VARIABLE BIT RATE VIDEO: CHARACTERISTICS, MODELING AND MULTIPLEXING

Daniel Reininger

David Sarnoff Reasearch Center
Princeton, NJ 08543-5300

Benjamin Melamed
Dipankar Raychaudhuri
Bhaskar Sengupta

C&C Research Laboratories
NEC USA, Inc.
Princeton, New Jersey 08540

1 SOURCE MODELING

This section describes the TES source model for MPEG-coded video, used in the performance study; TES models for other video coding schemes may be found in [8]. The model was constucted as a deterministic superposition of three component source models, one each for a subsequence of I-frames, P-frames and B-frames. The component bit rates were then interleaved (superposed) in the appropriate MPEG cycles as explained before. Each component subsequence was accurately modeled as a TES process; the details of TES process theory may be found in [6, 3, 4, 5], and an overview at an introductory level in [7]. The section is organized in two subsections. The first subsection describes the TES model construction of component bit rate subsequences for each type of frame. The second subsection describes the composite model and presents the modeling results.

1.1 TES Modeling of Component Frame Types

We begin by describing the generic construction of TES models from empirical bit rate measurements of MPEG-coded video. TES modeling differs from other modeling approaches in that it aims to fit a prescribed marginal distribution and a prescribed autocorrelation function *simultaneously*, via a stationary stochastic process, thereby capturing both first-order and second-order statistical properties of empirical time series.

The construction of TES models is a two-phase procedure. In the first phase, one defines a *background* TES process $\{U_n^+\}_{n=0}^\infty$ or $\{U_n^-\}_{n=0}^\infty$ of the form

$$U_n^+ = \begin{cases} U_0, & n = 0 \\ \langle U_{n-1}^+ + V_n \rangle, & n > 0 \end{cases} \qquad U_n^- = \begin{cases} U_n^+, & n \text{ even} \\ 1 - U_n^+, & n \text{ odd} \end{cases} \tag{1.1}$$

Here, U_0 is distributed uniformly on $[0, 1)$; $\{V_n\}_{n=1}^\infty$ is a sequence of iid random variables, independent of U_0, called the *innovation sequence*; and angular brackets denote the modulo-1 (fractional part) operator $\langle x \rangle = x - \max\{\text{integer } n : n \le x\}$. Background TES sequences serve an auxiliary role. The superscript notation in (1.1) is motivated by the fact that $\{U_n^+\}$ and $\{U_n^-\}$ can generate lag-1 autocorrelations in the range [0,1], and [-1,0] respectively. From now on, we always append plus or minus superscripts to other mathematical objects associated with $\{U_n^+\}$ and $\{U_n^-\}$, but the subscript is omitted when the distinction is immaterial. Intuitively, the modulo-1 arithmetic used in defining TES processes in (1.1) has a simple geometric interpretation as random walks on a circle of circumference 1 (unit circle), with mean step size $E[V_n]$. The walk will have zero, positive or negative drift around the unit circle according as $E[V_n] = 0$, $E[V_n] > 0$ or $E[V_n] < 0$, respectively.

In the second phase, the background sequence is transformed into a corresponding *foreground sequence* $\{X_n^+\}_{n=0}^\infty$ or $\{X_n^-\}_{n=0}^\infty$, respectively,

$$X_n^+ = D(U_n^+), \quad X_n^- = D(U_n^-), \tag{1.2}$$

where D is a transformation from $[0, 1)$ to the reals, called a *distortion*. Eq. (1.2) defines two classes of TES models, denoted TES$^+$ and TES$^-$, respectively, and these foreground sequences are the end-product TES models.

