SCENE ORIENTED MODEL FOR VBR VIDEO

E. Casilari, M. Lorente, A. Reyes, A. Díaz Estrella and F. Sandoval
Dpto. Tecnología Electrónica, E.T.S.I. Telecomunicación,
Universidad de Málaga, Campus de Teatinos, 29071 Málaga (Spain)
Telephone no.: 34-5-2132755; FAX 34-5-2131447; E-mail: casilari@dte.uma.es

INDEXING TERMS
Long Range Dependence, VBR video Traffic, MPEG, Projected Autorregresive Model.

ABSTRACT
In this letter a new model for variable bit rate (VBR) video traffic is presented. The model, which could be used as a traffic generator, considers two time scales: scenes, for periods of several minutes, and groups of pictures (GOP), for periods of half a second. To model the scene changes a Markov chain is used. For the GOP level a modification for the projected autorregresive (PAR) model is proposed so that the fitting of the autocorrelation function is improved. The model is utilised to imitate two real MPEG video signals, showing that it is able to accurately capture the behaviour of the real traffic in a queue.
INTRODUCTION

Among the most emerging real-time services to be conveyed by broadband networks, based on ATM, we have all the video applications, such as teleconferencing, video on demand, etc. These multimedia services will represent a major traffic source so it is necessary to establish accurate models which can approximate their statistical characteristics. These models could be used to dimension, simulate or test the performance of the networks.

A video sequence consists in a series of pictures (or frames) containing a bidimensional matrix of pixels. The pictures can be encoded with a constant bit rate (CBR) and variable image quality or, on the contrary, if a quality level is fixed (a constant signal-to-noise ratio) the size of the frame will depend on the movement and the complexity of the image (VBR video). This second type of video may benefit from the statistical multiplexing that ATM provides. But, in order to optimise the utilisation of the available bandwidth, the behaviour of this traffic must be properly characterised. Hence, the modelling and analysis of VBR video traffic is a still open issue which has focused the interest of researchers in broadband communications. In this letter we propose a two-level model to imitate VBR traffic. The proposed scheme includes the modelling of scene-changes with the aim of capturing long-term dependencies, and it introduces an improvement of the PAR model to generate traffic within each scene (in this case, the number of bits per frame). Finally, the model is used to approximate a real MPEG encoded film, proving that it is able to fit the behaviour of the actual traffic in a queue.

THEORETICAL FORMULATION

The final objective of modelling a traffic source, regarded as a random series $s[n]$ containing the size of the transmitted traffic within a period of time (in this case the GOP size), is the generation of another signal $s'[n]$ which must approximate the behaviour in a queue (losses, average delay, jitter) of $s[n]$. Different models [1] [2] [3] have been proposed for VBR video traffic. Among these models we have TES models, Markov chains, AR, ARMA and PAR models. As a general rule, these models are designed to adjust first order statistics, as the Probability Distribution Function ($F_S(x)$), the mean or the variance, as well as some initial points of the autocorrelation function $R_s[k]$, that is, the Short range Dependences (SRD) that the signal exhibits. If the buffer size is small, their behaviour in a queue usually approximates well that of the target traffic. However these models tend to underestimate the loss probability and the delay parameters as the buffer size increases. This is due to the fact that these models neglect the Long Range Dependence (LRD) that a video sequence presents, since the video signal progresses across scenes with different activity levels.

This LRD or self-similar nature of video traffic can be analytically observed [4] [5] [6] in the hyperbolical decay of the autocorrelation function. The preceding models normally exhibit exponentially decaying autocorrelations so they cannot match $R_s[k]$ for long values of the lag. Thus, to overcome this problem, it is necessary to model the short term variations within a scene but also the transition between different scenes [5]. In particular this letter proposes a two level model: one for scenes and another for GOPs, where the GOP is a fixed number of several frames with the same structure in a MPEG flow.

Scene level model: To model the scene-change we utilise a discrete Markov chain $M_k=\{1, 2, ..., N\}$ where the number of states $N$ corresponds to the number of activity levels to be distinguished. The changes between states will follow a transition matrix $P$ while the sojourn time for each state will be modelled with an exponential distribution with mean $t_i$. To calculate $P$ and $t_i$, we must generate from $s[n]$ a sequence $m[n]$, where $m[i] \in \{1, ..., N\}$ indicates the scene type in which the $i$th GOP is included. To estimate $m[i]$ we must consider not only the activity level of the $i$th GOP but also those of the adjacent GOPs. With this purpose we compute the signal $s_w[n]$ as the result of averaging $s[n]$ through a moving average filter with a window size $W$. Quantizing $s_w[n]$ in $N$ uniform levels and numbering them from 1 to $N$, we would obtain $m[n]$. From $m[n]$ the transition probability between any two states $i$ and $j$ could be obtained and, consequently, the transition matrix $P$. In the same way the mean sojourn times $t_i$ for each state can be computed. With $P$ and $t_i$ the scene level is completely defined.
**GOP level model:** Once we have divided $s[n]$ into different scenes, we group the GOPs belonging to the $i$th activity level in a vector $v_i$. Thus we obtain $N$ different vectors in which the long term dependencies have been cancelled. Consequently, these $v_i$, representing the traffic for each activity level, can be approximated with conventional models. In this letter we propose to use an improvement of the projected autoregressive models (PAR) model [1], which has been depicted in figure 1.

