components. To this end, we performed a periodogram analysis and estimated from the raw data a set of parameters K, ω_i, A_i and B_i, $1 \leq i \leq K$, such that the trigonometric sums $\{\hat{T}_n\}$, given by

$$\hat{T}_n \; = \; \sum_{i=1}^{K} \left[A_i \, \cos(\omega_i n) + B_i \, \sin(\omega_i n) \right], \tag{5}$$

tracked well the oscillations of the raw data $\{\hat{X}_n\}$. We then transformed $\{\hat{X}_n\}$ into a new sequence $\{\hat{R}_n\}$ (called the *residuals sequence*), where $\hat{R}_n = \hat{X}_n - \hat{T}_n$. The residuals sequence had a much faster decaying autocorrelation function as compared to $\hat{\rho}_X(\tau)$, and a considerably reduced magnitude. Based on the appearance of the residuals sample path, we fixed ξ at 0.5 and selected a simple class of uniformly distributed innovations of the form $V_n = L + (L+R)Z_n$, where $-0.5 \le L < R \le 0.5$ are parameters, and $\{Z_n\}$ is a sequence of i.i.d. variates, uniform on [0,1), commonly available on most computers as pseudo random number streams. Switching to the equivalent parameters $\alpha = R - L,\ \phi = \frac{R+L}{R-L}$, we performed a computer search of the parameter space $\{(\alpha, \phi):\ 0 < \alpha < 1,\ -1 \le \phi \le 1)\}$ for a good TES model $\{U_n^+\}$, such that $Y_n = D_H(S_{0.5}(U_n^+))$ had the same marginal distribution as $\{\hat{R}_n\}$, and the first few autocorrelations of $\{Y_n\}$ were a good match to their $\{\hat{R}_n\}$ counterparts. The final step was to add back the trigonometric components $\{\hat{T}_n\}$ to $\{Y_n\}$, leading to the final model of coded video GOB bit-rate

$$X_n = D_H(S_{0.5}(U_n^+)) + \hat{T}_n, \tag{6}$$

where D_H is given by Eq. (4), S_ξ by Eq. (3), U_n^+ by Eq. (2) and $\hat{T}_n$ by Eq. (5).

Fig. 5 depicts a sample path generated by a simulation of $\{X_n\}$ from Eq. (6), with the choice of α, ϕ and ξ indicated in the heading. Notice the similarity of the simulated sample path in Fig. 5 to its empirical counterpart in Fig. 3. A comparison of the corresponding autocorrelation functions, in Fig. 6, reveals a remarkable match, at least up to lag 100. Finally, Fig. 7 compares the simulation-based histogram (solid line) to its empirical counterpart (dashed line), again with good agreement. Note that the empirical histogram is far from being gaussian (as assumed in the simple AR model).

4 Numerical Results

A simulation study was conducted in order to assess the impact of using the more accurate TES source model instead of the AR model. As in [13, 14], the performance measures were obtained by simulation of a CSMA/CD LAN, supporting a mix of interactive data stations and *TES-modeled* VBR video stations. The details of this simulation study appear in [10]. The CSMA/CD network is assumed to have a collision detection delay of 1 μs, a channel speed of 10 Mbps with 50 data stations distributed uniformly over the length of 500 meters of cable. The number of video stations was varied to produce different load levels.

Fig. 8 shows the corresponding video packet loss probability *vs.* LAN throughput predicted by the AR and TES models (both are shown for buffer size of 100 ms, engineered for transport-level delay compensation). As in [14], we observe that the video packet loss rises quite rapidly at throughput levels higher than about 50% of capacity; an acceptable level of 10^{-5} is achieved at an operating throughput in the region of 55% of capacity. At throughput levels exceeding 60% of capacity, the AR and TES model results show some divergence, saturating to maximum throughput levels of 0.68 and 0.71 respectively.

In conclusion, we have presented the TES methodology for modeling of VBR sources and compared it to the AR method. The performance results obtained from the the two modeling approaches differed in capacity by about 4-5%. While this difference is not large, it gives more confidence in the results since the TES method is more accurate.

References

[1] D. Geist and B. Melamed, “TEStool: An Environment for Visual Interactive Modeling of Autocorrelated Traffic”, *Proceedings of ICC ‘92*, Vol. 3, 1285-1289, 1992, Chicago, Illinois.

[2] Hang, “Source Modelling for Packet Video,” *Conf. Rec. ICC*, Phila. PA, 1988, pp 38.7.1-38.7.6.

[3] D.L. Jagerman and B. Melamed, "The Transition and Autocorrelation Structure of TES Processes Part I: General Theory," *Stochastic Models*, Vol. 8(2), 193-219, 1992.

[4] D.L. Jagerman and B. Melamed, "The Transition and Autocorrelation Structure of TES Processes Part II: Special Cases," *Stochastic Models*, Vol. 8(3), 499-527, 1992.

[5] Joseph, Raychaudhuri and Zdepski, "Shared Access Packet Transmission systems for Compressed Digital Video," *IEEE J. Selected Areas in Commun.*, 1989, pp. 815-825.

[6] Kishimoto, Ogata and Inumaru, "Generation Interval Distribution Characteristics of Packetized Variable Rate Video Coding Data Streams in an ATM Network," *IEEE J. Selected Areas in Commun.*, 1989, pp. 833-841.

[7] D-S. Lee, B. Melamed, A.R. Reibman and B. Sengupta, "Analysis of a Video Multiplexer Using TES as a Modeling Methodology", *Proceedings of GLOBCOM '91*, pp. 1.3.1-1.3.5, 1991, Phoenix, Arizona.

[8] Maglaris, et.al., "Performance Models of Statistical Multiplexing in Packet Video Communications," *IEEE Trans. Comm.*, Vol. COM-36, Vol. COM-36, 1988, pp. 834-844.

[9] B. Melamed, "TES: A Class of Methods for Generating Autocorrelated Uniform Variates," *ORSA J. on Computing*, Vol. 3(4), 1991, pp. 317-329.

[10] B. Melamed, D. Raychaudhuri, B. Sengupta and J. Zdepski, "TES-Based Video Source Modeling for Performance Evaluation of Integrated Networks," *INFOCOM '92*, pp.75-84.

[11] M. Nomura, T. Fujii and N. Ohta, "Basic Characteristics of Variable Rate Video Coding in ATM Environment," *IEEE J. Selected Areas in Commun.*, 1989, pp. 752-760.

[12] G. Ramamurthy and B. Sengupta, "Modeling and Analysis Of A Variable Bit Rate Video Multiplexer," *INFOCOM '92*, pp. 817-827, 1992.

[13] J. Zdepski, K. Joseph and D. Raychaudhuri, "Packet Transport of Rate-Free Interframe DCT Compressed Digital Video on a CSMA/CD LAN," *IEEE GLOBECOM*, 1989, pp.44.3.1-7.

[14] Zdepski, Joseph, Raychaudhuri and Daut, "Prioritized Packet Transport of VBR H.261 Format Compressed Video on a CSMA/CD LAN," *Third International Workshop on Packet Video*, Morristown, NJ, 1990.