The TES modeling methodology used by us employed a composite distortion of the form

$$D_{Y,\xi}(x) = \hat{H}_Y^{-1}(S_\xi(x)), \quad x \in [0, 1). \tag{1.3}$$

Here, the inner transformation, S_ξ, is a "smoothing" operation, called a *stitching transformation*, parameterized by $0 \le \xi \le 1$, and given by

$$S_\xi(y) \;=\; \begin{cases} y/\xi, & 0 \le y \le \xi \\ (1-y)/(1-\xi), & \xi \le y < 1. \end{cases} \tag{1.4}$$

The outer transformation, $\hat{H}_Y^{-1}$, is the inverse of the empirical (histogram) distribution function computed from an empirical time series $\{Y_n\}_{n=0}^N$ as

$$\hat{H}_Y^{-1}(x) \;=\; \sum_{j=1}^{J} 1_{[\hat{C}_{j-1},\hat{C}_j)}(x)\,[l_j + (x - \hat{C}_{j-1})\,\frac{w_j}{\hat{p}_j}], \quad 0 \le x \le 1, \tag{1.5}$$

where 1_A is the indicator function of set A, J is the number of histogram cells, $[l_j, r_j)$ is the support of cell j with width $w_j = r_j - l_j > 0$, $\hat{p}_j$ is the probability estimator of cell j and $\{\hat{C}_i\}_{i=0}^J$ is the cdf of $\{p_j\}_{j=1}^J$, i.e., $\hat{C}_j = \sum_{i=1}^j \hat{p}_i$, $1 \le j \le J$ ($\hat{C}_0 = 0$ and $\hat{C}_J = 1$).

To understand the modeling procedure sketched above, we take note of the following facts.

1. It can be shown [3] that all TES background sequences are stationary Markovian, and their marginal distribution is uniform on $[0, 1)$, *regardless* of the probability law of the innovations $\{V_n\}$.
2. The *inversion method* [1] allows us to transform any uniform variate to others with arbitrary distributions: if U is uniform on $[0, 1)$ and F is a prescribed distribution, then $X = F^{-1}(U)$ has distribution F. In particular, $F = \hat{H}_Y$ is just a special case.
3. For $0 < \xi < 1$, the effect of S_ξ is to render the sample paths of background TES sequences more "continuous-looking". Because stitching transformations preserve uniformity [6], the inversion method can still be used on stitched background processes $\{S_\xi(U_n)\}$.

It follows that any foreground TES sequence of the form

$$X_n \;=\; \hat{H}_Y^{-1}(S_\xi(U_n)), \tag{1.6}$$

obtained from any background sequence $\{U_n\}$, is always guaranteed to have the empirical (histogram) distribution $\hat{H}_Y$, *regardless* of the innovation density f_V and stitching parameter ξ selected. The choice of a pair (f_V, ξ) will determine the dependence structure of (1.6), and in particular, its autocorrelation function. Thus, TES modeling decouples the fitting of the empirical distribution from the fitting of the empirical autocorrelation function. Since the former is guaranteed, one can concentrate on the latter.

In practice, approximating the empirical autocorrelation structure is carried out via a heuristic search for a suitable pair (f_V, ξ), under software support. We used the TEStool visual interactive software environment [2], designed to support TES modeling. TEStool allows the user to read in empirical time series and calculate their statistics (histogram, autocorrelation function and spectral density) in textual and graphical representations. It provides services to construct and modify TES models and to superimpose the corresponding TES statistics on their empirical counterparts. The search proceeds in an interactive style, guided by visual feedback: each model modification triggers a recalculation and redisplay of the corresponding statistics. TES model autocorrelations and spectral densities are calculated numerically from fast and accurate formulas developed in [3, 4].

While selecting a stitching parameter $0 \le \xi \le 1$ is a simple matter, the selection of an innovation density f_V is far more involved. TEStool restricts innovation densities to the class of step-function densities over $[-0.5, 0.5)$ with no practical loss of generality, since step-function densities can approximate any density arbitrarily closely. The advantage of this restriction is a conceptual and operational simplification. Step-function densities can be readily modified graphically, since steps are visually represented as rectangles, and those can be created, deleted, stretched and moved easily with the mouse.

1.2 A Composite TES Model for MPEG-Coded Video

The deterministic sequencing of MPEG frames renders the VBR process a non-stationary one. In order to faithfully model MPEG video, the correct interleaving of frame types should be effected. Figure 1 displays the results of of an MPEG case study with cycles of the form IBBPBBPBB....