The first order autoregressive (AR) model generates a series $g[n]$ as a linear function of a white gaussian noise and a delayed sample $g[n-k]$: $g[n] = a \cdot g[n-k] + b \cdot w[n]$. The parameters $a$ and $b$ are designed [1] in such a way that $g[n]$ adjust the mean, the variance and the autocorrelation $R_s[k]$ of the target series $v[n]$ for the elected lag $k$ which, for the case of a GOP video signal, is normally chosen to be 1. The PAR model proposes to project $g[n]$ onto its own distribution function $F_g(x)$ so that we will obtain a new series $u[n]$, uniformly distributed between 0 and 1. If we project $u[n]$ on the inverse distribution function $F_v(x)$ of $v[n]$ we will obtain a series $v'[n]$ which will perfectly adjust $F_v(x)$. The problem of this projection onto two non-linear functions ($F_g \circ F_v^{-1}$) is that the approximation of the autocorrelation that $g[n]$ performed is distorted, normally with a faster decay. To minimise this effect we propose a slight modification to the model, readjusting the parameter $a$ of the AR model. Multiplying $a$ by a factor $\eta$, defined as $\eta = \frac{R_s[k]}{R_v'[k]}$, where $R_s[k]$ is the autocorrelation of the actual signal for the lag $k$, and $R_v'[k]$ the distorted autocorrelation that the conventional PAR model would exhibit, we reinforce the autocorrelation of $g[n]$ for the lag $k$ so that, after being projected, the distortion of the autocorrelation will be compensated.

**SIMULATION AND RESULTS**

To prove the proposed scheme two different MPEG-I encoded sequences were used. The first one corresponds to the whole film “Star Wars”. The second sequence includes thirty minutes of the soccer World Cup 1994 final. Both traces were encoded with 24 frames per second and a GOP structure of twelve frames. The model was designed to distinguish between three different scenes types ($N=3$). Distribution functions were calculated using histograms of 40 levels. Results are presented for simulated queues with an utilization factor of 70%. In figures 2 and 3 the behaviour in a queue of the real traffic is compared with those of different models: a SRD model that does not consider scene changes ($N=1$) and the proposed model using both the conventional PAR model and the modified PAR as well as different windows sizes for the averaging filter. For the film “Star Wars”, figure 2 proves the accuracy of the modified PAR scheme to predict losses if the window size ($W=500$) is correctly designed. A model with a shorter window ($W=50$) or conventional PAR scheme and, of course, a SRD model underestimate the losses, especially when the queue size increases. On the other hand figure 3 proves that the proposed model is also able to adjust the behaviour of a video trace of a sport event, improving the adjustment that a SRD model performs. In this case a shorter window ($W=100$) is enough for a proper tuning of the model, as long as scenes are not so long as in a film.

**CONCLUSIONS**

A two level model for variable bit rate (VBR) video traffic has been presented. To cope with the existence of LRD within the video signal, the model considers two time scales: scenes and GOPs. To model the scene changes a Markov chain is used. For the GOP level a modification for the projected autoregressive (PAR) model is proposed so that the fitting of the autocorrelation function is improved. The ability of the model to adjust the behaviour of the real traffic in a queue is proved using two MPEG sequences with different characteristics. By this scheme the complexity of self-similar models and fractal calculations is avoided.
ACKNOWLEDGEMENTS

This work has been partially supported by the Spanish Comisión Interministerial de Ciencia y Tecnología (CICYT), Project No. TIC96-0743. We also wish to express our gratitude to M. Garret (Bellcore, USA) and O. Rose (Wurzburg University, Germany) for releasing the MPEG traces.

5. REFERENCES


FIGURE CAPTIONS:

Figure 1. Modified Projected Autorregresive (PAR) Scheme

Figure 2. Probability of bit loss for the trace “Star Wars”

Figure 3. Probability of bit loss for the sport trace.
\[
g(n) = a'g(n-k) + bw(n)
\]

\[
v'[n] = \left( F_g \circ F_v^{-1} \right)(g[n])
\]

Fig. 1
Fig. 2
Fig. 3