The modeling procedure proceeded in two stages. In the first stage, each frame type (I, P and B) was modeled by a TES process, of the form (1.6), as described in the previous section. The construction utilized empirical bit rate measurements of an MPEG-encoded video sequence. A separate TES$^+$ model was fitted to the I-frame and B-frame subsequences, while a TES$^-$ model was fitted to the P-frame subsequence. These will be denoted $\{X_n^I\}$, $\{X_n^P\}$ and $\{X_n^B\}$, respectively, and the corresponding background sequences will be similarly denoted $\{U_n^I\}$, $\{U_n^P\}$ and $\{U_n^B\}$. All TES models matched the the corresponding empirical distributions, and approximated well the respective empirical autocorrelation functions.

In the second stage, the three TES models were interleaved in the correct order above. However, cross-correlations were induced into bit rates comprising individual cycles as follows. The inaugurating I-frame bit rate X_{n+1}^I for the next cycle was generated via U_{n+1}^I from its I-subsequence predecessor U_n^I in the normal way, detailed in the previous section. The next bit rate, X_{6n+1}^B (recall that there are 6 B-frames in a cycle) did not use U_{6n+1}^B; rather, it set $U_{6n+1}^B = U_{n+1}^I$. The remaining B-frames in that cycle were generated normally from their predecessors within the cycle via the TES model $\{X_n^B\}$ for B-frames. Similarly, the first P-frame bitrate in the cycle X_{2n+1}^P (recall that there are 2 P-frames in a cycle) set $U_{2n+1}^P = U_{n+1}^I$ and the remaining P-frames were again generated normally from their predecessors within the cycle via the TES model $\{X_n^P\}$ for P-frames. Thus, rather than having independent bit rates for different frame types within cycles and across them, the inaugurating I-frame of each cycle and the corresponding P-frames and B-frames within that cycle were rendered dependent random variables; the B-frame and P-frame bit rates fluctuated within each cycle around a baseline set by the inaugurating I-frame, and dependece among cycles was driven primarily by I-frames.

The composite TES model statistics are compared to their empirical counterparts in Figure 1. The screen is sub-divided into four canvas areas displaying various types of empirical statistics (bullets) superimposed on their TES model counterparts (diamonds). The upper-left canvas contains the empirical and TES model sample paths, the latter being generated by simulation; the upper-right canvas contains the corresponding histograms; the lower-left and lower-right canvases contain, respectively, the empirical autocorrelation function and spectral density plotted against their simulation-based TES model counterparts. We point out that both the autocorrelation functions and spectral densities were formally computed from a single sample path *as if* the sequences were stationary, and therefore, represent averaged estimates of different correlation coefficients. Nevertheless, they give us an indication of how well the composite TES model captured temporal dependence in the empirical data, since they were all computed from sample paths in the same way.

Indeed, Figure 1 shows excellent agreement of all statistics, as well as a marked similarity of the functional form of the empirical and composite TES model sample paths.

Acknowledgments

We thank Jon Hill for implementing the composite TES model for MPEG-coded video.

References

[1] Devroye, L. 1986. *Non-Uniform Random Variate Generation*, Springer-Verlag.

[2] Geist, D.; and B. Melamed. 1992. "TEStool: An Environment for Visual Interactive Modeling of Autocorrelated Traffic", *Proceedings of the 1992 International Conference on Communications* (Chicago, Illinois). Vol. 3: 1285-1289.

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

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

[5] Jagerman, D. L.; and B. Melamed. 1992c. Spectral Analysis of Basic TES Processes. *Preprint*, NEC USA, Inc., Princeton, New Jersey.

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

[7] Melamed, B.; D. Goldsman; and J. Hill. 1992. "The TES Methodology: Nonparametric Modeling of Stationary Time Series". *Proceedings of the 1992 Winter Simulation Conference* (Arlington, Virginia), 135–144.

[8] Melamed B.; and B. Sengupta. 1992. "TES Modeling of Video Traffic". *IEICE Transactions on Communications*, Vol. E75-B, No. 12, 1292–1300.

Author Biography

BENJAMIN MELAMED is Head of the Performance Analysis Department at the C&C Research Laboratories, NEC USA, Inc., in Princeton, New Jersey. His research interests include systems modeling and analysis, simulation, stochastic processes, and visual modeling environments. Melamed received his Ph.D. in Computer Science in 1976 from the University of Michigan. Before coming to NEC, he taught in the Department of Industrial Engineering and Management Science at Northwestern University. He later joined the Performance Analysis Department at Bell Laboratories, where he became an AT&T Bell Laboratories Fellow in 1988. He is a Senior Member of IEEE.

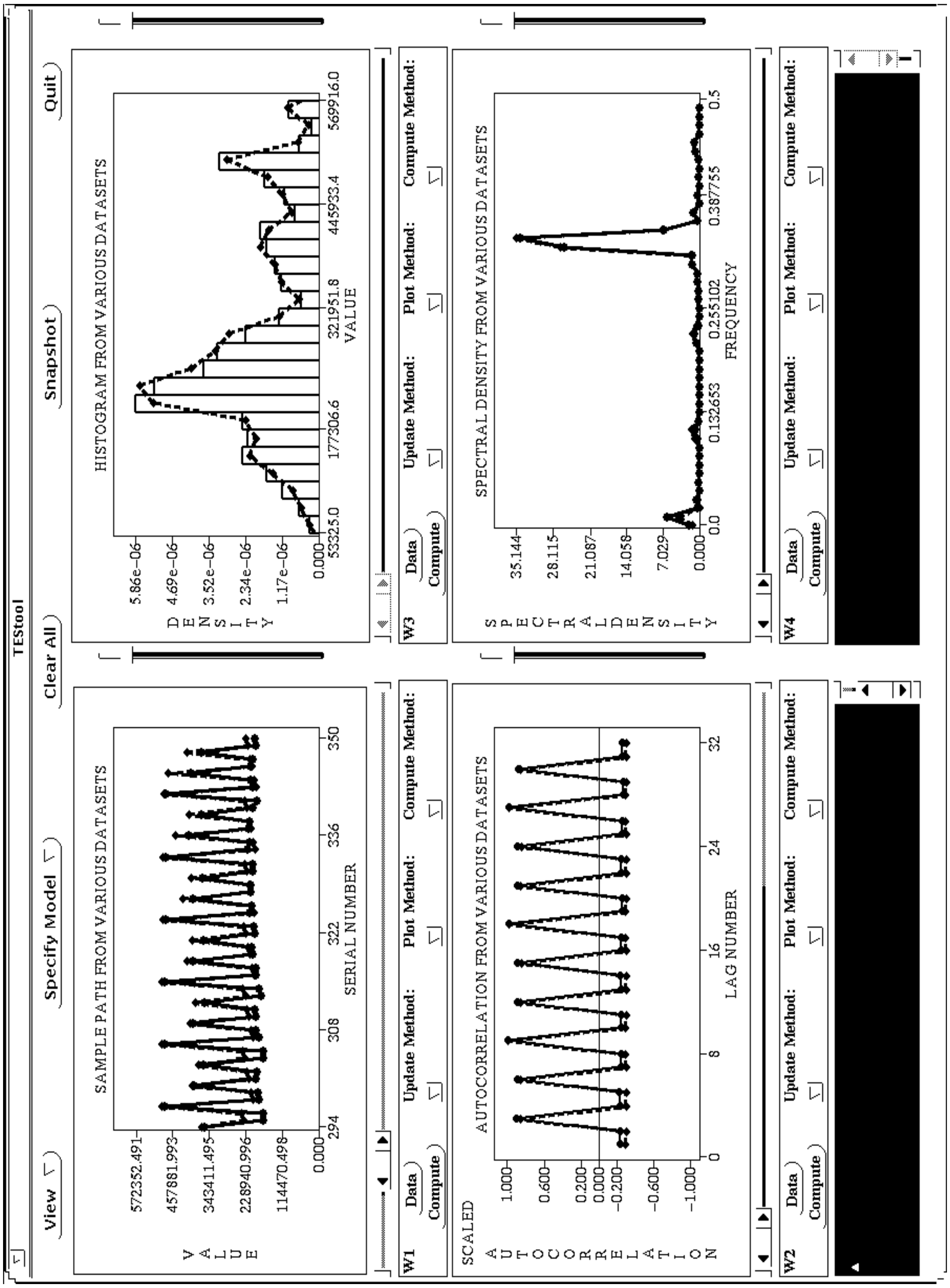

Figure 1: TES multiplex model of MPEG-compressed VBR video.